GAS RESPONSE TO A ROTATING STELLAR BAR

Mark Philip Schwarz

1979

A thesis submitted for the degree of
Doctor of Philosophy
to the Australian National University
Canberra.

To my mother,

and to Ray

The work reported in this thesis is that of the candidate alone except where acknowledged.

.....Philip Schwarz.....

June, 1979

ACKNOWLEDGEMENTS

It is a pleasure to thank Dr Agris Kalnajs for the invaluable assistance he gave during all phases of the work reported in this thesis.

I am grateful to Dr H. Visser for communicating results prior to publication and also to the staff at Mt Stromlo Observatory, particularly Dr G. Bosma, for valuable discussions.

I would like to thank the staff both at the Computer Services Centre, A.N.U., and the Computer Section at Mt Stromlo Observatory for their assistance, and I also much appreciate the work of Keith Smith in photographing the diagrams.

The support of a Commonwealth Postgraduate Research Scholarship is gratefully acknowledged.

ABSTRACT

We study the response of a gaseous disc to a rotating stellar bar. The gas is simulated by 2000 inelastic particles or "clouds", which are allowed to collide and in doing so can dissipate a large fraction of their relative motion. The response in the standard model with no inner Lindblad resonance (ILR) consists of a trailing spiral shock extending from corotation (CR) to just outside the outer Lindblad resonance (OLR), and can be interpreted as the result of the intersection of periodic orbits. The pattern evolves slowly towards a ring which is made up of a narrow range of periodic orbits belonging to one of two families. The ring at the OLR is usually aligned with the bar, but for weak bars can be at right angles. Recycling of the gas at a rate $1-2 M_{\odot} \text{ yr}^{-1}$ through star formation and stellar mass loss can inhibit ring formation and preserve the spiral structure indefinitely. The response at the ILR is similar: a ring forms aligned perpendicular to the bar axis, and gas can be forced into the centre on a timescale of a few bar rotations.

Circular orbits in a region between CR and the OLR become unstable even for quite moderate bar fields. Thus a collisionless sub-system, i.e. stars, will spread and form a hot disc with an exponential profile.

Our models bear remarkable resemblance to real galaxies with regard to both the "hydrogen" distribution and its velocity field. For example, radial velocity contour maps constructed for our models have kinks at the arms, as observed in the maps of M81. Also longitude-radial velocity maps

computed from the models have a striking similarity to those of the Galaxy, if the observer is placed near corotation.

The Cassini division in Saturn's rings occurs near an ILR. A simulation shows that, as in the galactic case, dissipation causes a gap to form on the outside of the resonance.

TABLE OF CONTENTS

	Page
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 COLLIDING PARTICLES IN AN EXTERNAL FIELD	8
2.1 - The Simulation	8
2.2 - Integration of Free Orbits	10
2.3 - Collisions	12
2.4 - The Potentials	16
CHAPTER 3 GAS RESPONSE TO A ROTATING STELLAR BAR - THE STANDARD MODEL	21
3.1 - Introduction	21
3.2 - The Standard Model	22
3.3 - Intersecting Rings	26
3.4 - Results and Discussion	30
3.5 - An Alternative Bar Field	40
3.6 - Gas Dynamics and the Collision Process	44
3.7 - The Initial Conditions	51
3.8 - The Effect of Stellar Mass Loss	56
CHAPTER 4 VARIATION OF THE BAR FIELD PARAMETERS	62
4.1 - Bar Strength	62
4.2 - A Strong Bar Field	68
4.3 - The Bar Rotation Rate	77
CHAPTER 5 NON-STANDARD MODELS	82
5.1 - Inverted Isochrone	82
5.2 - Hohl's Disc	88
5.3 - Spiral Forcing Field	96
5.4 - The Cassini Division in Saturn's Rings: Another Example of an Inner Lindblad Resonance	100

	Page
CHAPTER 6 VELOCITY FIELDS AND THE COMPARISON WITH OBSERVATION	108
6.1 - The Velocity Field of the Standard Model	108
6.2 - Torques exerted by the bar, and their Effects	110
6.3 - Rotation Curves	113
6.4 - Radial Velocity Maps	116
6.5 - Longitude-Radial Velocity Maps	120
CHAPTER 7 PERIODIC ORBITS IN A BAR FIELD	125
7.1 - A Second Order Theory	125
7.2 - The Outer Lindblad Resonance	128
7.3 - The Inner Lindblad Resonance	131
7.4 - Numerical Integration of Orbits	133
CHAPTER 8 CONCLUSIONS	138
APPENDIX	143
BIBLIOGRAPHY	148
ABSTRACT	iv

CHAPTER 1

INTRODUCTION

Spiral galaxies rotate differentially and consequently any material arms they might possess will wind up on a timescale of one galactic rotation, or about 10^8 years. This is much less than the age of the Galaxy (about 10^{10} years), so unless the patterns are continually rejuvenated, an alternative explanation must be sought. Lindblad (1948) was the first to introduce the idea that spiral structure was essentially a gravitational phenomenon: previously magnetic forces had been invoked to hold the arms together. Lindblad's main contribution was the notion of a density wave - a perturbation of the mass density which rotates as a solid body in the galactic disc, and through which stars and gas pass during their galactic orbits.

The subsequent progress in the field of spiral structure has been reviewed by Toomre (1977). The main thrust of this work has been to find long-lived non-axisymmetric oscillations of stellar discs (see for example Kalnajs (1971, 1978), Bardeen (1975) and Hunter (1972)). Self-consistent spiral modes have been found, but they are unstable and have large growth rates. Rather more successful have been the numerical simulations of stellar discs through large N-body calculations (Hohl, 1971; Miller (1976, 1979); Hohl and Zang, 1978). These studies, unlike the theoretical mode analyses, are of course capable of following the development of waves to finite amplitudes. Whilst the same unstable spiral waves form in the initial stages of the computer experiments, after a few revolutions they reach finite

amplitudes and end up as a steady rotating bar. It seems unlikely therefore that a long-lived spiral can be formed in a disc composed of stars alone.

This should hardly be surprising. It has long been recognized that spiral arms are associated with very young objects and gas - HII regions, neutral hydrogen and the young O and B stars recently formed out of the gas. Consequently it is important to consider how the dynamics of the gas differs from that of stars, and this is the main aim of the work described in the following chapters.

The essential difference between the motion of stars and that of gas is that gas can form shocks and dissipate energy. Since the early 1960's there had been suggestions that the dust lanes seen in the bars and along the arms of spiral galaxies might be due to shocks in the interstellar medium (Prendergast, 1962; Fujimoto, 1968). That shocks might occur is perhaps best illustrated by Kalnajs' (1973b) schematic picture of density waves as a superposition of elliptical orbits whose major axes shift in longitude as the radius of the ellipse changes. In the presence of a rotating non-axisymmetric perturbation varying with azimuth, θ , as $\cos 2\theta$, the periodic orbits in a disc potential are roughly ellipses whose eccentricities and orientations change rapidly near the principal resonances. If the perturbation field is sufficiently strong, or if the twist is tight enough, there will be a range of radii over which the orbits will cross just as in Kalnajs' picture. Now the ratio of pressure to gravitational force varies as the square of

the ratio of the sound speed ($\sim 8 \text{ km s}^{-1}$) to the typical non-circular velocity ($\sim 30 \text{ km s}^{-1}$). Hence the pressure force is small and gas flows as non-interacting particles except in the regions where their orbits cross. There we might expect shocks to form in the gas, and the resultant large density gradients will make pressure forces important.

The details of "galactic shocks" have been worked out in considerable detail for the case of a tightly wound spiral field, (Roberts, 1969; Shu, Milione and Roberts, 1973; and Woodward, 1975). In this case the time independent fluid dynamical equations governing the gas flow may be approximated by a one dimensional system. The well-known result of these integrations is a shock which follows the imposed spiral field. While the mathematical simplification associated with the tightly wrapped forcing field make it an attractive case to pursue, and will guarantee that any shocks that do form will be spiral, it is not the only way of generating spiral shocks. Several authors have pointed out (Kalnajs, 1973b; Sanders and Huntley, 1976) that bar-like density waves in the stellar disc will give rise to strong spiral density perturbations in the gas, chiefly around the Lindblad resonances. Kalnajs argued that gas would form shocks in these regions and conjectured that as a result of the ensuing dissipation, the shocks ought to be spiral. The main drawback of the bar-forced hypothesis is that the problem becomes two dimensional and time dependent. Nevertheless the bar-like distortion of the numerical experiments and the frequent occurrence of bar or oval distortions in flat galaxies

certainly suggests that it could be the relevant mechanism for spiral making, and that an effort should be made even if we could not hope to make a complete study in the time allocated for a Ph.D thesis. The aim is to obtain, not only a better understanding of gas flow in barred systems, but also insights into the problem of the origin of spiral structure itself.

Since this work was begun, some other studies of the gas response to a rotating bar have been published. Sørensen, Matsuda and Fujimoto (1976) used a fluid dynamical code with rather coarse resolution (a 30×50 Cartesian grid) to investigate the effect of a strong rotating bar. Despite numerical effects (a tendency of the gas to fall rapidly to the centre), they obtained shocks in roughly the positions in which dust lanes are observed along bars. Sanders and Huntley (1976) (also Huntley, Sanders and Roberts, 1977) have used the fluid dynamical code known as the "beam scheme" to integrate the flow of gas under the influence of a moderate oval distortion. They find that open spiral shocks develop at the ends of the bar, but are limited by numerical problems to a total integration time of only about two rotations of the bar. Finally Sanders (1977) has used a Lagrangian technique developed by Lucy for solving the hydrodynamic equations. The work demonstrated the importance of dissipation to spiral arm formation, but because of coarse resolution, shocks did not form, and the arms were very broad and indistinct. Also worth mentioning in this context are the early numerical experiments performed by Miller, Prendergast and Quirk (1970). Their self-gravitating N-body calculation

included a sub-system "gas", which was cooled at each step by reducing the local velocity dispersion. Although the calculations were numerically crude, they showed how a spiral pattern would last longer if dissipation were included. Miller et al. also demonstrated that it was the (bar-like) asymmetry in the "stars", (the non-dissipative sub-system) which in some way drove the spirality in the "gas".

The study described in the following chapters, pursued independently of those mentioned above, uses a technique by which the gaseous disc is simulated by a large number of particles (typically 2000). These represent gas clouds which can collide inelastically, losing a large fraction of their relative motion. This approach is useful not only because it enables a direct comparison between the stellar and gaseous response, but also because it should provide insights into the processes of galactic shock formation; for example the colliding orbits described above are directly calculable. The method is particularly appropriate to this problem since the low sound speed means that pressure forces will be small over most of the disc, and the gas streamlines will follow ballistic orbits except where these cross. This implies better resolution for a given amount of computing time compared with fluid dynamical codes. In addition there is the usual advantage of Lagrangian formulations, that less computation is needed in regions where the density is low.

Details of the simulation scheme are given in Chapter 2, and applied to a standard model in Chapter 3. The model has only an outer Lindblad resonance and the corotation

resonance. First we calculate the periodic orbits of collisionless particles in the bar potential and demonstrate the fact that they must cross. This is followed by a computation of the full model with and without collisions. Finally the effect of varying the collision parameters and initial conditions is investigated.

Chapter 4 studies the effects of varying the strength and pattern speed of the bar field. Further models are investigated in Chapter 5. The first uses an axisymmetric disc with a centrally peaked density distribution which has the effect of adding the inner Lindblad resonance. The second uses a stellar disc similar to that produced in the computer simulations of Hohl (1971). In the third the bar is replaced by a spiral field in order to determine if there is any major difference in the gas response to these two forms. The model was chosen to agree closely with the one used by Visser (1978) to model the galaxy M81. The chapter closes with a crude simulation of the rings of Saturn. Its relevance to the gas flow in a barred galaxy is that the Cassini division in the rings occurs near a resonance with the satellite Mimas. The frequency commensurability is just the same as that at the galactic inner Lindblad resonance. Now it has been suggested (for example Greenberg (1978)) that energy dissipation at the resonance is the cause of the gap. Simulations of the ring material by colliding particles moving in a central force field with a perturbation due to Mimas should resolve whether such a mechanism could work.

In Chapter 6, some comparisons are made with observations of gas distribution and flow in our Galaxy and in

external systems. Radial velocity contour maps are produced for the models, the effect of the arms on rotation curves discussed, and longitude-radial velocity ($l-v$) maps produced for comparison with those of our Galaxy.

Since we expect the gas streamlines to follow particle trajectories over most of the disc, we devote Chapter 7 to a study of periodic orbits in a bar field. Comparison is made between the predictions of the second order theory and numerically calculated orbits.

CHAPTER 2

COLLIDING PARTICLES IN AN EXTERNAL FIELD

2.1 *The Simulation*

All the models in this work are of infinitely thin discs. The hydrogen layer is indeed quite thin in our own galaxy: about 200 pc near the Sun, although the layer probably becomes much thicker at large distances from the centre. Furthermore, it is important to realize that the motion perpendicular to the plane is essentially uncoupled to that in the plane. In the case of gas, the z -motions are most likely dominated by the effects of supernova explosions and interstellar bubbles occurring more or less randomly over the disc (Cleary, 1977). Only in the outer parts of the disc, where the galactic potential might be more spherically symmetric (and the z -frequency and circular frequency thus more nearly equal), could we envisage a large-scale organized z -motion in the gas. For these reasons it seems reasonable to treat the gaseous disc as infinitesimally thin. The ensuing simplifications are of course enormous.

The other approximation which allows models to be calculated with moderate amounts of computer time on a medium-sized machine is the neglect of self-gravity of the gas. Certainly the stellar component is much more massive than the gaseous, and must dominate the axisymmetric force field. However the distortion of the stellar disc which constitutes the bar is only about 10% of the total surface density in most of the models calculated in the following chapters. Now this is of the same order as the gas content of spirals (5 - 10%),

so if there is an appreciable modulation of the gas density, it will produce a significant non-axisymmetric field. Thus if the gas response is in the form of a spiral, there will be a spiral component in the forcing field. In Chapter 5 we will see that the response to a spiral forcing field is similar in many respects to that produced by a bar, so that qualitative conclusions should be reliable despite the neglect of self-gravity.

Two approaches could be taken to simulate the dynamics of the gaseous disc of a spiral galaxy. The first is to use a hydrodynamic code to integrate the partial differential equations which govern the behaviour of a fluid. The second is to integrate the equations of motion of many particles which one considers to be fluid elements or gas clouds. This is the approach taken here.

The orbit approach is particularly applicable to this situation as the pressure forces can be neglected over most of the disc. This is because the ratio of pressure to gravitational force goes as $(v_s/v_r)^2$, where the sound speed, v_s , is only about 8 km s^{-1} compared with the non-circular velocities, v_r , of the order of 30 km s^{-1} . This implies that the streamlines of the gas flow are just the orbits of free particles in the gravitational field except where these cross. The shocks which would form in the highly supersonic gas flow at these points are simulated here by introducing collisions between the particles.

Furthermore, the fluid dynamic approach assumes that the interstellar medium can be represented by a smoothly varying gas with a known equation of state, whereas simulating the motions of many clouds might seem a more realistic way of

treating this "variously damped and heated multi-phase medium" (Toomre, 1977).

2.2 *Integration of Free Orbits*

Between collisions the orbits of the particles are integrated in the fixed stellar potential due to an assumed stellar distribution. This potential is composed of an axisymmetric part, Φ , and a part, ψ , which varies with angle, θ , around the disc as $\cos 2[\theta - \Omega_p t - \alpha(r)]$, i.e. which rotates with angular velocity Ω_p . In most of the cases considered, this non-axisymmetric part is in the form of a bar so that it may be written

$$\psi = \psi_0(r) \cos 2(\theta - \Omega_p t).$$

In a frame rotating with angular velocity Ω_p , the equations of motion are

$$\ddot{\tilde{\mathbf{r}}} = -\nabla(\Phi + \psi) + \Omega_p^2 \tilde{\mathbf{r}} + 2\dot{\tilde{\mathbf{r}}} \times \Omega_p,$$

which we integrate in Cartesian co-ordinates:

$$\ddot{x} = -\frac{\partial}{\partial x}(\Phi + \psi) + \Omega_p^2 x + 2\Omega_p \dot{y},$$

$$\ddot{y} = -\frac{\partial}{\partial y}(\Phi + \psi) + \Omega_p^2 y - 2\Omega_p \dot{x}.$$

These equations are integrated using a fourth-order Runge-Kutta scheme given in Abramowitz and Stegun (1964), p.897. With the notation

$$\tilde{\mathbf{F}}(\tilde{\mathbf{r}}, \dot{\tilde{\mathbf{r}}}) = (\ddot{x}, \ddot{y}),$$

the scheme becomes

$$\tilde{r}_{n+1} = \tilde{r}_n + h[\dot{\tilde{r}}_n + \frac{1}{6}(K_1 + K_2 + K_3)] + O(h^5)$$

$$\text{and } \dot{\tilde{r}}_{n+1} = \dot{\tilde{r}}_n + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4) + O(h^5),$$

$$\text{where } K_1 = hF(\tilde{r}_n, \dot{\tilde{r}}_n),$$

$$K_2 = hF(\tilde{r}_n + \frac{1}{2}h\dot{\tilde{r}}_n + \frac{1}{8}hK_1, \dot{\tilde{r}}_n + \frac{1}{2}K_1)$$

$$K_3 = hF(\tilde{r}_n + \frac{1}{2}h\dot{\tilde{r}}_n + \frac{1}{8}hK_1, \dot{\tilde{r}}_n + \frac{1}{2}K_2),$$

$$\text{and } K_4 = hF(\tilde{r}_n + h\dot{\tilde{r}}_n + \frac{1}{2}hK_3, \dot{\tilde{r}}_n + K_3)$$

The timestep, h , is the same for all particles, and is held constant during the integration. It is chosen so that there are between 30 and 100 steps per bar rotation period, the exact number depending on how fast the central rotation rate is compared with the pattern speed, Ω_p .

The accuracy obtained has been checked by integrating several orbits with both typical and extreme values of parameters, for timesteps h , $h/2$ and $h/4$. Assuming that the error is dominated by truncation error (rather than round-off error), we have used Richardson's deferred approach to the limit to estimate the error, and have also verified its h^5 dependence. In the integrations with collisions we should certainly satisfy the criterion that the cumulative error in position between collisions should not be greater than the collision box length (see Section 2.3). The probability that a particle does not

collide within two rotations of the bar is small over most of the disc. In all the models described in the following chapters, the accuracy is considerably better, and in some cases approaches the best obtainable within the limits of the fourth-order routine and the nine digit floating-point representation used. That is to say that for much smaller timesteps the round-off error would become important, and as is well known, this rapidly increases, dominating the error as the timestep is further decreased.

Jacobi's integral,

$$J = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}\Omega_p^2(x^2 + y^2) + \Phi(r) + \psi(r, \theta)$$

which takes the place of an energy integral in a rotating frame of reference, serves as a useful monitor of the accuracy of an integration. Of course, the Jacobi integral varies as the bar is turned on, but after this is completed, it is constant to better than 1 part in 3×10^4 over ten bar rotations for the crudest timestep used (about 30 steps per bar revolution). This is sufficient to ensure the reliability of individual orbits in the collisionless case. Round-off error becomes important when the step size is reduced by more than a factor of four below this.

2.3 Collisions

In order for an ensemble of particles to behave like a gas, the trajectories must not cross. To prevent this we simulate inelastic collisions between particles. Rather than continuously check the distance between individual particles, we use the less time consuming procedure described below, in which particles interact only when in the same box of a fixed

Cartesian grid. They may then be thought of as gas clouds with an effective size $\sim R_{\text{COLL}}$, the length of the box side. *The hope is that the structure and evolution of the model will not depend critically on the exact details of the collision but only on gross parameters such as inelasticity and cross-section.*

This approach to gas dynamics has precedent in other applications. The first simulations with small numbers of particles (30-50) moving in three dimensions were carried out by Ulam (1968) in the study of the nuclei of galaxies. More sophisticated collision procedures and larger numbers of particles (about 100) were used by Trulsen (1972 a, b, c) in studying the dynamics of jet streams and of Saturn's rings. More recently Brahic (1975, 1976, 1977) has made a systematic study of the flattening of a similar system of colliding particles in an external gravitational field, and has pointed out the relevance to the formation of Saturn's rings and the collapse of a proto-galaxy. In addition to the expected flattening, Brahic found that inelastic collisions tend to spread out the particles, extending the disc which has formed both inwards and outwards. On the contrary, Trulsen found that radial focussing occurred if the collisions were sufficiently inelastic, and a particular type of collision model (the "snowflake model") was used.

Lin and Pringle (1976) employed a similar scheme to study mass transfer and accretion disc flow in binary systems. Their simulation, involving about 2700 particles moving in a plane, demonstrated the role of a viscous interaction in the accretion of gas onto the primary. The ring which forms in the

absence of dissipation, spreads out in the radial direction, allowing gas to spiral into the primary. In later work, Lin and Papaloizou (1979) actually obtained a spiral near the 2:1 resonance around the primary this being just the equivalent of the inner Lindblad resonance in disc galaxies.

Most recently, Larson (1978a, b) has simulated collapse and fragmentation of a gas cloud by a system of about 100 self-gravitating particles. Unlike the above studies, Larson simulates pressure and viscosity separately, and both by means of separation-dependent forces rather than by collisions. However, as in the other calculations, dissipation plays a crucial role, in this case by allowing strongly bound condensations to form.

We describe here the standard procedure used to treat collisions in the modelling of later chapters. The effect of varying the various parameters is investigated in Chapter 3.

At each r -th step of the integration (where r is generally unity), a fixed Cartesian grid with box side $RCOLL$ is laid over the disc, and each particle is assigned to the appropriate box. If two particles are in the same box, then they are made to "collide". Only binary collisions are treated, so that if either $2N$ or $2N + 1$ particles occur in the same box, then N binary collisions take place, with no particle being involved in more than one collision. The algorithm used to compute which particles are in the same box is important, as this is a time consuming operation. The boxes are numbered with a

single integer co-ordinate, and the box number associated with each particle is calculated and stored. The particles are then sorted in order of increasing box number. The sort is quite fast when the efficient sort-merge-sort algorithm is used, since the particle positions do not change greatly at each step and so the order is not completely randomized.

The standard procedure for dealing with each binary collision follows closely the approach of Brahic (1977). The component of relative velocity, V_r , along the line joining the particles is reversed and reduced in magnitude. The component, V_t , perpendicular to V_r (the grazing component in the picture of two spheres colliding) is also reduced. The components after collision, V_r' and V_t' are thus

$$V_r' = - (1 - f_r)V_r$$

$$\text{and } V_t' = (1 - f_t)V_t$$

In most of the simulations both f_r and f_t are equal to 0.8. The correspondence between f_r and f_t and destruction fractions of dilation and shear respectively is at best very crude, and calculation of equivalent viscous flow parameters has not been attempted. However by dimensional arguments, the "Reynolds number" must be of order 200.

It should be noted that collisions conserve (for small $RCOLL$) total angular momentum about the centre of the galaxy almost exactly (and in the case $f_t = 0$ exactly). However significant amounts of energy will be lost whenever the colliding particles have substantial relative motion. We assume that the

energy liberated is radiated away instantaneously: the radiation timescale in the real case is indeed much shorter than the dynamical timescale.

If the collision process is thought of in terms of spheres colliding, then they effectively have a distribution of radii ranging from zero to R_{COLL} . The radius that a particular particle takes on at any time is more or less random, depending entirely on how close it is to the other particle it is about to collide with in the box.

Finally one necessary refinement is to disallow collisions for which the relative velocity along the line joining the two particles is such that they are separating. Since this component, V_r , is reversed during a collision, this feature prevents two particles losing all their relative motion by many consecutive collisions and sticking together forever, i.e. after a collision the particles are allowed to separate without suffering further collisions although they may remain in the same collision box.

2.4 *The Potentials*

It is well known that the surface brightness of the discs of spiral galaxies decreases exponentially, and it has been argued (Freeman, 1970) that this implies a surface density distribution of the same kind. For the axisymmetric disc component we use the isochrone distribution which is roughly exponential over a limited range of radii. The potential is given by

$$\Phi = -1/[1 + (1 + r^2)^{\frac{1}{2}}]$$

The isochrone possesses some nice analytic properties and was first introduced in connection with globular clusters by Hénon (1959). (The name arose because all orbits with the same energy have the same period independent of angular momentum). The surface density of the flat disc which gives rise to the isochrone potential is

$$\mu(r) = (2\pi G r^3)^{-1} \{ \ln[r + (1 + r^2)^{\frac{1}{2}}] - r/(1 + r^2)^{\frac{1}{2}} \}$$

This looks roughly exponential near the maximum of the rotation curve, although for large r decreases more slowly (as $\ln(2r)/r^3$) with the result that the rotation curve also decreases more slowly outside the peak.

The circular angular velocity,

$$\begin{aligned} \Omega &= \left(\frac{d\Phi}{dr} / r \right)^{\frac{1}{2}} \\ &= \{ [1 + (1 + r^2)^{\frac{1}{2}}]^2 (1 + r^2)^{\frac{1}{2}} \}^{-\frac{1}{2}} \end{aligned}$$

is plotted in Figure 2.1 together with the Lindblad precession frequencies, $\Omega \pm \kappa/2$, where the epicyclic frequency, κ is given by

$$\kappa^2 = 4\Omega^2 + r \frac{d\Omega^2}{dr} = (1 + r^2)^{-\frac{3}{2}}$$

The inner and outer Lindblad resonances (ILR and OLR) occur where $\Omega \pm \kappa/2 = \Omega_p$ respectively; that is where the forcing frequency due to the bar as seen by a particle in a circular orbit equals its natural radial frequency, κ . The

FIGURE 2.1

The angular frequency, Ω , and Lindblad precession frequencies, $\Omega \pm \kappa/2$, as a function of radius for the isochrone mass distribution.

corotation resonance (CR) occurs where $\Omega = \Omega_p$.

The angular velocity, Ω , in the isochrone is quite flat near the centre and affects the possible resonances in the presence of a rotating bar. In particular no ILR exists except for low rotation rates, Ω_p , of the bar. This is useful as it allows us to study first the simpler situation of models with a CR and OLR only. To simulate the gas response near an ILR, we use an axisymmetric disc whose density is peaked near the centre. The distribution, which we will call the inverted isochrone because it is so obtained, is described in Chapter 5.

The non-axisymmetric part of the potential used in the models is generally in the form of a bar or oval distortion.

$$\psi = \psi_0(r) \cos 2\theta$$

Two specific forms of $\psi_0(r)$ are referred to in the following chapters. The first is the $k = 3$ bar ($n = 0$) from the complete set of biorthonormal surface density-potential pairs generated by Kalnajs (1976):

$$\psi(r, \theta) = -0.53q[1 - 2.5(r/a)^2 + 2.1875(r/a)^4 - 0.65625(r/a)^6] \cos 2\theta, \quad r < a$$

$$2\pi G\sigma(r, \theta) \begin{cases} = 2.2613q(r^2/a^3) [1 - (r/a)^2]^{\frac{5}{2}} \cos 2\theta, & r < a, \\ = 0 & , \quad r > a \end{cases}$$

where q defines the strength and a the scale length of the bar. For computational purposes, the potential outside $r = a$ is approximated by an inverse power law with continuity of ψ and

and its derivative being satisfied at the boundary.

The detailed nature of the non-axisymmetric parts of galactic discs is unknown, even for galaxies with a strong bar. We use a second bar potential in order to check that our results are not peculiar to the above bar. Henceforth called the "deformed isochrone" bar, it is derived from the isochrone by stretching it along one axis by a factor $(1 + \epsilon)$ and squeezing it along the orthogonal direction. Because of the extended mass distribution, such a deformation gives rise to r^{-2} terms in the potential. These are removed by the application of the operator $(1 + \frac{1}{2}r\frac{\partial}{\partial r})$ which enhances the field near the centre and for large r the potential decreases like a quadrupole (as r^{-3}):

$$\psi_0(r) = -\epsilon\{3(r/a)^2 [1 + (r/a)^2]^{-\frac{5}{2}}\}$$

The density distribution of this bar is

$$2\pi G\sigma(r, \theta) = 15\epsilon(r^2/a^3) [1 + (r/a)^2]^{-\frac{7}{2}} \cos 2\theta.$$

It must be stressed that this is merely a convenient analytic form and while one can provide a physical description for generating it in terms of infinitesimal deformations, we do not wish to suggest that this is how the bar comes into existence.

Like the deformed isochrone, the Kalnajs bar decreases like a quadrupole at large r , but it differs in that it falls off faster just outside the bar maximum.

CHAPTER 3

GAS RESPONSE TO A ROTATING STELLAR BAR - THE STANDARD MODEL

3.1 *Introduction*

Spiral galaxies, and in particular barred systems, are sufficiently similar in form that it is not unreasonable, as an initial attempt, to model a typical galaxy. As has long been recognized, two-armed spirals predominate, and rings or pseudo-ring (broken ring) structures are very common especially amongst barred spirals. Even when the arms are not traceable continuously over large radial ranges, there is often a remarkable bisymmetry. It is this grand design which suggests that large-scale spiral structure results from an underlying bisymmetry of the stellar disc, rather than from a random process propagating through the interstellar medium. The latter, when combined with differential rotation and a suitable time-scale, does produce suggestive short spiral filaments (Gerola and Seiden, 1978), but lacks the large-scale bisymmetry.

In this chapter we construct a "standard model" using the following constraints as a guide in choosing the relevant parameters:

- (1) In observations of barred galaxies, a maximum variation of about one magnitude in luminosity around a circle is common;
- (2) The unstable modes for the isochrone (Kalnajs, 1978) suggest that corotation occurs near the peak of the rotation curve, and the bar potential maximum somewhat inside this radius; and

- (3) In the numerical simulations of stable bars (Hohl, 1971), the density maximum is inside corotation though at a comparable radius.

The parameters we choose lead to a model with no inner Lindblad resonance (ILR). This arises naturally out of the use of the isochrone axisymmetric mass distribution as pointed out in Section 2.4, and allows us to concentrate on the region between corotation (CR) and the outer Lindblad resonance (OLR) which has received little attention in the past. Indeed some authors have proposed that spiral structure is unlikely to exist outside the CR radius (see e.g. Roberts et al. (1975)), and that the outermost H II region or the edge of the prominent spiral structure or the easily visible disc could be used as a marker for this radius.

3.2 *The Standard Model*

The axisymmetric mass distribution used in the standard case is the isochrone, upon which the Kalnajs bar field (see Section 2.4) is superimposed. The numerical simulations of stellar systems (e.g. Hohl, 1971) show that the bar which forms, rotates so that it ends roughly co-rotate with the disc material. With this in mind, the bar field is stretched so that the potential maximum occurs on CR which is placed at $r = 1.5$ (in the natural isochrone units). The associated angular rotation rate, $\Omega_p = .2657$, is probably an overestimate; CR occurs at more than twice the radius of the peak of the rotation curve in Hohl's models. (Models with lower pattern speeds will be presented in Chapter 4). Fitting the rotation curve to that of the galaxy, the rotation rate corresponds very roughly to $25 \text{ km s}^{-1} \text{ kpc}^{-1}$.

The strength of the bar field is $q = 0.1$. The contours of total stellar surface density are shown in Figure 3.1, and it is clear that the deformation from axisymmetry is quite moderate: the maximum density contrast around a circle is about 0.25 magnitudes. Galaxies classified SB would mostly have rather stronger bars. Indeed this slight ellipticity would be indistinguishable in the central parts from the effects of projection on the plane of the sky, were it not for the outer isophotes. The ratios of maximum radial and azimuthal bar forces to the axisymmetric radial force at the principal resonances are given below:

r	arizmuthal	radial
1.5 (CR)	0.090	0.0
2.6 (OLR)	0.0094	0.043

A 10% perturbation field was used, as the noncircular motions in external galaxies are typically of this order.

The non-axisymmetric field is turned on linearly over two rotations of the bar. This is slow compared to all the characteristic timescales which are simple linear combinations of $\Omega - \Omega_p$ and κ , except near a resonance where one of them vanishes.

Since we do not expect the initial distribution of particles to affect the final pattern nor velocity field greatly, we use a simple form, viz. constant surface density in the central part tapering linearly to the edge. For reasons which will become clear later, two initial distributions are used: in model A the density tapers linearly from radius 1.9 to 2.6, in model B from

FIGURE 3.1

Contours of total stellar surface density for the isochrone mass distribution with the Kalnajs bar of strength $q = 0.1$ superimposed. The contour interval is 0.25 in $\ln$ (surface density), and the total length of each side of the border is 7.0 units. The bar lies along the x-axis.

2.1 to 3.3. Because the force field and initial conditions are symmetric with respect to 180° rotation about the centre, the resultant model will also have this property. Thus it is only necessary to load one half of the plane (we use about 2000 particles); the other half is obtained from the symmetry.

The standard binary collision procedure described in Section 2.3 is used, with a fraction, 0.8, of both components of relative velocity being destroyed in each collision. Such a large fraction is to be expected because of the highly supersonic nature of the flow. We may test for collisions at each timestep since the danger of the same two particles colliding in consecutive timesteps is all but removed by reversing the relative velocity component along the line joining the particles on the first collision, and allowing collisions between approaching particles only. The collision box size is set such that each particle suffers one collision on average per passage through an arm. $RCOLL = 0.035$ fits this criterion. It should be noted that this is about one-hundredth of the radius of the model, and so should allow sharp resolution of any narrow arms.

For comparison, model C is computed in which collisions between particles have been suppressed. This gives essentially the stellar response to a bar field, since this component of a galaxy is collisionless. In all other respects models C and A are the same.

3.3 *Intersecting Rings*

Before presenting the results of the full model calculations we examine the behaviour of initially filled circular orbits of Kalnajs (1973). The particles are arranged so that the ring separation is greater than the particle spacing along the ring. In this way, we can follow the deformation of the circular orbits as the bar field is turned on. Away from resonances, each ring behaves in an adiabatic manner: the initially singly periodic orbits remain singly periodic as the bar field grows, provided the rate is sufficiently slow. The stellar potentials are as given in the previous section, but collisions are suppressed. A similar procedure has been followed by Lindblad (1960), who used a dispersion ring rather than a bar. His Figure 18 particularly, demonstrates the same behaviour as those obtained in this section, especially with regard to the particles near corotation.

Figure 3.2 shows snapshots of the particle distributions at intervals of one bar revolution. The deformation of the rings into distorted ellipses occurs as the bar is being turned on. Eventually the ellipses cross neighbouring rings since their eccentricity and orientation change with radius, particularly near the OLR. The situation rapidly becomes quite confused near CR, since some particles move forward and others backward in the frame rotating with the bar. Eventually most of the particles there phase mix and all trace of the original rings is lost.

Figure 3.2

Intersecting rings of particles for the standard model with $q = 0.1$. The time in bar rotations is given in the lower right hand corner of each frame, and the axes extend to ± 4 natural isochrone units. The radii of CR and OLR are indicated by the arrows, and the bar lies along the x-axis.

↑ c o ,

2

3

4

5

Mention must be made of the orientation of the ellipses. As the bar is being turned on, they are aligned at an angle to the bar axis controlled by the growth rate of the bar. When the bar is at full strength the ellipses are at either 0° or 90° (or wobble about one of these angles). The reasons for this will become clear in Chapter 7. It is the crossing of these orbits which fuels expectations of shocks in gas forced by a bar.

Now dissipation or viscosity has a similar effect on the orbits as does a growing field; see Sanders and Huntley (1977) for a simple example treated with the epicyclic theory. So under the influence of (inelastic) collisions, we would expect the "orbits" to lead or lag the bar field by some angle just as in the growing field (e.g. the snapshot after 2 bar rotations). Under such circumstances the bar must exert a net torque on the ring, and this implies that there will be secular effects in gas forced by a bar.

Figure 3.3 shows a series of snapshots with the bar twice as strong as the previous one. This is included with the hope that it will whet the reader's appetite. Here the behaviour of at least two of the rings is quite different from the above case. The orbits between CR and the OLR become unstable. Needless to say, the integration scheme has been checked very carefully for programming errors or numerical problems. The behaviour of the offending particles is virtually independent of timestep (provided it is sufficiently small). Further discussion of this phenomenon is postponed to Section 4.2.

FIGURE 3.3

As for Figure 3.2, with $q = 0.2$.

↑ c 0 1

2

3

4

5

3.4 *Results and Discussion*

A time sequence of particle distributions for the collisionless case, model C, is shown in Figure 3.4. The spiral structure which appears at about two bar revolutions is due to the phase difference of orbit orientations in a growing field mentioned in the previous section. After 10 bar rotations the distribution is quite symmetrical about both minor and major axes of the bar; it displays no spirality. This is exactly what we expect from a system in which there is no dissipation e.g. a stellar system. After the initial transient, the stars settle down into a quasi-periodic motion in which they neither gain nor lose any net angular momentum from or to the bar.

The patterns which persist near the stable points at CR (i.e. the points 90° from the bar axis) indicate that particles have been trapped around these points, a phenomenon upon which much has been written (Contopoulos, 1973). An associated problem which is not so well understood is the density response at CR. In model C, the density just outside the stable points is about half that just outside the unstable points (the on-axis points). Inside CR the difference is rather less. However this is just the opposite of the prediction of Barbanis (1970) who argued for a density maximum at the stable point for a spiral field turned on suddenly. We must remember that the final distribution around resonances depends not only on the applied fields and the initial distribution, but because the behaviour is not adiabatic, also on the way the perturbation is turned on. These last two factors should be less important

FIGURE 3.4

Time sequence of distributions for the collisionless model C. The time in bar rotations is given in the lower right hand corner of each frame, and the axes extend to ± 4.5 units. The bar lies along the x-axis.

2
C O
↑

4

6

8

10

18

in models A and B because of the momentum exchange which occurs during collisions.

Figure 3.5 shows a sequence of distributions obtained for model A. As anticipated, the dissipation introduced through collisions has altered the evolution so that spirality remains for many rotations of the bar. The pattern after 5 bar revolutions is dominated by a basically two-armed spiral which extends from CR to outside the OLR. The arms are quite sharp indicating that a "shock" has formed in the gas. A "shock" in the sense used here is a sharp increase in the particle density at which considerable energy dissipation occurs by means of collisions. The second frame of Figure 3.6 plots the positions at which collisions occur (at time 5 bar rotations), and as one might expect, since collisions necessarily occur where the density is highest, they trace out a narrow spiral. Furthermore if we plot only those collisions in which more than 0.0002 and 0.001 energy units respectively are lost we obtain the lower frames of Figure 3.6. (Positions from several time-step have been superimposed to improve the statistics). These collisions account for 96% and 80% respectively of the total energy loss at each timestep. Here most of the collisions inside CR disappear, and what remains defines the narrow region in which large changes in velocity occur, i.e. shocks in the usual sense. Since we expect star formation to occur in these regions, collision plots give a better representation of the expected appearance of the *optical* arms. Further discussion of the velocity fields is given in Chapter 6.

FIGURE 3.5

As for Figure 3.4, but for the standard model A.

↑ c ↑

2

3

4

5

7

10

FIGURE 3.6

Particle distribution for the standard model A after 5 bar rotations (top left hand frame); plot of all collision positions for the same model (top right hand frame); and plots of positions of collisions in which the energy dissipated is greater than E_c , where E_c is given the lower right hand corner (bottom frames). The bar lies along the x-axis.

↑
C ↑
O

0.

.0002

.001

In addition to the two-armed spiral, transitory spiral arcs occasionally form. The particles in the central region where there are no resonances, remain in collisionless periodic elliptical orbits aligned with the bar major axis.

The pitch angle of the arms near the OLR tends to 0° , so they appear to turn back towards the centre. This is a tendency often found in the outer parts of spiral galaxies, and in particular barred systems, where the effect sometimes gives the impression of an outer ring (de Vaucouleurs, 1959). NGC 4151 is a good example of a galaxy with such a "pseudo-ring". This turning-back effect goes by the name of the "Stock effect", and has been interpreted as an artifact of projection (Stock, 1955). Here it is real.

There are many features in the particle plots for which one can find the counterpart in a real galaxy; for example the way the arms continue inwards past the ends of the bar near CR is reminiscent of NGC 4151. We point out one feature though, which appears in almost all the simulations and (together with the outer ring discussed below) is the one which appears most often in external systems. This is the depopulated region which lies just upstream of the stable point at CR. As the bar becomes stronger (Chapter 4), the region becomes larger and takes on the same appearance as the holes which are common to almost all SB galaxies, the spaces in the "0" shape. The same phenomenon can be seen in the outer parts of many galaxies classified as normal spirals (e.g. M100) and is associated with the tendency of the pitch angle of arms to decrease in the outer parts.

The time evolution of the pattern is dominated by an outward flux of particles from the region between CR and OLR, to the OLR. While the evolution is evident in the density plots, it becomes more striking if we examine the angular momentum distributions which reflect more accurately the frequencies of the particles, rather than the spatial locations. The comparison of the histograms for model A (Figure 3.8) with those for model C (Figure 3.7) demonstrate the efficiency with which collisions move angular momentum to the OLR. In model C there is some transfer but the effect is miniscule in comparison. Figure 3.5 shows that the particles at the OLR form into a tightly wound ring aligned at right angles to the bar major axis. [We use the word "ring" despite the ellipticity of the structure as this is the term used to describe such features in real galaxies. There, it is not immediately clear whether the "rings" are circular or not because of projection effects]. Since the particles in the ring have orbital angular momenta less than that of circular orbits at the OLR, we will refer to the ring as being "inside the OLR".

The situation is complicated by the results of model B, shown in Figure 3.9. Here the ring which forms is aligned along the bar major axis, and inspection of the histograms in Figure 3.10 reveals that the angular momenta of ring particles is greater than that of circular orbits at the OLR. It will be shown in Chapter 7 that the periodic orbits outside the OLR are aligned with the bar whereas those inside are generally aligned at right angles. Presumably then, the ring particles move

FIGURE 3.7

Histograms of number of particles against momentum for the collisionless model C. Arrows indicate the momenta of circular orbits at CR and the OLR. Each filled box corresponds to 6 particles. The time is given in bar rotations above each frame.

FIGURE 3.8

As for Figure 3.7, for the standard model A.

FIGURE 3.9

As for Figure 3.4, for the standard model B.

↑ c ↑
1

2

3

4

5

10

FIGURE 3.10

As for Figure 3.7, for the standard model B.

roughly in one of these sets of orbits, the other being depleted by collisions. The two types of ring never co-exist, as they would intersect (compare Figures 3.5 and 3.9). Which of these orbits is populated depends, in this case, on the initial conditions, the outer orbits being preferred whenever particles are loaded outside the OLR.

The timescale for ring formation is short compared to the age of the Galaxy (50-100 rotations). Since all spirals do not possess outer rings, the models must be inadequate in at least one of the following ways:

- (1) the stellar distribution (and in particular the non-axisymmetric component) might not be a good representation of that in spiral galaxies;
- (2) the collision process might not simulate the gas dynamics accurately; or
- (3) some other physical process not modelled changes the gas distribution at a rate comparable to that of ring formation.

3.5 *An Alternative Bar Field*

The effect of varying bar field parameters will be described in Chapter 4. Here we describe a model in which the Kalnajs bar has been replaced by a deformed isochrone. This field, which decreases at a slower rate outside the bar maximum, was used because it was thought that the depopulation near CR might be due to too high a perturbation field strength

there. A bar potential which fell off more slowly would allow a weaker field to be applied at CR whilst preserving the field strength at the OLR.

Contours of the total surface density for a field $\epsilon = 0.006$ are shown in Figure 3.11. Echoing the slower decrease in potential, the eccentricity of the contours decreases more slowly outside CR than for the Kalnajs bar. Furthermore the contours near the centre are more eccentric. Both these effects give the deformed isochrone model a "smoother" appearance. The maximum density contrast around a circle is only about 0.13 magnitudes. The ratios of bar force to axisymmetric force at the resonances are given below:

r	azimuthal	radial
1.5 (CR)	0.073	0.0
2.6 (OLR)	0.047	0.032

The radial force at OLR is roughly as for the Kalnajs bar, but the azimuthal force is increased so as to maintain a reasonable field at CR.

The depopulation of the region near CR in the particle distributions (Figure 3.12) is no longer as rapid. In particular, the particles near the stable Lagrangian point now appear to be trapped despite collisions. However the shocks near CR are rather weaker than in the Kalnajs bar. The ultimate fate of the model is familiar: a broad ring outside the OLR with a depleted region (albeit narrower) just inside the resonance. The ring is aligned with the bar major axis independent of the initial density

FIGURE 3.11

As for Figure 3.1, for the deformed isochrone bar field, $\epsilon = 0.006$.

FIGURE 3.12

As for Figure 3.4, for the deformed isochrone bar field, $\varepsilon = 0.006$.

↑ C ↑

2

4

6

8

10

12

distribution. As will become clear later, this difference is due to the higher field strength at the OLR which alters the possible angular momenta of stable periodic orbits of a particular orientation near the OLR. An increase in the field decreases the maximum angular momenta of stable periodic orbits aligned perpendicular to the bar major axis. A general conclusion which I draw from the several dozen models not discussed, is that rings tend to form outside the OLR and aligned with the bar major axis. Those which form perpendicular to the bar tend to be less robust and are easily disrupted if there is just a little gas outside the OLR. This is in disagreement with Kormendy's (1979) conclusion that most outer rings (R) are aligned at right angles to the bar axis. Accurate deprojection is impossible, so to obtain his result he assumes that inner rings (r) are all circular. But this is surely an unlikely circumstance, if only because in my experience rings seem to form near those resonances where orbits attain maximum (local) eccentricity.

3.6 *Gas Dynamics and the Collision Process*

How well do inelastic collisions of particles simulate the gas dynamics? At least two separate questions must be answered before we can have much faith in the collision process. First, we would want to know how the amount of energy dissipated in collisions affected the result. The viscosity of the interstellar medium is not a quantity which is easily estimated, though we suspect that it is greater than simple molecular viscosity. It is probably associated with larger scale motions (turbulent viscosity) or with the random motions of the interstellar clouds. In any case we become suspicious when rings

form in simulations such as those described in the previous sections; after all, even in an axisymmetric disc if the viscosity were large enough, we might expect all the particles to condense into rings, or fall to the centre (Trulsen, 1972). Secondly, we need to know if the spiral pattern and its evolution depend on the exact way in which the disc is cooled. If the result depended critically on how the collisions were treated, we could not hope that our simple approach would be very realistic.

It turns out that the collision box size, $RCOLL$, is the main parameter determining the amount of energy dissipated at each timestep. A sequence of distributions at time 5 bar rotations, for models with increasing values of $RCOLL$ are displayed in Figure 3.13. As $RCOLL$ is decreased the pattern gradually becomes less distinct, eventually tending to the collisionless case. The collision frequency decreases roughly as $(RCOLL)^{1.7}$, more slowly than the squared relation expected for a random distribution. This is presumably due to saturation effects which also appear in the amount of energy dissipated per timestep. If we increase $RCOLL$ above the average spacing, 0.063, the arms become broader and more diffuse. Of course this is just what we expect for lower effective resolution, but the situation is aggravated by the fact that binary collisions alone are unable to reduce to zero the velocity dispersion in a box containing more than two particles. As the box size is increased not only do more boxes contain three or more particles, but the velocity dispersion of particles within the box also increases. The

FIGURE 3.13

The effect of varying RCOLL, which is given in the lower right hand corner of each frame. The particle distributions are for time 5 bar revolutions. The size of the collision box is shown for the largest value of RCOLL.

↑ c
↑ o

.01

.02

.04

.08

.10

.15

o

scheme of Lin and Pringle (1976), which treats the case of many particles in a box by destroying a fraction, α , of the local dilatation and shear was used for $\text{RCOLL} = 0.12$. However when α was set to 0.8, large-scale clumping occurred particularly near the OLR. The value of RCOLL in our simulation which gives the highest density contrast across an arm is between one-half and three-quarters of the average spacing: values on either side degrade the contrast.

The experiment with $\text{RCOLL} = 0.15$ has been followed a further 20 bar rotations. It is found that the migration of particles to the OLR is slow on this timescale though at the expense of course of broader arms (the first frame of Figure 3.14). Despite the diffuse nature of the particle distribution, collisions, and particularly those in which large amounts of energy are lost, still occur along a very narrow region. In the two lower frames of Figure 3.14, collisions are plotted only if more than 0.001 and 0.002 energy units respectively are lost, and these make up 49% and 30% in turn of the total dissipated at each timestep. The width is in fact limited by the resolution. Of course there is no direct observational evidence for sharp HI arms (although dust lanes are certainly suggestive), and narrow optical arms could certainly be produced by a concentration of the *energy loss* to these regions.

The frames in Figure 3.15 show models calculated with representative values of (f_r, f_t) after 5 bar rotations. The parameters must be less than unity if the particles are to separate after collision. It is difficult to see from the case

FIGURE 3.14

As for Figure 3.6, for the simulation with
RCOLL = 0.15 at time 25 bar rotations.

FIGURE 3.15

The effect of varying the parameters f_r and f_t , which are given in the lower right hand corner of each frame. The particle distributions are for time 5 bar rotations.

↑ C ↑ O .8, .8

1, 1.

.4, .4

.8, 0.

$f_r = f_t = 1$ in Figure 3.15 that particles have actually stuck together, especially as large-scale clumping does not occur immediately. However the record of energy dissipated and number of collisions per timestep suggests the following sequence of events. The energy dissipated per timestep is at first greater than for the $f_r = f_t = 0.8$ case as expected, but after about four bar rotations it decreases, presumably because of the smaller number of effectively independent particles. For $f_r = f_t = 0.4$, a hotter disc is produced, and the resultant arms are broader. The heating of the disc is due not only to the lower collision frequency, but also to the reversal of the relative velocity component, V_r , after collision. If the grazing component, V_r , is not reduced during the collision ($f_t = 0$), the results are similar to the standard case, although the arms are not quite as sharp. In summary then, a large fraction of the relative velocity must be destroyed if narrow arms are to form.

The time between testing for collisions has only a mild effect on the overall result. Of course, if it is increased sufficiently, the integration approaches the collisionless case, and the arms weaken and disappear as for decreasing $RCOLL$. That decreasing the testing timestep does not double the number of collisions implies that the velocity dispersion is so small that two nearby particles will move less than a distance $RCOLL$ relative to each other between testing; or expressed in simpler terms, we are not missing many collisions as a result of testing only at discrete intervals.

In order to answer the second question as to whether the result depends on exactly how we cool the disc, we have computed models using several different methods of calculating the velocity change on collision, all with essentially the same result. Here we mention one which is of particular relevance to ring formation: we suppress any change in the angular momenta of individual particles, destroying only their relative radial velocity. This may seem a silly way of treating collisions, but it does show that the ring formation process is not dependent solely on the angular momentum exchange which occurs in a collision. Indeed, we could have suspected this when Figure 3.8 showed that ring formation was associated with a net increase in angular momentum, whilst collisions very nearly conserve total angular momentum. The initial condition used for the model displayed in Figure 3.16 is the same as for model B. It seems that ring formation is a necessary consequence of sharp arms in our models.

3.7 *The Initial Conditions*

Initial conditions should not affect the long-term evolution of the gaseous discs modelled above. Collisions will result in mixing of particles in angular momentum, and large-scale non-circular motions induced by the bar field should result in a redistribution of mass density over the disc to some extent. However we have already seen how the initial loading of particles can affect the orientation of the ring which formed. We should check that the results are not dependent on other aspects of the initial condition.

FIGURE 3.16

As for Figure 3.4, for a collision process in which angular momenta of particles are individually conserved (see text)

↑
C O

2

4

5

6

8

10

As remarked earlier, the final population of the various orbits in the collisionless case will depend on how the bar is turned on. Although the effect should be less marked when collisions are simulated, we have computed a model in which the bar is turned on over 6 bar rotations as a check. The evolution (Figure 3.17) is similar to that of model A, with the timescale for ring formation merely lengthened by the extra turn-on-time.

The loading of particles in rings is useful in that it allows us to view the way in which periodic orbits deform. However it also means that the first collision each particle in a particular ring suffers will be atypical: it will be with a particle whose initial angular momentum differed by either zero or $\pm \Delta h$, where Δh is constant along a ring. Now this may be of no long-term consequence, but the evolution of the central region in the standard model is certainly dependent on the initial conditions, in that they are responsible for the complete absence of collisions. A model in which the particles are initially distributed randomly on the plane but with the same average density distribution as for model A, is reproduced in Figure 3.18. The most remarkable feature is the outward flux of particles from the central region. Although suggestive, this should not be immediately identified with the "holes" in the HI maps of external galaxies, as it is likely that, unlike the standard model, many galaxies have an ILR.

We also include in this section mention of a model in which the constraint on the disc to remain symmetric with respect

FIGURE 3.17

As for Figure 3.4, for a model in which the bar is turned on over 6 bar rotations.

↑ C ↑

2

4

8

10

12

16

FIGURE 3.18

As for Figure 3.4, for a model in which the particles are initially distributed randomly over the disc (see text).

↑ C
↑ O

2

4

5

6

8

10

180° rotation about the origin has been removed. The initial condition is a disc containing 3000 particles whose centre of mass has been displaced from that of the stellar disc by superimposing a $\cos\theta$ variation of 10% amplitude on the density. The point is to determine whether collisions amplify the skewness of the distribution. Figure 3.19 shows that this is not the case. In fact the initial disturbance is damped at most radii. Even where the $\cos\theta$ component grows slightly it is never by more than a factor of two, and it remains smaller than the 2 θ component.

3.8 *The Effect of Stellar Mass Loss*

The formation of an outer ring (and the depletion of the region near CR) in the models of Section 3.4 presents a problem in that all spiral galaxies do not possess such rings. One process which we have not simulated in our experiments but which will certainly affect the gas distribution is recycling through star formation and stellar mass loss. Star formation occurs mainly in the arm regions where the gas density is highest, and mass loss should occur over the whole stellar disc; this is just what is required to inhibit the formation of the ring.

The actual rate of mass loss from stars is not well determined from direct observation, though a rate of $1 M_\odot \text{ yr}^{-1}$ for the galaxy is not unreasonable (Pottasch, 1970). This agrees quite well with estimates based on the birth rate of stars in the range $1.5\text{--}4 M_\odot$: from our present knowledge of the end states of

FIGURE 3.19

As for Figure 3.4, for a model which is not constrained to be bisymmetric (see text).

↑
C ↑
O

2

4

6

stellar evolution, Salpeter (1977) estimates that about $1.5 M_{\odot} \text{ yr}^{-1}$ must be lost to the interstellar medium. Furthermore we expect the rate of mass loss to be approximately equal to the rate of star formation which is estimated variously to be about $1 M_{\odot} \text{ yr}^{-1}$ (Mezger, 1975), and $2 M_{\odot} \text{ yr}^{-1}$ (Salpeter, 1977). Over 10 bar revolutions of our model, or 2×10^9 yrs say, a mass of between 2×10^9 and $4 \times 10^9 M_{\odot}$ will thus be cycled through stars. This is a large fraction of the gas mass of the Galaxy, $4 \times 10^9 M_{\odot}$ (Gordon and Burton, 1976), and so could have a significant effect on the accumulation of gas in a ring.

To test whether this process could avoid the ring formation, we make the simplest assumptions possible. We remove a fixed number of particles at each timestep from the regions where the gas density is highest (in practice where there are three or more particles to a collision box), and return them randomly over the disc. If we take the recycled number per bar rotation to be 217 (roughly equivalent to $2 M_{\odot} \text{ yr}^{-1}$), then we obtain the results shown in Figure 3.20. The recycling clearly preserves the sharp inner spiral structure and inhibits the formation of the outer ring. Moreover, the plot of the collision positions after 55 bar revolutions (Figure 3.21) shows that the energy is dissipated along a narrow front in the same way as at much earlier times in the standard model. Removing collisions in which little energy is lost (as in the lower two frames of Figure 3.21) emphasizes this fact. Dissipation in collisions of the sort plotted in these lower pictures represents 96% and 82% respectively of the total.

FIGURE 3.20

As for Figure 3.4, for a model in which gas is recycled (see text).

↑ c ↑ o

5

10

20

30

40

50

FIGURE 3.21

As for Figure 3.6, for a model in which gas is recycled (see text). The distributions are for time 55 bar rotations.

0.

.0002

.001

Accounting for star formation and stellar mass loss is thus one possible way of overcoming the drift of gas to the OLR.

VARIATION OF THE BAR FIELD PARAMETERS

Despite the similarities, spiral galaxies display a rich variety of patterns, as a glance through the Hubble Atlas will prove. Even amongst the galaxies which are classified as barred, the apparent strength of the bars, and the pitch angle, extent and multiplicity of the arms vary remarkably. The Hubble classification (Sandage, 1961) and the more elaborate revised de Vaucouleurs (1963) classification attempt to divide galaxies into groups each described by a set of characteristics. As Kormendy (1979) has pointed out, because of the apparent independence of many of the characteristics, the number of groups required is large. Kormendy considers galaxies to be composed of a number of distinct components or building blocks.

In this chapter we investigate how the gas response depends on two fundamental parameters of the bar field, bar strength and pattern speed. In this way we hope to understand some of the characteristics used in morphological classification and how they relate to each other. By using this procedure, we do not mean to imply that in real galaxies, the properties of the bars are independent of the axisymmetric disc. On the contrary, they will depend entirely on the nature of the disc in which the distortion is embedded. However since this connection is not well understood we shall treat them as independent parameters.

4.1 *Bar Strength*

Figure 4.1 shows particle distributions after 6 bar rotations for models using the deformed isochrone bar (Section

FIGURE 4.1

The effect of varying the bar strength, ϵ , which is given in the lower right hand corner of each frame. The distributions are for time 6 bar rotations. The axes extend to ± 4.5 units and the bar lies along the axis.

2.4) at different strengths, $\epsilon = 0.002, 0.004, 0.006, 0.007,$ and 0.01 . The ratio of bar force to axisymmetric force can be obtained by simply scaling the $\epsilon = 0.006$ values given in Section 3.5. The range covered is roughly 3% to 14% at CR. The value of the collision box size RCOLL has been varied slightly from model to model to produce the sharpest arms. Two features which change noticeably with bar strength are the pitch angles of the arms and their width. The first of these has been determined for each model by fitting a logarithmic spiral

$$\theta = \cot i \ln(r/r_0)$$

to the main arc of the arm, and is plotted against bar strength in Figure 4.2. [We use the term "pitch angle" here as "inclination" has another meaning. The pitch angle, i , is just $(90^\circ - \mu)$, where μ is the characteristic angle used by Danver (1942)]. Although the pitch angle decreases at both ends of the arms (and quite rapidly at the outer edge) the logarithmic spiral is a good fit over most of their length. (The values of pitch angle are uncertain to $\frac{1}{2}^\circ$, depending on the exact length of arm used in the fit).

Danver (1942), in a study of 98 spirals, found that arms could be fitted by logarithmic spirals, and reported a similar decline of i in the outer parts. His average pitch angles for various Hubble types are reproduced below:

	Sb	Sc	SBb	SBc	all
i (deg)	14.8 ± 0.9	17.5 ± 0.6	14.8 ± 1.0	17.0 ± 2.3	16.6 ± 0.4
σ	5.9	6.4	5.0	8.3	6.1

FIGURE 4.2

Pitch angle, in degrees, plotted against bar strength, ϵ .

The arms are more open in later types, but the pitch angle shows no significant difference between SA and SB systems.

The range spanned in Figure 4.2 is comparable with Danver's, with the exception of some very open spiral arms. Furthermore the pitch angle is correlated with the bar strength in the sense that the weaker the bar, the more tightly wound the spiral arm. The weakest field gives rise to a spiral which is restricted to the OLR region. (It is not discernible until later than 6 bar revolutions, and even then the pitch angle determination is uncertain). For the other cases, the pitch of the arms just outside CR (where the arm crosses the bar major axis) is almost independent of bar strength. Most of the change occurs near the OLR where the larger pitch angles occur in models in which most of the particles are participating in the flow characterized by periodic orbits outside the OLR (i.e. aligned along the bar axis). Conversely the smaller pitch angles correspond to flow aligned perpendicular to the bar axis. In this regard, the jump which occurs in Figure 4.2 at $\epsilon \sim 0.0065$ is probably real.

The situation is clearly complex. The low field cases may be interpretable in terms of a 90° shift in the orientations of the streamlines across the region where periodic orbits intersect at a resonance. This is inapplicable though to moderate to strong field cases where the spiral extends across two resonance regions. The existence of two different families of periodic orbits appears to be the most important factor controlling the pitch angle near the OLR in these cases. If so, we might expect the pitch of the arms to depend on initial

loading and collision parameters as well as bar strength, since these factors also affect the distribution of particles between the two families. A comparison of models A and B in Chapter 3, and of the frames in Figure 3.13 (where RCOLL is varied) confirms this.

We remark that the lower pattern speed models of Section 4.3 and those of Chapter 5 suggest that pitch angle is sensitive to the relative positions of the resonances, and we suspect that the variation with Hubble type apparent in Denver's data reflects this dependence rather than a trend in the non-axisymmetric field strength.

As the bar strength is increased, the arms also become broader (even though we have varied RCOLL to obtain the sharpest arms possible). Presumably the collision scheme we have used is unable to reduce the larger velocity dispersions which result from the stronger bar to the same level as for the standard case. However quantities such as the thickness of arms depend on the effective equation of state used, and we cannot claim that our scheme exactly represents that of the interstellar gas.

Some other observations can be made regarding the models of Figure 4.1. The total angular winding of the arms shows little dependence on bar strength, except for the weakest field, in which case the spiral is restricted to the region near the OLR. We do not attempt to quantify what would be a subjective measure except to say that in general the total winding is greater than the 180° predicted on the basis of 90°

winding per resonance (Sanders and Huntley, 1976). Values for our models are closer to the data obtained by Danver (1942), i.e. $277^\circ \pm 17^\circ$ and $211^\circ \pm 24^\circ$ for SBb and SBc galaxies respectively.

Finally, though not shown, the character of the ring which eventually forms changes at about $\epsilon = 0.006$. For weaker bars the ring is aligned perpendicular to the bar; for stronger ones, along the bar. This is hardly surprising in view of the remarks we made when discussing the pitch angles. Also, like the pitch of the arms, the final orientation of the ring can be affected by initial loading and the collision parameters. That is to say, we can create a ring aligned along the bar axis for weak fields by varying these factors; we have never obtained a ring perpendicular to the axis for strong fields.

4.2 *A Strong Bar Field*

To study the response to a strong bar field in more detail we use the Kalnajs bar at twice the strength of that used in the standard model (Section 3.2). The ratios of bar to axisymmetric force can be obtained by doubling the values given in Section 3.2. Contours of equal surface density for the total mass distribution shown in Figure 4.3 are more representative of barred galaxies than is the mild distortion of the standard model. The maximum density contrast between points on the major and minor axes of the bar at the same radius is about 0.5 magnitudes.

Evolution of a collisionless disc in the field (Figure 4.4) highlights the phenomenon already briefly mentioned in

FIGURE 4.3

Contours of total stellar surface density for the isochrone mass distribution with the Kalnajs bar of strength $q = 0.2$ superimposed. The contour interval is 0.25 in $\ln$ (surface density), and the total length of each side of the border is 7.0 units. The bar lies along the x-axis.

FIGURE 4.4

Time sequence of particle distributions for the collisionless model using the strong bar field $q = 0.2$. The time in bar rotations is given in the lower right hand corner of each frame, and the axes extend to ± 4.5 units for the top frames and 9.0 units for the lower frames. Only particles within these dimensions are shown. The bar lies along the x-axis.

↑ C ↑ O

4

6

8

10

20

30

Section 3.3, where the integration of rings of particles suggested that some orbits between CR and the OLR were unstable. The result is a spreading of the disc reminiscent of that of the self-gravitating stellar discs of Hohl (1971) after the formation of a bar. In that case the result was an exponential disc,

$$\mu \propto e^{-\alpha r}$$

with a change of slope just inside CR. The length scale α^{-1} of the outer component was 8.6 kpc, that of the inner 1.44 kpc.

The azimuthally averaged density for our model is plotted against radius in Figure 4.5. Four distributions at times between 20 and 22 bar rotations are averaged to improve the statistical significance in the outer parts. The densities inside CR are largely dependent on the initial conditions and are not of much significance. The peak just inside the OLR is evidence of the ring clearly visible in Figure 4.5. The alignment perpendicular to the bar major axis is reminiscent of barred galaxies of early Hubble type. Outside, the fall-off in density is roughly exponential with scalelength 1.24 (or 8.3 kpc if we put CR at 10 kpc). The same spreading of the disc occurs under the influence of the weaker bar of the standard model, albeit more slowly, and the circles in Figure 4.5 give the average density profile at time 20-22 bar rotations. The scalelength in this case is 0.69 (4.6 kpc). For the strong bar, there is a slow increase in the angular momenta of the unstable orbits as can be seen in Figure 4.7. This is hardly detectable in the standard model (Figure 3.7).

FIGURE 4.5

Azimuthally averaged density for the standard bar, $q = 0.1$ bar (circles) and the strong bar, $q = 0.2$, (squares) plotted against radius. The units of density are arbitrary. Collisions have been suppressed.

Could it be that this is the origin of Freeman's (1970) exponential disc? If so, we can make a strong prediction: this component must be hot in so far as the eccentricities are large. Certainly if one is prepared to accept that there are oval distortions in the discs of most galaxies, then the process must certainly occur. Blackman (1979a, b) has found a double exponential component structure in the discs of two of the three galaxies for which he has published surface photometry. It is likely that the outer exponential has been formed by the process occurring in our experiments, since the boundary between the two components lies at the outer edge of the spiral structure where we expect the OLR to occur. Furthermore the scalelengths (normalized by the radius of the peak of the rotation curve) are comparable though larger than the values we obtain. The outer exponential in Hohl's experiments are almost certainly caused by the same process. The behaviour of the individual unstable orbits which produce the spreading is described in Chapter 7.

A model using the strong bar field in which collisions have been simulated is shown in Figure 4.6. As in the collisionless case, the turn-on time is 2 bar rotations, and the initial distribution is as for the standard model A (Section 3.2). The angular momentum distributions are given in Figure 4.8. The most striking feature is the rapid depopulation of the regions around the Lagrangian points L_4 , L_5 (i.e. the points at CR on the bar minor axis). Even in the collisionless case (Figure 4.4) these points are definitely density minima. In the Appendix, the condition is derived for which stable "banana" orbits exist at CR.

FIGURE 4.6

As for Figure 4.4, for the strong bar field model with collisions. The axes extend to ± 4.5 units and only particles within that area are plotted.

↑ c ↑ o

2

4

6

8

10

20

FIGURE 4.7

Histograms of number of particles against momentum for the collisionless case with the strong bar field, $q = 0.2$. Arrows indicate the momenta of circular orbits at CR and the OLR. Each filled box corresponds to 6 particles. The time is given in bar rotations above each frame.

FIGURE 4.8

As for Figure 4.7, but for the strong bar model with collisions.

It is

$$(K_1 + K_2 - 4\Omega_p^2)^2 - 4K_1 K_2 < 0,$$

$$\text{where } K_1 = \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} \bigg|_{L_4, L_5},$$

$$K_2 = - \left[\frac{\partial^2}{\partial r^2} (\Phi + \psi) \bigg|_{L_4, L_5} - \Omega_p^2 \right]$$

This condition in the regime of interest, effectively limits the curvature of the Jacobi integral at L_4, L_5 . Calculation of K_1 and K_2 for $q = 0.2$ (the strong bar) yields

$$(K_1 + K_2 - 4\Omega_p^2)^2 - 4K_1 K_2 = -0.407,$$

so that no stable orbits should exist at CR. In fact the critical value of q is 0.189, so that the L_4, L_5 points are stable for the standard case.

Also derived in the Appendix, are the energies associated with each of the normal modes of oscillation near CR. The energy of the lower frequency mode is negative definite, that is it decreases as the amplitude of the motion increases. This might explain why collisions are so effective at clearing out the region around the L_4, L_5 points even for weaker bars. In the process of removing energy, collisions can actually increase the amplitude of the lower frequency mode.

We stress that these results are derived from the linearized equations, and can be applied to orbits only while they remain near the hills of the potential in the rotating frame.

Finally we remark on the ultimate fate of gas in the model. The ring which forms outside the OLR is almost circular, and the region inside completely devoid of gas after 20 bar rotations. The stronger bar has moved the gas well beyond the OLR, and since the Kalnajs bar decreases rapidly, the periodic orbits near the ring radius have small eccentricity. Of course the depopulated region will be replenished with gas from stellar mass loss at a rate of $1-2 M_{\odot} \text{ yr}^{-1}$ as discussed in Section 3.7.

4.3 *The Bar Rotation Rate*

The pattern speed in external galaxies cannot be measured directly. The only method of estimating it is to compare easily identified density and velocity characteristics with similar features in numerical experiments such as those described in these pages. The rings which we have seen form at the OLR are but one example. It is with this aim in mind that we compare four models with different rotation rates.

In the numerical experiments of Hohl, the CR radius was more than twice the radius of the peak of the rotation curve, and by this criterion the pattern speed adopted for the standard model is high. We choose three lower rates, $\Omega_p = 0.1, 0.06,$ and 0.04 , for which CR is at $r = 3.0, 5.86$ and 8.0 respectively. The deformed isochrone bar is used, stretched so that the potential maximum is at the peak of the axisymmetric rotation curve, $r = 2.2$. Placing the bar maximum at CR leads to unreasonably large fields at the OLR. It would have been preferable if the ratios of bar to axisymmetric force at the resonances had been the same for each pattern speed, but this

is not possible as the radial and azimuthal components vary differently. The exact values are given below:

Ω_p	r	radial	azimuthal
0.14	3.0 (CR)	0.0033	0.092
	4.5 (OLR)	0.055	0.070
0.06	5.9 (CR)	0.047	0.043
	8.0 (OLR)	0.035	0.027
0.039	3.8 (ILR)	0.107	0.164
	8.0 (CR)	0.10	0.078
	11.0 (OLR)	0.061	0.044

The distributions in Figure 4.9 are at time 5 bar rotations. However care must be taken when making comparisons, as the inner parts of the models with lower pattern speeds will have rotated many more times, so their stage of evolution is not identical.

As the pattern speed decreases, the central regions gradually become more important. First we notice near CR a doubling of the arms, intimately connected with a ring-like structure which defines the outer edge of the undisturbed disc for the high pattern speed models. The ring itself does not lie at CR but slightly inside the ultraharmonic resonance where $(\Omega - \kappa/4) = \Omega_p$. We tentatively identify this feature with the inner rings (r) from which spiral arms are observed to emerge. NGC 2523 is the example usually quoted.

We digress for a moment to consider the work of Kormendy (1979) on such rings. On the basis of the distribution of

FIGURE 4.9

Effect of varying the pattern speed, Ω_p , which is given in the lower right hand corner of each frame. ILR, CR and OLR are indicated the arrows associated with the symbols I, C and O respectively. The axes extend to ± 4.5 , ± 9 , ± 13.5 and ± 18 units in turn. The bar lies along the x-axis in each case.

.27

.14

.06

.04

apparent axial ratios, he finds that these rings are intrinsically circular but with low statistical significance. From this he deduces that the outer rings are prolate (in the sense of being aligned perpendicular to the bar major axis). If however the inner rings were elliptical with eccentricity greater than that of the outer rings (as those in Figure 4.9 are), then the outer rings would be intrinsically oblate, as is more common in our experiments.

For even lower pattern speeds an ILR exists ($\Omega_p = 0.06$ is close to the critical value). Its effect is to clear an annulus by sweeping particles towards the centre, forming a small ring in the process. In the frame $\Omega_p = 0.14$ only a narrow region has been cleared, whilst for the most slowly rotating bar, most of the particles are in a lump at the centre.

We make the observation that the inner ring for the most slowly rotating bar is rather weaker than for the other models. In this connection it is instructive to compare the barred galaxies on pages 45 and 46 of the Hubble Atlas (variety (s) of SBb galaxy) with those on pages 47 and 48 (varieties (r) and (rs)). Not only do the former have less prominent (or absent) inner rings, but also much more dust lane structure within these rings. This is suggestive of an OLR in (s) varieties of SBb types, and its absence in (r) varieties, especially as the central regions of the former are dominated by a bright condensed nucleus. (The condition that an ILR exist is that the density be peaked near the centre). The evolution near an ILR is investigated more thoroughly in the next chapter.

Finally notice in the $\Omega_p = 0.04$ model, the short filaments or "hairs" in the region where the main outer arm turns back: the gas still seems to want to shock along trailing spiral segments. This occurs often when the bar field at the OLR is large (and fairly flat, unlike the Kalnajs bar). A similar phenomenon is observed in many barred and normal spirals, a striking example being seen in the print of NGC 1398 in the Hubble Atlas.

NON-STANDARD MODELS

5.1 *Inverted Isochrone*

In the case of some barred galaxies, particularly of the Seyfert type, there is a small spiral or ring near the centre of the bar. NGC 4151 provides an example of such inner structure. In the light of the results of Section 4.3, this is suggestive of an ILR. The isochrone distribution is not suitable for studying this behaviour because of its flat angular velocity near the centre. Here we introduce an axisymmetric disc which is highly peaked near the centre.

Starting from a known disc galaxy model, it is possible to generate another by inversion (Kalnajs, 1976). This process turns the given surface density $\sigma(r)$ into $\sigma(1/r)/r^3$. Because the surface density of the isochrone falls off as $|\ln r|/r^3$ for large r , the inverted density will have a mild singularity at the centre, and the inverted potential

$$\begin{aligned}\psi(1/r)/r &\equiv \psi(r) \\ &= -1/(r + \sqrt{1 + r^2}) \\ &= -\sqrt{1 + r^2} + r\end{aligned}$$

will have a finite slope at $r = 0$. Thus the angular velocity of circular orbits in the field,

$$\Omega = [r \sqrt{r^2 + 1} (r + \sqrt{r^2 + 1})]^{-1/2}$$

is singular at the origin. Although this is unrealistic, a steep rise is just what is needed to produce an ILR near the centre. Figure 5.1 plots the angular velocity and the Lindblad

FIGURE 5.1

The angular frequency, Ω , and Lindblad precession frequencies, $\Omega \pm \kappa/2$, as a function of radius for the inverted isochrone mass distribution.

precession frequencies $\Omega \pm \kappa/2$; the epicyclic frequency, κ , is given by

$$\kappa = \Omega [3 - r^2/(r^2 + 1) - r/\sqrt{r^2 + 1}]^{1/2}.$$

The pattern speed is chosen to be the highest which is compatible with a reasonably low ratio of the radii of OLR and ILR, viz.

$\Omega_p = 0.2$. This latter condition is merely a computational necessity if the regions near both OLR and ILR are to be simulated with any accuracy. The resonances occur at 3.02 (OLR), 2.2 (CR) and 0.88 (ILR).

The bar field used is the Kalnajs mode (Section 2.4) which is scaled so that the potential maximum occurs just inside CR ($r = 2.0$), and the maximum bar force is roughly 10% of the axisymmetric. The ratios at the resonances are tabulated below:

r	azimuthal	radial
0.88 (ILR)	0.266	0.226
2.2 (CR)	0.103	0.032
3.0 (OLR)	0.056	0.116

The particle distributions of Figure 5.2 show that the initial gas response is a spiral extending from inside the ILR to beyond the OLR. There is a break in the spiral near CR which appears as a doubling of the arm structure. The outer arm evolves into an eccentric ring inside CR, and the inner into a small spiral inside the ILR. A broad ring outside the OLR forms in the same way as in the simulations of Chapter 3. The clearing out of the interarm regions is almost complete, implying that the gas flow is entirely along the arms.

FIGURE 5.2

Time sequence of particle distributions for the inverted isochrone model. The time in bar rotations is given in the lower right hand corner of each frame, and the axes extend to ± 4.5 units. ILR, CR and OLR are indicated by arrows associated with the symbols I, C, and O respectively. The bar lies along the x-axis.

↑ ↑ ↑
I C O

2

3

5

7

The distribution of particles with angular momentum (Figure 5.3) is also instructive. The depletion of some angular momentum states occurs as particles are moved both beyond the OLR and towards the centre. The ILR appears to behave like an OLR "turned-inside-out". The ring inside CR is similar to those found for the isochrone models with low pattern speed (Section 4.3), and once again is probably connected with the second harmonic resonance ($r = 1.7$, $h = 8.24$) rather than with CR.

The ILR is extremely efficient at moving particles into the centre of the model. After 6 bar revolutions, all the particles initially inside the ILR are essentially in one lump at the centre. If $4 \times 10^9 M_{\odot}$ of gas is initially distributed uniformly over the disc $r < 3.0$, then this is roughly equivalent to moving $0.4 M_{\odot} \text{ yr}^{-1}$ (for one bar revolution = 1.5×10^8 yrs). There is insufficient resolution to examine the detailed structure at the centre, however before the lump forms, the gas forms a contracting ring in which large amounts of energy are dissipated. Could this be the origin of the Seyfert phenomenon? At least the process is in some sense self-sustaining, since as more mass is moved to the centre, the density distribution becomes more peaked, and this is just the condition for an ILR to exist. It should be noted that the bar field at the ILR is not abnormally strong, as the Kalnajs bar falls off rapidly inside its maximum. Finally, although the orientation of the ILR ring is at an oblique angle to the bar axis during its evolution, at later times the ring is aligned almost perpendicular to the bar.

FIGURE 5.3

Histograms of number of particles against momentum for the inverted isochrone model. Arrows indicate the momenta of circular orbits at the ILR, CR and OLR. Each filled box corresponds to 6 particles. The time is given in bar rotations above each frame.

5.2 *Hohl's Disc*

The exact form and strength of stellar bars in spiral galaxies is not well known. Accurate photometry is essential if the non-axisymmetric component of the disc potential is to be estimated. In addition the measurements must be taken in the red so as to minimize the contribution from ionized gas and young stars. Finally, methods of deprojection involve assumptions about the structure and dynamics of the galaxy, the usual practice being to make it appear as circular as possible. This has the effect of reducing the apparent strength of bar-like distortions.

In the absence of such information we have determined the field of one of the bars arising from a numerical simulation of a stellar disc. Hohl's (1971) flat discs are more convenient than Miller's (1979) three dimensional bars. The difference in density between major and minor axes of the bar derived from Hohl's Figure 11 is plotted against radius in Figure 5.4. It is assumed that the azimuthal variation is caused solely by a 2θ component (as suggested by Hohl's Figure 5). We also plot the density of the deformed isochrone (Section 2.4) suitably scaled in radius and strength. The fit is reasonably good, except perhaps near the centre, and in order to avoid the non-trivial task of calculating the forces from Hohl's data, we will use the deformed isochrone bar. Corotation in Hohl's model is at roughly five times the radius of the bar density maximum. At such radii we may assume that the potential decreases like that of a quadrupole (as $1/r^3$) as does the deformed isochrone. Local variations might of course occur, especially near a resonance.

FIGURE 5.4

Density difference between major and minor axes of Hohl's bar (squares), with that of an appropriately scaled version of the deformed isochrone bar (solid curve).

The axisymmetric disc formed in Hohl's experiment can be fitted very well with one of the higher- n Toomre disc models (Toomre, 1963). The first member of Toomre's family, with

$$2\pi G\sigma_1(r) = C_1^2(1 + r^2)^{-\frac{3}{2}},$$

$$\text{and } \phi_1(r) = -C_1^2(1 + r^2)^{-\frac{1}{2}}$$

is very similar to the isochrone. The densities of higher members fall off with increasing rapidity. In Figure 5.5 we plot the axisymmetric rotation curve derived from Hohl's Figure 12, $t = 8.0$, and also a suitably scaled version of Toomre's fourth model,

$$V_4^2(r) = 15r^2(1 + r^2)^{-\frac{3}{2}} + 27r^2(1 + r^2)^{-\frac{5}{2}} + 45r^2(1 + r^2)^{-\frac{7}{2}} + 105r^2(1 + r^2)^{-\frac{9}{2}}.$$

The corresponding density is

$$2\pi G \sigma_4(r) = 105(1 + r^2)^{-\frac{9}{2}}.$$

We remark that the Toomre disc does not fit the details of the axisymmetric density profile, but the important quantity to be reproduced by the model is the force field (or equivalently the rotation curve).

The angular velocity and Lindblad precession frequencies are plotted against radius in Figure 5.6. The corotation radius, 12.5 kpc, corresponding to $\Omega_p = 18.2 \text{ km s}^{-1} \text{ kpc}^{-1}$ becomes $r = 2$ in the natural Toomre model units. The principal resonances then occur at 2.56 (OLR), 2.0 (CR) and 1.41 and 0.41 (ILR). The separation between the resonances is rather smaller than in any of the isochrone models. The bar to axisymmetric force ratios at

FIGURE 5.5

The rotation curve of Hohl's disc (squares), with that of an appropriated scaled version of the Toomre $n = 4$ disc.

FIGURE 5.6

The angular frequency, Ω , and Lindblad precession frequencies, $\Omega \pm \kappa/2$, as a function of radius for the Toomre $n = 4$ disc.

the resonances are given below:

r	azimuthal	radial
1.41 (ILR)	0.039	0.052
2.0 (CR)	0.033	0.053
2.56 (OLR)	0.021	0.032

The particle distributions which result (Figure 5.7) present a rather different picture from the other models we have produced. The ring near the ultraharmonic resonance has almost completely disappeared and with it the bar-like nature of the gas response. Without the aid of the central lump, one would have difficulty identifying the bar axis. Furthermore the ring which forms at the ILR (in fact just inside as can be seen in the momentum distributions of Figure 5.8) is of a different kind from those in Chapter 4 and Section 5.1 which we identified with the inner rings (r) of SB galaxies. Whilst those were aligned along the bar major axis, the ILR ring is more nearly perpendicular. Associated with this, is the fact that the outer arms join the ring near the bar minor axis rather than the major axis as for the ultraharmonic resonance ring. In the case of the inverted isochrone model, an ILR ring formed but was swept into the centre on a timescale of a few bar revolutions.

One of the differences between this model and the others we have computed lies in the condensed nature of Hohl's bar. The bar maximum lies so close to the centre that the orientations of the periodic orbits at the ILR are reversed. We show in Chapter 7 that the behaviour at an ILR depends on the

FIGURE 5.7

As for Figure 5.2, for the model based on Hohl's disc. The axes extend to ± 4 units.

↑ ↑ ↑
↑ ↑ ↑
↑ ↑ ↑

1

2

3

4

5

7

FIGURE 5.8

As for Figure 5.3, for the model based on Hohl's disc

sign of the quantity

$$\Psi = \frac{1}{2} \frac{\partial \psi_0}{\partial r} + \frac{\psi_0}{r_0} \frac{2\Omega}{\kappa} ,$$

where ψ_0 is the radial part of the perturbation and r_0 the resonant radius. When Ψ is negative, the situation is just like that at an OLR "turned-inside-out", except that periodic orbits outside are aligned with the bar, those inside perpendicular.

If however the bar is centrally condensed so that $\frac{\partial \psi_0}{\partial r}$ is sufficiently large negative for Ψ to be positive at the ILR, the orientations are just reversed. Furthermore there is an extra 90° change in orientation of periodic orbits outside the ILR where $\frac{\partial \psi}{\partial r} = -\frac{\Omega}{\Omega - \Omega_p} \frac{2\psi_0}{r}$ according to the epicyclic theory. This reversal occurs at zero eccentricity unlike the situation at a resonance.

We conclude this section by asking whether at least some galaxies classified as normal spirals could have central bars driving the spiral pattern. The Hohl model, in which an essentially continuous spiral terminates in a ring at the ILR, does not easily fit into the SB classification.

5.3 *Spiral Forcing Field*

A considerable amount of work has been done investigating the gas response to a tightly wound spiral density wave of the sort studied by Lin and Shu (1964). Roberts (1969) integrated the time independent hydrodynamic equations in such a field, and the time development of the gas flow was calculated by Woodward (1975). This approach has been used to model a variety of

galaxies whose rotation curves are reasonably well determined (Roberts, Roberts and Shu, 1975). The most detailed comparison has been made by Visser (1977) with high resolution HI observations of M81 (Rots and Shane, 1975). M81 is one of the few galaxies for which high enough resolution is obtainable (and presumably in which the non-circular motions are large enough) for the abrupt change in flow velocity at the arms predicted by Roberts to appear as wiggles in the isovelocity contours.

Visser's model was produced by piecing together one dimensional local solutions, and the result is assumed to be stationary. With our scheme, we are able to construct a fully two dimensional time dependent model. In particular we wish to determine the secular effects in the gas. The asymptotic spiral shocks are computed neglecting the azimuthal force component, and one might suppose that if the latter were indeed negligible, then there would not be the angular momentum redistribution associated with ring formation, and the gas response would be stationary.

We can fit Visser's M81 model force field well by the isochrone (at least outside $r = 2$ kpc), if we set 2 kpc = 1 isochrone unit. We adopt Visser's pattern speed of $18 \text{ km s}^{-1} \text{ kpc}^{-1}$ resulting in CR at 11.3 kpc, OLR at 15.5 kpc, and an ILR almost occurring at 3.2 kpc. Visser's spiral field can also be fitted by an analytic form. The spiral has a pitch angle of 15° , and the potential is chosen to be of the form

$$\psi = Ar^2 \exp(-r^2/r_0^2) \cos 2(\theta + \alpha \ln r),$$

with $r_0 = 7.0$ kpc and $\alpha = 3.73$. The radial force is

$$F_r = 2r (r^2/r_0^2 - 1) \exp(-r^2/r_0^2) \cos 2(\theta + \alpha \ln r) \\ + 2\alpha r \exp(-r^2/r_0^2) \sin 2(\theta + \alpha \ln r) .$$

Over the range of radii of interest, the second term is dominant and is given below as a fraction, F , of the axisymmetric force with the average of Visser's F_1 and F_2 models for comparison:

r (kpc)	F	F (Visser)
2	0.025	0.0
4	0.055	0.063
6	0.085	0.080
8	0.095	0.094
10	0.080	0.072
12	0.052	-

We make no attempt to reproduce Visser's model exactly in the central regions.

Particle distributions for the model are shown in Figure 5.9. At early times the pattern is very similar to that obtained by Visser (see his Figure 6.8(a) or 6.3). The spiral shock lies very close to the minimum of the potential near the ILR. In the outer parts, where there is a doubling of the arm, what appears to be the primary arm lies well inside the potential minimum as in Visser's case. The bifurcation of the arm is just what we observed in the bar driven models of Chapter 4 in the region around the second harmonic resonance. The temporal evolution of the model follows the familiar pattern. Rings form inside the second harmonic resonance, and outside the ILR as can be seen more clearly in the angular momentum distributions

FIGURE 5.9

As for Figure 5.2, for the simulation based on Visser's model of M81. The axes extend to ± 7 units.

↑
I

↑
C

1

2

3

4

(Figure 5.10). Indeed a doubling of the arms in M81 at about 8.3 kpc gives the appearance of a ring (see e.g. the deprojected HI maps of Rots (1975)). A radial velocity contour map for this model showing the same kinks as those obtained by Visser is presented in Section 6.3.

Thus contrary to initial expectations, a spiral field turns out to be just as efficient at transporting gas as a bar field. This example suggests that the inclusion of the spiral field produced by the gas, which we have ignored in the bar field models, will not change the qualitative behaviour of the response.

5.4 *The Cassini Division in Saturn's Rings:*

Another Example of an Inner Lindblad Resonance

The Cassini division in the rings of Saturn lies close to the radius at which ring particles have one-half the period of the satellite Mimas. Several authors have suggested that it is this resonance which is responsible for clearing the gap, although Alfvén (1968) favours a theory in which the gap is a relic from the time of formation of the satellites. Two important observations should be made concerning the structure of the rings (Dolfus, 1968). First, the position of the gap is not exactly on resonance, but is at a slightly larger radius. Secondly, the brightness inside the division (the outer edge of ring B) greatly exceeds that outside (ring A).

The first attempt to explain these observations in detail using a resonance model was made by Franklin and Colombo

FIGURE 5.10

As for Figure 5.3, for the simulation based on Visser's model of M81.

(1970), who calculated the periodic orbits of particles moving under the influence of Saturn and Mimas. Then, in order to avoid collisions between adjacent particles in their static model, they predicted a density proportional to $1/\Delta R$, where ΔR is the radial range swept out during a complete periodic orbit. The profile they obtained across the gap is in good agreement with brightness observations. More recently, Greenberg (1978) has shown how a dissipative process could lead to the clearing of the gap. The orbital resonance forces the larger bodies in the rings into non-circular orbits. The more numerous smaller bodies collide more often and so respond less strongly to the resonance and can thus provide an eccentricity-reducing medium for the larger bodies. These then experience a secular decrease in semi-major axis.

Now the frequency commensurability at the resonance is just the same as at the inner Lindblad resonance in the galactic case. We have seen how material is moved inwards at such a resonance if there is dissipation (Section 5.1). In this section then, we simulate the ring matter by particles which collide inelastically to ascertain whether such a process might have caused the Cassini division.

If we neglect the oblateness of Saturn, the force field is just the Keplerian

$$\Phi = - GM_S/r ,$$

M_S being the mass of Saturn. We can transform to a frame of reference centred on Saturn and rotating with Mimas (we neglect

the small eccentricity of Mimas' orbit, ~ 0.02). In this frame, the gravitational potential due to the satellite is

$$\psi = - \frac{GM_m}{R} [r^2/R^2 + 1 - 2(r/R) \cos\theta]^{-\frac{1}{2}}$$

where R is the radius of the orbit of Mimas, and θ the azimuth measured from the Saturn-Mimas radius. This expression may be decomposed into its Fourier components, after putting $u = \ln r$, $u' = \ln R$, by using the series

$$[2 \cosh(u-u') - 2\cos\theta]^{-\frac{1}{2}} = \pi^{-1} \sum_m Q_m |m|^{-\frac{1}{2}} [\cosh(u-u')] e^{im\theta}$$

where

$$Q_m |m|^{-\frac{1}{2}} (\cosh u) = \pi \Gamma(m+\frac{1}{2}) [\Gamma(\frac{1}{2}) \Gamma(m+1)]^{-1} {}_2F_1[\frac{1}{2}, m+\frac{1}{2}; m+1; \exp(-2|u|)] \exp[-(m+\frac{1}{2})|u|] .$$

The leading $m = 1$ term is exactly cancelled by the acceleration of the reference frame. Thus the 2θ term is the most important one, and it has the form

$$- \frac{3}{4} GM_m R^{-3} [1 + \delta(r)] r^2 \cos 2\theta ,$$

where $\delta(r)$ is small near the resonant radius $r/R = .62996$. We take $\delta(r)$ constant over the small range of radii of interest so that, rescaling so that $R = 1$ and $GM_S = 1$, we find

$$\Phi = - 1/r$$

and
$$\psi = - C (M_m/M_S) r^2 \cos 2\theta$$

where $C = 0.9102\dots$

Now the Mimas-Saturn mass ratio is 0.67×10^{-7} , but to simulate such a system would require more computer time and core memory than is available. We set $M_m/M_s = 0.002$ and compute the evolution of 2000 particles with an initial spacing of 0.012 (in the above units). The collision box size is chosen to be 0.004 and a fraction 0.8 of the relative motion is destroyed on each collision.

In the particle distributions of Figure 5.11, a spiral forms initially, and, as in the galactic case, a region outside the resonance is subsequently cleared out. The high densities in the ring, and the partial clearing of a region inside the ring are probably due to our particles being rather too sticky. Furthermore in the real case, the "collision box size" and the distance between ring particles is much smaller than we have used. As expected the ring inside the resonance is aligned perpendicular to the Mimas-Saturn axis. The particles outside the gap appear to move in orbits aligned with the axis: it is the region where the orbits intersect which has been depleted.

Because of the eccentricity of the orbits, the position of the gap relative to the resonance is a little unclear. However when the distributions in angular momentum are plotted (Figure 5.12) it becomes apparent that the gap lies just outside the resonance, as is indeed observed.

The width of the gap is about 0.05, and for a simulation with $M_m/M_s = 0.01$ it is about 0.12. We can find no reasonable scaling law which fits these results to the width of

FIGURE 5.11

As for Figure 5.2, for the simulation of the Cassini division of Saturn's rings. The arrow indicates the position of the resonance. The axes extend to the radius of Mimas' orbit, and Mimas lies on the x-axis.

FIGURE 5.12

As for Figure 5.3, for the simulation of the Cassini division in Saturn's rings. The arrow indicates the momentum of circular orbits at the resonance.

the gap which Franklin and Colombo (1970) claimed to be between 0.01 and 0.02 of the radius of Mimas' orbit. Of course, unlike in our experiment, the ring particles have undergone some 10^{12} periods of bumping! However we have at least shown that a dissipative process is capable of clearing a gap just outside the resonance.

CHAPTER 6

VELOCITY FIELDS AND THE COMPARISON WITH OBSERVATION

6.1 *The Velocity Field of the Standard Model*

The discussion of velocity fields in the gas has been left to this chapter because the observation of spiral structure both in our own galaxy and in external systems is very closely connected with their velocity fields. In external galaxies, non-circular motions observed in the HI radial velocity maps have provided one of the most crucial tests of the density wave theory so far. Yet the same non-circular velocities present the major obstacle to the unravelling of the spiral structure of our own galaxy.

The velocity field of the standard model is presented in Figure 6.1. The major point to note is the change in direction of the velocity at the arms as one expects to occur at a shock. It is indeed possible to identify the two flow patterns near the OLR corresponding to the two orientations of periodic orbits: the one in the outer regions aligned with the bar, the inner one perpendicular, with the appearance of a cusp on the major axis. It is where these two flows intersect that the spiral shocks occur. The flow is subsequently along the arm. The change in velocity, and with it the energy dissipated, occurs along a very narrow front. Just how narrow is highlighted by the plots of collision positions, and especially of those in which large amounts of energy are lost (see Figure 3.6). Even in the experiment with collision box size $RCOLL = 0.15$, in which the arms are quite broad with little suggestion of a shock (the first frame of Figure 3.14), such plots indicate abrupt changes

FIGURE 6.1

The velocity field for the standard model A at time 5 bar rotations. The length of the vectors is proportional to the velocity in the frame rotating with the bar, and the sense of rotation is clockwise outside CR and anti-clockwise inside CR. The bar lies along the x-axis.

in velocity along a narrow spiral front whose width is limited by the collision box size (lower frames of Figure 3.14). This sharp arm lies on the outside of the broader HI arm, as one expects for flow outside CR.

The field near CR is as expected: gas circulates around the L_4 , L_5 Lagrangian points. Because of the depopulation which eventually occurs of these points, the gas must travel a considerable distance within an arm during its orbit around the galaxy. This will lead to mixing of the gas on a length scale similar to the separation of the resonances, and the consequent implications for radial abundance variations should not be overlooked. Of course for weaker bars the mixing will not be as effective (see for example the deformed isochrone model of Chapter 3), but for the type of bar seen in SB systems, the streamlines must run along the arm for at least 90° . It is of interest in this connection that Pagel (1978) has found no abundance gradient along the spiral arms of the barred spiral NGC 1365.

6.2 *Torques Exerted by the Bar, and their Effects*

In our simulations we have assumed the bar potential to be fixed, yet in reality the torque it exerts on the gas will affect its own evolution. That this effect could be significant was pointed out by Kalnajs (1972a) who showed that the tightly wrapped spiral shocks of Roberts (1969) would damp the density wave which produced them on a timescale of one galactic year. Essential to the argument is the fact that the waves have negative energy (Lynden-Bell and Kalnajs, 1972).

We have seen that under the influence of bar forcing in the standard model, gas between CR and the OLR is moved to the OLR (Figure 3.8). At the same time the total angular momentum of the gas is increased by 10% over the 8 bar rotations after turn-on. This implies that the bar does (positive) work on the gas, and in the process will amplify itself. To be more precise, we have calculated the torque exerted on particles of a given angular momentum as a function of time, for the standard model. In the initial stages as the bar is being turned on, the gas near CR does work on the bar and the bar does work on the outer regions. In quasi-equilibrium (4 or 5 bar rotations), the torque exerted on the gas is positive with the exception of that in a narrow region of angular momentum just outside CR. Certainly the net torque (which is dominated by the contribution from the OLR region) is positive as we expect from the secular increase in angular momentum of the particles. Furthermore if we estimate the timescale for moving particles from midway between CR and the OLR (say) to the OLR, we obtain about 13 bar revolutions, in good agreement with the timescale for ring formation. [OLR and CR refer to the angular momentum scale. These can roughly be identified in the plots of particle position by the arrows which represent the corresponding radii for objects in *circular* orbits].

Now we can calculate the effect of the torque, L , exerted on the bar on its growth rate

$$\dot{H}/H = \Omega_p L/E$$

where H and E are the angular momentum and energy of the bar mode (Lynden-Bell and Kalnajs, 1972). The difficulty arises in

estimating E , i.e. the energy change in the disc produced by the presence of the mode. Following Kalnajs (1973a) we find that if we had a stable mode, then the energy is

$$E = \omega \frac{\partial \lambda(\omega)}{\partial \omega} \frac{1}{2} \int_0^{2\pi} \int_0^{\omega} 2\pi \psi(r, \theta) \sigma(r, \theta) r dr d\theta ,$$

where $\lambda(\omega)$ is the mathematical eigenfunction defined by Kalnajs (1976). We assume that the term $\omega \frac{\partial \lambda(\omega)}{\partial \omega}$ to be of order unity as is the case for the stable slow bar modes in the uniformly rotating disc (Kalnajs, 1972b). [It varies from zero for an unstable mode, to a large positive value for a mode for which $|\Omega - \kappa/2 - \Omega_p| \ll \Omega$]. Since the energy is negative, it follows that the bar mode will grow if it exerts a (positive) torque on the gas, and vice versa. The integral, which is the potential energy of the bar, can be calculated, and assuming 5% gas density in the disc, we can estimate the growth rate to be of order 5 bar rotations. It is indeed heartening that the response outside the OLR is such as to maintain the bar-like distortion of the disc which produced it.

For the strong bar ($q = 0.2$) mode (Section 4.2), the torque is fairly constant over the period 5-10 bar rotations, and its value then is about four times that for the standard case is quasi-equilibrium. This of course implies that the growth rate of the bar is the same in both cases. After ring formation (15-20 bar rotations) the net torque drops effectively to zero. In the strong bar experiment without collisions a net (time-averaged) torque is still exerted on the particles after 20 bar rotations. The corresponding growth rate is 40 bar

rotations. This is presumably due to the unstable orbits, and has implications for the maintenance of a bar in the purely stellar case.

In the inverted isochrone model (Section 5.1), the torque exerted by the bar on the gas after turn-on is dominated by the contribution from the ILR region and is negative. Here the gas response would damp the bar mode which caused it. The evolution of the model is rapid, so the torque never reaches a quasi-equilibrium value.

The situation is similar for Visser's spiral field (Section 5.3). Here there is not even any contribution from the OLR to partially cancel the work done by the gas on the stellar spiral at the ILR. Some help comes from the ultraharmonic resonance region but the net effect is still a damping of the wave. Using arguments similar to those above we can estimate the timescale to be about 1.5 bar rotations, although the value of $\omega \frac{\partial \lambda(\omega)}{\partial \omega}$ is likely to be greater than unity for this case.

6.3 *Rotation Curves*

It has become apparent recently that the rotation curves of many galaxies appear to remain flat to quite large radii (Bosma, 1978). This would imply more mass in the outer regions than can be accounted for by the luminosity, which of course falls off exponentially. The initial reaction was to look for massive haloes around galaxies, but as yet no conclusive evidence has been found. We do not intend to provide any solution to this problem in this section, but merely to issue a cautionary warning concerning methods of determining rotation curves.

Of course non-circular motions in the discs of spiral galaxies mean that it is not possible to determine the rotation curve from the run of radial velocities along the major axis, even if this were well determined (see the remarks in the next section). Some procedure must be followed in which all the velocity data from an HI radial velocity map (say) is used. The situation is further complicated by the possibility of warps in the outer parts, so that the inclination and position angle may vary with radius in the galaxy.

In order to illustrate the problem quantitatively, we examine the procedure used by Bosma (1978) to determine the rotation curve from HI maps of external galaxies. At each radius he assumes various values of inclination, i , and position angle, $p.a.$, and for each a cosine is fitted to the run of radial velocity (weighted by HI mass) around a deprojected ring. The values of i and $p.a.$ which give the best fit in the least squares sense are then accepted for that radius. The rotation velocity is of course simply the amplitude of the cosine fit.

We have projected one of our models (the isochrone disc with a deformed isochrone bar rotating with $\Omega_p = 0.136$) at 45° to the line of sight, and at various position angles, and then applied the above procedure. The rotation curves obtained are shown in Figure 6.2. For position angle 90° , the derived curve is almost flat, whilst for 0° , it falls off more rapidly than the axisymmetric curve. The derived values of i and $p.a.$ often varied by 10° or more from the real values. No evidence was obtained which would suggest that such effects could masquerade as the kinematic warps of Bosma however. (Kinematic

FIGURE 6.2

Rotation curves obtained by applying Bosma's procedure to a projected model (squares), together with the curves corresponding to the applied (isochrone) axisymmetric field (solid line).

warps are those determined by the above procedure rather than by visual inspection). Neither could the preponderance of flat rotation curves be ascribed to such procedures since one would expect roughly an equal number which decreased more rapidly than expected. Nevertheless one should be aware of the problems involved in determining accurate rotation curves.

6.4 *Radial Velocity Maps*

In order to compare the velocity fields of our models with those from HI observations we have produced "radial velocity contour maps". Particles are binned on a rectangular grid, and the average velocity in each bin corrected to the radial velocity which would be observed if the galaxy was inclined to the line of sight at a given position angle. Unfortunately we have the same problem as radio astronomers: insufficient signal, so we must add together 10 distributions separated by 2 timesteps. This is sufficient to randomize the particle positions with respect to the bin boundaries. The resolution which is then achieved is roughly the same as in the observations of M81 by Rots and Shane (1975). The results for the standard case with position angles 45° and -45° (measured from the bar axis) are shown in Figure 6.3.

The contours follow the normal "spider" pattern as expected with wiggles appearing at the arms as a result of the sudden change in velocity there. These kinks have provided one of the few observational tests of the density wave theory of spiral structure. Visser (1978), on the basis of gas flow in a tightly wrapped spiral field, has been able to fit the wiggles in the observed maps of M81 on the eastern arm but not on the

FIGURE 6.3

Contours of equal radial velocity as would be seen by an observer at position angles 45° and -45° to the bar axis for the standard model. The resolution element or "beam size" is shown by the box.

western arm. This is an important point as the exact details of the contours near an arm will depend on whether the arm is inside or outside CR. Unfortunately the only galaxy for which the details of the velocity contours at the arms can be resolved, M81, is obviously tidally distorted by a nearby companion, which accounts for the lack of bisymmetry.

We have plotted the radial velocity contours for our simulation based on the spiral forcing model of Visser (Section 5.2) and the results are shown in Figure 6.4. The position angle used is roughly equivalent to that used by Visser. The results confirm the shapes of his wiggles. Furthermore the exact position of the kinks in relation to the arms is different from that in the standard case. (To convince oneself of this, two regions at the same angle to the inclination major axis must be compared). This reflects the opposite sense of flow around the galaxy with respect to the pattern inside and outside CR.

A second important feature in the contour maps of the standard model is the non-alignment of the kinematic major axis the true major axis. In the inner parts the former may be determined from the normal to the velocity contours (about 48° for position angle -45°), in the outer parts from the position of the radial velocity hills (at the most 37° for the same case). As Kalnajs (1978) has pointed out, this effect should be easily observable in radial velocity contour maps. Indeed Bosma (1978) has used the phenomenon to identify oval distortions in external galaxies.

FIGURE 6.4

Contours of equal radial velocity as would be seen by an observer at position angle indicated by the arrow for the simulation based on Visser's model of M81. The resolution element or "beam size" is shown by the box.

6.5 *Longitude-Velocity Maps*

There have been many attempts to determine the large-scale structure of our Galaxy from HI observations, most recently by Weaver (1970) and Simonson (1977). Basically the brightness temperature, T_B , as a function of longitude and radial velocity is transformed to hydrogen density as a function of position on the galactic disc. The major problem is determining the radial velocity as a function of distance. An axisymmetric model is not adequate as the early maps (Kerr and Westerhout, 1965) show. The situation is further complicated by velocity crowding - that is a region where radial velocity is fairly constant with distance appearing as an arm-like feature in the maps of brightness temperature in the longitude-radial velocity (l-v) plane.

We have produced l-v maps of our models for various solar positions and compared these with the map for the galaxy produced by Valdes (1978). His presentation is most appropriate as z-effects have been eliminated by integrating between $\pm 2^\circ$ latitude. Our maps are produced by binning the particles in longitude around the assumed solar position and in radial velocity as would be seen by an observer moving with gas at that position. To produce a quantity comparable with brightness temperature we use a weighting factor proportional to $1/r^2$, r being the distance from the "sun". Six distributions separated by two timesteps are added together to improve the statistics. Because of the discrete nature of our particles, the local gas ($r = 0$) cannot be well represented. To obtain the equivalent of integrated brightness temperature between $\pm 2^\circ$ latitude in a disc of thickness about 200 pc, we apply a linear tapering function, zero at the sun and

unity at one quarter of the radius of the solar circle from the sun.

Figure 6.5 shows a series of maps for the standard model (at time 3 bar rotations) with the sun at various positions on a circle of radius 1.5. For this model we should compare only the regions outside the solar circle with those in our galaxy i.e. only the regions $360^\circ > l > 180^\circ$, $v > 0$ and $180^\circ > l > 0^\circ$, $v < 0$. The resulting map for a solar position 67.5° from the bar major axis, bears most resemblance to that of the Galaxy. The main arm crosses the galactic centre direction ($l = 180^\circ$) at almost exactly zero radial velocity, as is observed. This is not the case for other solar positions. Of course this arm is composed of several filaments in the galactic case but it is the gross structure with which we are concerned here. At higher longitude (smaller radius) the arm fades out before it has reached $l = 270^\circ$; in the other direction it continues to about $l = 30^\circ$, slightly further than in the Galaxy. In this region a weak higher velocity arm also appears. This arm, which is much stronger in the standard model B (Section 3.2) corresponds to gas outside the OLR in the flow which is aligned with the bar. The l - v map for model B is reproduced in Figure 6.6. Both arms occur in the galactic map. The hole in the HI at $l = 120^\circ$, $v = 25 \text{ km s}^{-1}$ in Valdes' map is presumably due to the clearing out of gas from the L_4 , L_5 Lagrangian points, though at the early stage of our model this has not yet occurred. Between $l = 270^\circ$ and $l = 360^\circ$ there is a stubby feature which can just be resolved into two arcs. These correspond to the doubling of the arms just outside CR. There is indeed a similar two-fold structure in the galactic

FIGURE 6.5

Longitude-radial velocity maps produced from the standard model A for "Sun" at the position angles with respect to the bar major axis shown. Longitude in the model increases clockwise, and the radial velocity is reversed in order for the maps to be comparable with those of Valdes'.

FIGURE 6.6

Longitude-radial velocity maps produced from (a) the standard model B; (b) the inverted isochrone model and (c) the inverted isochrone model with velocities replaced by the appropriate circular values. All are for the "Sun" position angle 67.5° from the bar major axis (measured anticlockwise).

maps, although the lower velocity arm is much stronger. We can only hope to fit general features without varying the parameters of our model. Nevertheless the similarity in the regions outside the solar circle is remarkable.

We also include in Figure 6.6 a map derived from the inverted isochrone model of Section 5.1, to compare some of the features inside the solar circle. The shoulders which occur in the "rotation curve" edge of the galactic map ($90^\circ > l > 0^\circ$, $v > 0$ and $270^\circ < l < 360^\circ$, $v < 0$) arise in our model from the inner spiral or ring at the ILR. Furthermore, although our maps are unreliable for local gas, the scarcity of positive velocity HI near $l = 240^\circ$ is strikingly similar to that in the Galaxy and remarked upon by Valdes. Finally for comparison, a map is shown in which the velocities have been replaced by the appropriate circular values. This demonstrates how the non-circular velocities have been increased near the galactic centre by the action of the bar.

These l - v maps give us further confidence in our models, especially as we had no intention that our simulations should model the Galaxy: the production of l - v maps was very much an afterthought. Most importantly, we obtain the major features of our own galaxy with a basically two-armed spiral. We have also obtained maps for the sun at different radii in the standard model, though for the rough comparison we have made, we can only limit this parameter to roughly $1 < r < 2$: the sun is about at CR. The future aim should be to combine the features of the two models into one.

CHAPTER 7

PERIODIC ORBITS IN A BAR FIELD

7.1 *A Second Order Theory*

The epicyclic theory of orbits assumes that the deviations from circular motion are small and is of very limited applicability to our models. Here we investigate periodic orbits using a second order theory developed by Lynden-Bell and Kalnajs (see Lynden-Bell, 1973). They utilize the ideas of action-angle variables which are dealt with in any text on advanced dynamics (e.g. Goldstein, 1959) or in Born's book "The Mechanics of the Atom".

Hamilton's equations for a one-dimensional system are just

$$\frac{\partial H}{\partial p} = \dot{q} \quad , \quad \frac{\partial H}{\partial q} = -\dot{p}$$

for a Hamiltonian $H(p,q)$, where the co-ordinates p,q are canonically conjugate. Then if the motion is periodic we define the action variable

$$J = \oint p dq$$

where the integration is over a complete cycle of q . The co-ordinate, w , conjugate to the momentum J is called the "angle", and is found by solving the Hamilton-Jacobi equation

$$H\left(\frac{\partial S}{\partial q}, q\right) = H'(J)$$

where $p = \frac{\partial S}{\partial q}$ and $w = \frac{\partial S}{\partial J}$.

It follows then from Hamilton's equations in the new action-angle variables that J is constant in time, so that $\frac{\partial H'}{\partial J}$ is also.

Thus we may write

$$w = \omega t + \beta$$

where ω is the generalized frequency for the motion in the "angle" w , and β is a constant of integration.

All the above holds in the multi-dimensional case, except that it is then possible that the frequencies are commensurable, i.e. that integers n_i exist such that $\sum n_i \omega_i = 0$. In that case we may transform to new action-angle variables such that $w_s = \sum n_i w_i$ is one of the new angles. Under the influence of a perturbation, the new ω_s will still be small and the corresponding w_s will change only slowly. The other angles, w_f , will be circulating more rapidly, so the corresponding J_f will be adiabatically invariant. The motion in the slow variable is found by averaging over the rapid circulation in the fast variables.

Now if we consider a perturbation

$$\psi = \psi_0(r) \operatorname{Re}\{\exp[2i(\theta - \Omega_p t)]\}$$

of the axisymmetric disc, then the resonances occur when the forcing frequency of the bar is commensurable with the natural radial frequency, i.e.

$$m(\Omega - \Omega_p) + l\kappa = 0, \text{ for integer } l, m.$$

Clearly the new slow variable should be chosen as

$$w_s = w_\theta - \Omega_p t + (1/m)w_r$$

where w_θ and w_r are just the angles corresponding to the actions

$$J_r = \frac{1}{2\pi} \oint v_r dr$$

and $J_\theta = \frac{1}{2\pi} \oint v_\theta d\theta = h$, the angular momentum. w_s may be derived from a generating function

$$S = (w_\theta - \Omega_p t + 1/m)J_s + w_r J_f$$

from which we find the new fast angle

$$w_f = w_r$$

and the new actions

$$J_s = J_\theta$$

$$J_f = J_r - (1/m)J_\theta$$

at co-rotation $l = 0$ and so $J_f = J_r$, but the more interesting cases in our context are the principal Lindblad resonances where $m = 2$, $l = \pm 1$. Now we expand the perturbing bar field in action-angle co-ordinates:

$$\psi = \sum_k \psi_k(J_f, J_s) \exp(ikw_f + (2iw_s)),$$

and then average over the fast motion which leaves us with $\psi(J_r, J_s) \exp(2iw_s)$. For numerical purposes we can evaluate the perturbation using the epicyclic approximation:

$$r - r_0 = a \cos w_r$$

$$\theta = w_\theta t - \frac{2\Omega}{\kappa} \frac{a}{r_0} \sin w_r$$

$$\text{with } a \equiv \sqrt{2J_r/\omega_r}$$

$$\text{and } r_0 = r_0(J_s)$$

(see Lynden-Bell and Kalnajs, 1972).

So we obtain to first order in a

$$\begin{aligned} \psi &= \psi_0(r_0, w_\theta) \exp(2iw_\theta - 2i\Omega_p t) \\ &+ \left[\frac{a}{2} \frac{\partial \psi_0}{\partial r} \right]_{r=r_0} - \frac{a}{r_0} \frac{2\Omega}{\kappa} \psi_0 \exp[iw_r + 2i(w_\theta - \Omega_p t)] \\ &+ \left[\frac{a}{2} \frac{\partial \psi_0}{\partial r} \right]_{r=r_0} + \frac{a}{r_0} \frac{2\Omega}{\kappa} \psi_0 \exp[-iw_r + 2i(w_\theta - \Omega_p t)] \\ &\dots (1) \end{aligned}$$

7.2 The Outer Lindblad Resonance

At the OLR, $\Omega + \kappa/2 = \Omega_p$, so $l = +1$ and

$$w_s = w_\theta - \Omega_p t + w_r/2,$$

$$\text{and } J_f = J_r - J_\theta/2$$

The second term of equation (1) is the slowly varying one and its real part is given by

$$\psi_p = \left(\frac{a}{2} \frac{\partial \psi_0}{\partial r} \right)_{r=r_0} - \frac{a}{r_0} \frac{2\Omega}{\kappa} \psi_0 \cos 2w_s$$

We consider the isochrone disc which is expressible in terms of J_r and J_θ as

$$H_0 = -\frac{1}{2}(J_r + |J_\theta|/2 + \sqrt{J_\theta^2/4 + 1})^{-2}$$

or as a function of J_f and J_s ,

$$H_0 = -\frac{1}{2}(J_f + J_s + \sqrt{J_s^2/4 + 1})^{-2}$$

Now if we turn the perturbation on slowly, a circular orbit with actions $J_r^0 = 0$, J_θ^0 , will deform into a periodic orbit with $J_f = J_f^0$ since J_f is adiabatically invariant. So

$$J_f^0 + J_s^0/2 = 0$$

and

$$J_s - J_s^0 = 2J_r = \kappa a^2$$

in the epicyclic approximation.

Hence the Hamiltonian, written as the difference between the energy of the perturbed orbit and the reference orbit is, to second order in $(J_s - J_s^0)$

$$\begin{aligned} H_s &= \left[\left(\frac{\partial H_0}{\partial J_s} \right)^0 - \Omega_p \right] (J_s - J_s^0) + \frac{1}{2} \left(\frac{\partial^2 H_0}{\partial J_s^2} \right)^0 (J_s - J_s^0)^2 + \psi_p \\ &= a^2 \kappa [F(J_s^0) - \Omega_p] + \frac{a^4 \kappa^2}{2} G(J_s^0) + \psi_p \quad \dots (2) \end{aligned}$$

where $F(J_s^0) = (\omega_\theta + \omega_r/2)^0$

and $G(J_s^0) = \left[\frac{\partial}{\partial J_s} (\omega_\theta + \omega_r/2) \right]^0$.

F and G may be written as explicit functions of J_s^0 (or equivalently r_0).

The periodic orbits are found by looking for the turning points in H_s , i.e. $\frac{\partial H_s}{\partial w_s} = 0$ and $\frac{\partial H_s}{\partial a} = 0$. This yields

$$\cos 2w_s = \pm 1 \quad \dots (3)$$

and

$$2a\kappa [F(J_s^0) - \Omega_p] + 2\kappa^2 a^3 G(J_s^0) \pm \left(\frac{\partial \psi_0}{\partial r} \right)_{r=r_0} - \frac{1}{r_0} \frac{2\Omega}{\kappa} \psi_0 = 0 \quad \dots (4)$$

(3) and (4) are the fundamental equations describing the behaviour of the periodic orbits near the resonance.

Outside the OLR, $F(J_s^0) - \Omega_p < 0$, and $G(J_s^0) < 0$, so the cubic equation (4) has one root for each sign of $\cos 2w_s$. When $\cos w_s = +1$, a is positive and so the orbit is aligned along the bar major axis. "On resonance", $F(J_s^0) = \Omega_p$, the alignment of the orbits is the same, and importantly, their eccentricity varies as the cube root of the perturbation field. Inside the resonance, $F(J_s^0) - \Omega_p > 0$, and three positive roots of equation (4) exist for small enough perturbations. It is easily verified that the largest value of a gives a stable orbit aligned with the bar. This orbit is a continuation of those outside the resonance. The other two solutions represent orbits aligned perpendicular to the bar axis, and calculation of the second derivative, $\frac{\partial^2 H}{\partial a^2}$, shows that the smaller of the two values of a gives a stable orbit, whilst the other an unstable orbit. As

the perturbation field is increased the two roots coalesce and subsequently become imaginary. Thus in our parameterization of the orbits in terms of the fast action $J_f = J_s^0$, there is a region inside the OLR, in which stable orbits aligned perpendicular to the bar do not exist. For strong perturbations the above analysis can be used only as a guide, because one must worry about higher order terms in the expansion of H_s , and about the "fast" perturbation.

7.3 The Inner Lindblad Resonance

At the ILR, $\Omega - \kappa/2 = \Omega_p$, so the new slow variable is

$$w_s = w_\theta - \Omega_p t - w_r/2,$$

and the associated actions

$$J_s = J_\theta,$$

$$J_f = J_r + J_\theta/2.$$

After averaging equation (1) over the fast motion, we obtain for the real part of the perturbation

$$\psi_p = \left(\frac{a}{2} \frac{\partial \psi_0}{\partial r} \right) \Big|_{r=r_0} + \frac{a}{r_0} \frac{2\Omega}{\kappa} \psi_0 \cos 2w_s.$$

As for the OLR, if we turn the perturbation on slowly, a circular reference orbit with actions $J_r^0 = 0$, J_θ^0 , will deform into an orbit with $J_f = J_f^0$. However in this case

$$J_f^0 = J_s^0/2$$

and $(J_s - J_s^0) = -2J_r = -a^2\kappa.$

The Hamiltonian is then given by

$$H_s = -a^2 \kappa [f(J_s^0) - \Omega_p] + \frac{a^4 \kappa^2}{2} g(J_s^0) + \psi_p$$

where $f(J_s^0) = (\omega_\theta - \omega_r/2)^0$

and $g(J_s^0) = \left[\frac{\partial}{\partial J_s} (\omega_\theta - \omega_r/2) \right]^0$.

Finding the turning points of H_s in the usual way yields

$$\cos 2w_s = \pm 1$$

$$\text{and } -2a\kappa \left[f(J_s^0) - \Omega_p \right] + 2a^3 \kappa^2 g(J_s^0) \pm \Psi = 0 \quad \dots (5)$$

$$\text{where } \Psi = \left. \frac{1}{2} \frac{\partial \psi_0}{\partial r} \right|_{r=r_0} + \frac{\psi_0}{r_0} \frac{2\Omega}{\kappa} \quad \dots (6)$$

Considering first the outer ILR of a pair, $f(J_s^0) - \Omega_p$ is positive inside the resonance and $g(J_s^0) < 0$, so the equation (5) is of the same form as equation (4) outside the OLR. However in this case, provided ψ_0 does not change too rapidly at the resonance, $\Psi < 0$, and the periodic orbit is aligned perpendicular to the bar axis. Outside the resonance, the behaviour is the same as inside the OLR except for a similar reversal of orbit alignment. However it is possible in the case of a central bar, that $\Psi > 0$. In this case, the behaviour at the ILR is exactly similar to that at the OLR, "turned inside out". This situation was discussed by Contopoulos and Mertzianides (1977) who derived the same condition for change of orbit orientation. The possible relevance to galaxy morphology has been pointed out in Chapter 5. The other ILR should behave presumably just like an OLR, but at

any rate the resolution in our models is insufficient at those radii and the behaviour is probably dominated by finite collision box effects.

7.4 *Numerical Integration of Orbits*

The method for locating periodic orbits in this section is to integrate an initially circular orbit in a perturbation which is turned on slowly (over a time between 2 and 8 bar revolutions).

We have already seen in Chapter 3 how the particles in the central regions of the standard model follow almost elliptical orbits. As the radius increases towards CR, the orbits become more pointed at the bar major axis. Eventually the velocity on this axis becomes zero, and the orbits are no longer closed around the centre of the model in the rotating frame of reference. These are the familiar banana orbits (Figure 7.1(a)) which circulate about the stable Lagrangian points L_4 , L_5 (those 90° from the bar axis). The fact that some of these orbits do not remain trapped in the standard model (Figure 7.1(b)) is the first sign that the theory can only be used as a guide even at quite moderate bar field strengths. The real surprise occurs for circular orbits just inside the OLR. These clearly become unstable and their eccentricity increases without limit (Figure 7.1(c)). The particle appears to return to almost exactly the same radius after each excursion though at random phase with respect to the bar. Kalnajs has studied these orbits in some detail and found that at perigalacticon the particle may receive a little kick from the bar, increasing or decreasing the

FIGURE 7.1

Orbits in the standard bar field. The radius of the initial circular orbits is given in each case.

d

e

f

angular momentum depending on the phase. This results in a random walk in angular momentum, and because Jacobi's integral is conserved, also in radial action.

Outside the region where the orbits become unstable, though still inside the OLR, is a region where the major axis appears to wobble about the bar axis though it is difficult to find the periodic orbit (Figure 7.1(e)). Inside the unstable region, the alignment is perpendicular to the bar as expected. One might expect the second order theory (Section 7.2) to predict the radius at which this orbit becomes unstable, since it will occur where there are two equal roots, a , of the cubic equation (4), i.e. where

$$\frac{4}{3} \left[F(J_s^0) - \Omega_p \right]^{\frac{3}{2}} \sqrt{\frac{\kappa}{3G(J_s^0)}} + \left(\frac{1}{2} \frac{\partial \psi_0}{\partial r} \right) \bigg|_{r=r_0} \frac{\psi_0}{r_0} \frac{2\Omega}{\kappa} = 0.$$

However, because of the strong field, the predicted radius is too far from the resonance. At fields one tenth as strong, the agreement is good ($r \sim 2.54$). The eccentricity of the inner orbit at that point, 0.15, is also close to that predicted

$$a = \left[\frac{F(J_s^0) - \Omega_p}{3\kappa G(J_s^0)} \right]^{\frac{1}{2}}$$

$$= 0.1398 \quad \text{at} \quad r = 2.54.$$

In the case of the standard model with the Kalnajs bar, orbits are also unstable near the $(\Omega + \kappa/4) = \Omega_p$ resonance, which occurs at $r = 2.1$ (Figure 7.1(f)). Indeed for the strong bar field of section 4.2 (twice the standard strength) the

entire region between CR and the OLR becomes unstable.

The situation at an ILR is more complicated and is not studied in detail here. We have however confirmed the general behaviour expected for weak fields; a detailed comparison between predicted and numerical eccentricities has been made by Contopoulos and Mertzanides (1977). In particular we find, as was predicted for a central bar, that outside the ILR, the orientation changes by a further 90° through an intermediate circular orbit. If the bar is placed suitably, this can markedly reduce the eccentricity of orbits near the ILR. For strong bars we find no unstable behaviour like that observed at the OLR, although trapping of the orientation of an orbit along either the major or minor axes of the bar becomes impossible to achieve.

CHAPTER 8

CONCLUSIONS

As clearly pointed out by Oort (1962), gas is essential to the formation of spiral structure. The most obvious property which distinguishes it from the stellar subsystem is that it can dissipate energy. The aim of this study was to simulate the effect of dissipation in the gaseous subsystem and investigate its role in the formation of spiral structure.

By forcing the gas with a bar field, we have shown that spirality arises naturally from dissipation. This is in contrast to the asymptotic theory, where a spiral field is assumed at the outset. Furthermore, the necessity of dissipation has been demonstrated by suppressing collisions, thus simulating stars. While both stars and gas show spiral structure in the initial phases while the field is changing, once the field becomes steady, stars settle into a state which shows no spirality.

We interpret the development of spiral structure in the following way. If particles loaded uniformly on the disc are to participate in periodic flow without their trajectories crossing, then these trajectories must in fact be the periodic orbits in the bar field. But, as we have shown, a periodic orbit near a Lindblad resonance intersects one with slightly higher or lower angular momentum. Furthermore, close to such a resonance there exist more than one family of periodic orbits aligned at different angles to the imposed bar field, and these families also cross in space. Thus a periodic flow without crossing

trajectories is not possible: the gas clouds must bump, and the locus of these collisions turns out to be spiral in shape.

The spiral nature of the response allows the bar to exert a net torque on the particles, with the result that they move towards the Lindblad resonances. In the final state, these particles all move in a small range of periodic orbits, thus forming a ring. In this way the energy lost through collisions is minimized. During the evolution towards this end state the ring captures particles from other orbits.

The ring which forms at the OLR can be aligned along or perpendicular to the bar, depending on bar strength, initial conditions, and the collision process used, the former orientation being the more robust one. The appearance of the arms in the spiral phase depends upon which family of periodic orbits dominates. The archetypes of the two cases are models A and B of Chapter 3.

The secular effects mentioned above move particles to the OLR and to the inside of the ILR. The surprise is not so much that such effects occur, but rather the rate at which collisions cause particles to be moved. The timescale for ring formation at the OLR is short compared to the lifetime of a galaxy and there are many galaxies without rings. The process which is most likely to slow the evolution towards that final state is recycling of gas through star formation and stellar mass loss. A crude simulation of recycling showed that spiral structure could be maintained indefinitely. Other processes which could have a similar effect are circulation of gas out of the galactic

plane, and the locking up of a significant fraction of gas in small, dense (and hence effectively collisionless) clouds. However their contributions are difficult to estimate.

In the case of an ILR, we have shown that gas in the ring can be forced towards the centre of the model on a time-scale of a few bar rotations. It is likely that this will trigger star formation, and since many Seyferts have small spirals or ring-like structures embedded within a bar, this could be the origin of their activity.

Secular effects are not peculiar to a bar field: we have observed them in the case of the spiral field employed by Visser to model M81.

Secular effects in the gas imply that the bar is exerting, in most cases, a net torque which will influence its own evolution. We have shown that the torque on the gas at the OLR is positive and will amplify the bar mode, while that at the ILR is negative and will damp it. The net effect in the standard model was a growth of the bar on a timescale of about 5 bar rotations; for Visser's fields, the stellar spiral will decay in roughly 1.5 rotations. In this sense gas at the OLR seems to be important to the maintenance of spiral structure.

[The argument may also be given in terms of energy exchange, though it may at times appear to be somewhat paradoxical. For example, kinetic energy is always destroyed in collisions, yet the quasi-static configuration which forms as a result of these collisions around the OLR is such that the energy gain from the bar exceeds that loss.]

While studying free orbits in the bar field, we discovered that there is a region between CR and the OLR in which no stable orbits exist. Even for quite moderate bar fields, circular orbits just inside the OLR become unstable and grow in eccentricity without limit. For stronger bars the region of instability extends inward to CR. It has been shown that this effect produces an exponential tail on the disc, a frequent occurrence in real galaxies (Freeman, 1970). If such an instability is responsible for the tails, then because they are made up of highly eccentric orbits, they must be "hot".

Our models bear remarkable resemblance to real galaxies. This is notable in view of the few assumptions made in constructing them: the parameters of the bar field were chosen to be consistent with observations, the results of numerical experiments involving stellar discs, and bar mode analyses. The pitch angles obtained are similar to those observed, and this angle decreases in the outer parts, a common feature of both barred and normal spirals. Furthermore the depopulated regions on either side of the bar at CR are reminiscent of the holes which occur in almost every SB galaxy (as well as in some classified as normal).

Sharp changes in velocity occur at the arms and these show up as wiggles in radial velocity contour maps of the models. Similar kinks occur in the observed maps of M81. We have investigated the relation of the wiggles to the arms and how it changes across the CR radius.

As a bonus, longitude-radial velocity ($l-v$) maps constructed for the models have a striking similarity to those of the Galaxy. All the main features can be reproduced by our basically two-armed model if the "Sun" is placed about CR and 70° off the bar axis (in the sense of galaxy rotation).

We have also simulated the effect of an ILR in another context, namely the clearing of the Cassini division in the rings of Saturn. The forcing field in that case is due to the satellite Mimas. We have no difficulty accounting for the location of the observed gap just outside the resonance if we assume the ring material to be dissipative, but we have not been able to reproduce its width. The explanation of this failure may well lie in our inability to simulate the long time scales involved.

ORBITS AROUND COROTATION

The equations of motion of a particle in the potential $U(r, \theta) = \Phi(r) + \psi_0(r) \cos 2(\theta - \Omega t)$ are given by

$$\ddot{r} - r\dot{\theta}^2 - \Omega^2 r - 2\Omega r\dot{\theta} = -\frac{\partial U}{\partial r} \quad \dots (1)$$

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} + 2\Omega\dot{r} = -\frac{1}{r} \frac{\partial \psi}{\partial \theta} \quad \dots (2)$$

in a frame of reference rotating with the bar-like perturbation $\psi(r, \theta)$.

Now expanding around the extrema of the effective potential, $U(r, \theta) - \frac{1}{2}\Omega^2 r^2$, i.e. where $\theta = \theta_0 = 0, \pm \frac{\pi}{2}$ or π , and $r = r_c = \frac{1}{\Omega^2} \frac{\partial \Phi}{\partial r} \Big|_{r=r_c}$, we obtain

$$\frac{\partial U}{\partial r} = \Omega^2 r_c + \frac{\partial^2 (\Phi + \psi)}{\partial r^2} \Big|_{r=r_c} (r - r_c),$$

and $\frac{\partial \psi}{\partial \theta} = \frac{\partial^2 \psi}{\partial \theta^2} \Big|_{\theta=\theta_0} (\theta - \theta_0).$

Linearizing (1) and (2) gives

$$\ddot{r} - 2\Omega r_c \dot{\theta} - \Omega^2 (r - r_c) = -\frac{\partial^2 (\Phi + \psi)}{\partial r^2} \Big|_{r=r_c} (r - r_c) \quad \dots (3)$$

and $r_c \ddot{\theta} + 2\Omega \dot{r} = -\frac{1}{r_c^2} \frac{\partial^2 \psi}{\partial \theta^2} \Big|_{\theta=\theta_0} r_c (\theta - \theta_0) \quad \dots (4)$

Putting $r = r_c + \text{Re } e^{i\omega t}$,

and $r_c \theta = r_c \theta_0 + \theta e^{i\omega t}$,

we obtain

$$(\omega^2 + K_1)R + 2i\omega\Omega\theta = 0$$

and $(\omega^2 + K_2)\theta - 2i\omega\Omega R = 0$

where $K_1 = -\frac{\partial^2(\Phi + \psi)}{\partial r^2} \Big|_{r=r_c} + \Omega$

$$= -\frac{\partial^2\psi}{\partial r^2} \Big|_{r=r_c} - \kappa^2 + 4\Omega^2$$

and $K_2 = -\frac{1}{r_c^2} \frac{\partial^2\psi}{\partial \theta^2} \Big|_{\theta=\theta_0}$.

The general solution is

$$R = A \cos(\omega_1 t + \epsilon_1) + B \cos(\omega_2 t + \epsilon_2)$$

$$\theta = -k_1 A \sin(\omega_1 t + \epsilon_1) - k_2 B \sin(\omega_2 t + \epsilon_2),$$

where

$$k_1 = 2\omega_1\Omega/(\omega_1^2 + K_2) = (\omega_1^2 + K_1)/2\omega_1\Omega$$

$$k_2 = 2\omega_2\Omega/(\omega_2^2 + K_2) = (\omega_2^2 + K_1)/2\omega_2\Omega,$$

and where the frequencies ω_1, ω_2 satisfy

$$\omega^4 + \omega^2(K_1 + K_2 - 4\Omega^2) + K_1K_2 = 0 \quad \dots(5)$$

i.e. $\omega^2 = \frac{1}{2}(4\Omega^2 - K_1 - K_2) \pm \frac{1}{2} \left[(4\Omega^2 - K_1 - K_2)^2 - 4K_1K_2 \right]^{\frac{1}{2}}$.

In the limit of vanishing bar field strength the two frequencies become κ and zero, just the normal radial and azimuthal frequencies as seen in the co-rotating frame. At the equilibrium points $\theta_0 = 0, \pi$ (the Lagrangian points L_1 and L_2), $K_1 > 0$ and $K_2 < 0$ for moderate bar strengths. Thus one value of ω^2 is negative and the points are unstable. There may be a range of fields, sufficiently strong that $K_1 < 0$, for which these points are stable, although there are no such local minima at these points in any of our models. At the Lagrangian points L_4, L_5 ($\theta_0 = \pm\pi/2$), $K_2 > 0$ and the stability criterion is

$$(4\Omega^2 - K_1 - K_2)^2 > 4K_1K_2 \quad \dots(6)$$

This limits the curvatures K_1, K_2 , of the effective potential, or equivalently, given Φ and the form of ψ , it limits the strength of the perturbation. For example if we set

$$K_1 = K_1^0 (1 + \lambda) = (4\Omega^2 - \kappa^2) (1 + \lambda)$$

$$\text{and} \quad K_2 = \lambda K_2^0,$$

then for small λ the condition (6) becomes

$$\lambda < \kappa^2 / [2\kappa^2 (4\Omega^2 - \kappa^2 + K_2^0) + 4(4\Omega^2 - \kappa^2) K_2^0] .$$

Now the linearized equations (4), (5) may be derived from the Lagrangian

$$L = \frac{1}{2}(\dot{R}^2 + \dot{\theta}^2) + \Omega(R\dot{\theta} - \theta\dot{R}) \\ + \frac{1}{2}K_1 R^2 + \frac{1}{2}K_2 \theta^2,$$

so that the momenta conjugate to R and θ are

$$p_r = \dot{R} - \Omega\theta$$

$$p_\theta = \dot{\theta} + \Omega R ,$$

and the Hamiltonian is

$$H = \frac{1}{2}(p_r^2 + p_\theta^2) + \Omega(\theta p_r - R p_\theta) + \frac{1}{2}\Omega^2(R^2 + \theta^2) + \frac{1}{2}K_1 R^2 + \frac{1}{2}K_2 \theta^2 .$$

Following Freeman (1966), we decouple the doubly periodic motion into two simple harmonic oscillations or modes by means of the canonical transformation of the co-ordinates:

$$Q_1 = \sigma[(k_2\omega_2 - \Omega)R + p_\theta]$$

$$Q_2 = \sigma[(-k_1\omega_1 + \Omega)R - p_\theta]$$

$$P_1 = (-k_1\omega_1 + \Omega)\theta - p_r$$

$$P_2 = -(k_2\omega_2 - \Omega)\theta - p_r ,$$

where σ is given by

$$1/\sigma = k_2\omega_2 - k_1\omega_1 .$$

Using the following identities,

$$K_1 = \omega_1\omega_2 k_1 k_2$$

$$K_2 = \omega_1\omega_2/k_1 k_2$$

$$2\Omega = \sigma(\omega_2^2 - \omega_1^2)$$

and $\omega_1 k_1 + \omega_2 / k_2 = 2\Omega = \omega_1 / k_1 + \omega_2 k_2$,

we can express the Hamiltonian in terms of the new variables:

$$H(P_i, Q_i) = - \left\{ \frac{1}{2} \frac{\sigma \omega_1}{k_1} P_1^2 + \frac{1}{2} \frac{\omega_1 k_1}{\sigma} Q_1^2 \right\} \\ + \left\{ \frac{1}{2} \frac{\sigma \omega_2}{k_2} P_2^2 + \frac{1}{2} \frac{\omega_2 k_2}{\sigma} Q_2^2 \right\} .$$

Without loss of generality, we put $\omega_2 > \omega_1$, so that $\sigma > 0$.

The low frequency mode then has negative definite energy.

Coupling between the two modes might allow energy to feed from the lower frequency mode to the higher, so increasing the amplitude of both modes. Collisions provide one source of coupling, non-linear effects which have been neglected in the above analysis, another.

REFERENCES

- Abramowitz, M., and Stegun, I.M. 1964, *Handbook of Mathematical Functions*, (Washington: National Bureau of Standards).
- Alfvén, H. 1968, *Icarus*, 8, 75.
- Barbanis, B. 1976, *Astr. Ap.*, 46, 269.
- Bardeen, J.M. 1975, in *IAU Symposium No. 69, Dynamics of Stellar Systems*, ed. A. Hayli (Dordrecht: Reidel), p.297.
- Blackman, C.P. 1979a, *M.N.R.A.S.*, 186, 701.
- Blackman, C.P. 1979b, *M.N.R.A.S.*, 186, 717.
- Born, M. 1926, *The Mechanics of the Atom* (Constable).
- Bosma, A. 1978, Ph.D Thesis, Groningen.
- Brahic, A. 1975, in *IAU Symposium No. 69, Dynamics of Stellar Systems*, ed. A. Hayli (Dordrecht: Reidel), p.287.
- Brahic, A. 1976, *J. Comp. Phys.*, 22, 171.
- Brahic, A. 1977, *Astr. Ap.*, 54, 895.
- Cleary, M.N. 1977, Ph.D Thesis, Australian National University.
- Contopoulos, G. 1973, *Ap. J.*, 181, 657.
- Contopoulos, G., and Mertzianides, C. 1977, *Astr. Ap.*, 61, 477.
- Danver, C-G. 1942, *Lund Ann.*, No. 10, 135.
- de Vaucouleurs, G. 1959, *Handbuch der Physik*, 53, 311.
- de Vaucouleurs, G. 1963, *Ap. J. Suppl.*, 8, 31.
- Dolfus, A. 1968, *L'Astronomie*, 82, 253.
- Franklin, F.A., and Colombo, G. 1970, *Icarus*, 12, 338.
- Freeman, K.C. 1966, *M.N.R.A.S.*, 133, 47.
- Freeman, K.C. 1970, *Ap. J.*, 160, 811.
- Fujimoto, M. 1968, in *IAU Symposium No. 29, Non-Stable Phenomena in Galaxies* (Yerevan: Academy of Sciences of Armenian SSR), p.453.
- Gerola, H., and Seiden, P.E. 1978, *Ap. J.*, 223, 129.
- Goldstein, H. 1959, *Classical Mechanics* (Addison-Wesley).
- Gordon, M.A., and Burton, W.B. 1976, *Ap. J.*, 208, 346.

- Greenberg, R. 1978, *Icarus*, 33, 62.
- Hénon, M. 1959, *Ann. d'Astr.*, 22, 126.
- Hohl, F. 1971, *Ap. J.*, 168, 343.
- Hunter, C. 1972, *Ann. Rev. Fluid Mech.*, 4, 219.
- Huntley, J.M., Sanders, R.H., and Roberts, W.W. 1978, *Ap. J.*, 221, 521.
- Kalnajs, A.J. 1971, *Ap. J.*, 166, 275.
- Kalnajs, A.J. 1972a, *Astrophys. Lett.*, 11, 41.
- Kalnajs, A.J. 1972b, *Ap. J.*, 175, 63.
- Kalnajs, A.J. 1973a, *Ap. J.*, 180, 1023.
- Kalnajs, A.J. 1973b, *Proc. Astron. Soc. Aust.*, 2, 174.
- Kalnajs, A.J. 1976a, *Ap. J.*, 205, 745.
- Kalnajs, A.J. 1976b, *Ap. J.*, 205, 751.
- Kalnajs, A.J. 1978, in *IAU Symposium No. 77, Structure and Properties of Nearby Galaxies*, ed. E.M. Berkhuijsen and R. Wielebinski (Dordrecht: Reidel), p.113.
- Kerr, F.J. and Westerhout, G. 1965, in *Stars and Stellar Systems*, Vol. 5, *Galactic Structure*, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p.167.
- Kormendy, J. 1979, *Ap. J.*, 227, 714.
- Larson, R.B. 1978a, *J. Comp. Phys.*, 27, 397.
- Larson, R.B. 1978b, *M.N.R.A.S.*, 184, 69.
- Lin, C.C. and Shu, F.H. 1964, *Ap. J.*, 140, 646.
- Lin, D.N.C. and Papaloizou, J. 1979, *M.N.R.A.S.*, 186, 799.
- Lin, D.N.C. and Pringle, J.E. 1976, in *IAU Symposium No. 73, Structures and Evolution of Close Binary Systems*, ed. P. Eggleton, S. Mitton and J. Whelan (Dordrecht: Reidel) p.237.
- Lindblad, B. 1948, *M.N.R.A.S.*, 108, 214.
- Lindblad, P.O. 1960, *Stockholms Obs. Ann.*, 21, No. 4.
- Lynden-Bell, D. 1973, in *Dynamical Structure and Evolution of Stellar Systems*, ed. L. Martinet and M. Mayor (Geneva Observatory), p.91.

- Lynden-Bell, D. and Kalnajs, A.J. 1972, *M.N.R.A.S.*, 157, 1.
- Mezger, P.G. 1975, in *Structure and Evolution of Galaxies*, ed. G. Setti (Dordrecht: Reidel), p.143.
- Miller, R.H. 1976, *Ap. J.*, 207, 408.
- Miller, R.H. 1979, *Ap. J.*, 227, 407.
- Miller, R.H., Prendergast, K.H. and Quirk, W.J. 1970, *Ap. J.*, 161, 903.
- Oort, J.H. 1962, in *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: Benjamin), p.234.
- Pagel, B.E.J. 1978, General Lecture at 4th European Regional Meeting in Astronomy, Uppsala.
- Pottasch, S.R. 1970, in *IAU Symposium No. 39, Interstellar Gas Dynamics*, ed. H. Habing (Dordrecht: Reidel), p.272.
- Prendergast, K.H. 1962, in *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: Benjamin), p.217.
- Roberts, W.W. 1969, *Ap. J.*, 158, 123.
- Roberts, W.W., Roberts, M.S. and Shu, F.H. 1975, *Ap. J.*, 196, 381.
- Rots, A.H. 1975, *Astr. Ap.*, 45, 43.
- Rots, A.H. and Shane, W.W. 1975, *Astr. Ap.*, 45, 25.
- Salpeter, E.E. 1977, *Ann. Rev. Astr. Ap.*, 15, 267.
- Sandage, A. 1961, *The Hubble Atlas of Galaxies* (Washington: Carnegie Institute of Washington).
- Sanders, R.H. and Huntley, J.M. 1976, *Ap. J.*, 209, 53.
- Sanders, R.H. 1977, *Ap. J.*, 217, 916.
- Shu, F., Milione, V. and Roberts, W.W. 1973, *Ap. J.*, 183, 819.
- Simonson, S.C. 1976, *Astr. Ap.*, 46, 261.
- Sørensen, S.A., Matsuda, T. and Fujimoto, M. 1976, *Ap. Space Sci.*, 43, 491.
- Spitzer, L. 1978, *Physical Processes in the Interstellar Medium* (New York: Wiley).
- Stock, J. 1955, *A.J.*, 60, 216.
- Toomre, A. 1963, *Ap. J.*, 138, 385.
- Toomre, A. 1977, *Ann. Rev. Astr. Ap.*, 15, 437.
- Trulsen, J. 1972a, *Ap. Space Sci.*, 17, 241.

- Trulsen, J. 1972b, *Ap. Space Sci.*, 17, 330.
- Trulsen, J. 1972c, *Ap. Space Sci.*, 18, 3.
- Ulam, S.M. 1968, *Bull. Astr. 3-ième Ser.*, 3, 265.
- Valdes, F. 1978, *Publ. Astr. Soc. Pac.*, 90, 397.
- Visser, H.C.D. 1978, Ph.D Thesis, Groningen.
- Weaver, H. 1970, in *IAU Symposium No. 38, The Spiral Structure of our Galaxy*, ed. W. Becker and G. Contopoulos (Canberra: Australian Academy of Sciences), p.126.
- Woodward, P.R. 1975, *Ap. J.*, 195, 61.
- Zang, T.A. and Hohl, F. 1978, *Ap. J.*, 226, 521.