DISSIPATION IN BARRED GALAXIES: THE GROWTH OF BULGES AND CENTRAL MASS CONCENTRATIONS

DANIEL PFENNIGER Geneva Observatory

AND

COLIN NORMAN Johns Hopkins University; and Space Telescope Science Institute Received 1990 January 11; accepted 1990 May 3

ABSTRACT

Basic dynamical mechanisms that produce an amplification of the accretion rate of gas clouds into the central regions of barred galaxies, and their subsequent effects on the evolution of barred galaxies, are discussed. Weakly dissipative orbits, representing gas clouds, are computed in a barred galaxy model with a central mass concentration such as a black hole, or a secondary small inner bar. We find amplified accretion across resonances that is especially rapid inside the Inner Lindblad Resonance, large excursions outside the galactic plane, and the existence of nontrivial attractors like strange (chaotic) attractors or limit cycles. The underlying physical mechanisms are, in general, due to the presence of broad horizontal and vertical resonances through which weakly dissipative particles can rapidly traverse. The two principal physical implications are first that the growth of a significant central mass concentration in a barred galaxy induces broad radial resonance regions which can act to enhance the fueling rate of AGNs and starbursts. Second, the broad vertical resonances allow stars to diffuse into the bulge from the disk, indicating the possibility that a significant component of all bulges has formed in this way and should possess a metal-rich rotationally supported component. Secular evolution along the Hubble sequence due to increasing bulge-to-disk ratio with time is a natural consequence of this intrinsic resonant heating.

Subject headings: galaxies: internal motions — galaxies: structure — stars: stellar dynamics

I. INTRODUCTION

The majority of disk galaxies are barred or ovally distorted in the region where their rotation curve is steep (see de Vaucouleurs 1963; Kormendy 1982; Gerhard, Vietri, and Kent 1989). Dynamical studies of barred galaxies have shown that bars are, in general, rapidly rotating and are principally supported by quasi-periodic orbits elongated along the bar (Contopoulos and Papayannopoulos 1980; Athanassoula et al. 1983; Pfenniger 1984b; Teuben and Sanders 1985). N-body simulations show that nondissipative bars are robust and can persist over at least a Hubble time (Sparke and Sellwood 1987; Combes et al. 1990). A distinctive feature of barred galaxy dynamics is that a significant fraction of the phase space is occupied by stochastic orbits (Athanassoula et al. 1983; Pfenniger 1984a, hereafter P84). This is seen not only in singleparticle studies of the phase space but also in orbit analysis of N-body simulations of bars (Pfenniger and Friedli 1991). The chaotic behavior is due to the presence of many strong resonances such as corotation or Inner Lindblad Resonance (ILR) and the apparent absence of additional global integrals of motion beside the Hamiltonian (or Jacobi) integral (Contopoulos 1983).

Gas flows in barred systems are particularly interesting for the studies of the long-term evolution of barred galaxies. Since *N*-body bars are essentially stable over a Hubble time, dissipative effects due to the gas may alter this stability over a shorter time scale. Earlier studies (see Prendergast 1983) have shown that significant gas flows and associated mass transfer can be achieved; this may indicate the possibility of long-term evolution along the Hubble sequence. A number of numerical experiments suggest that when a galaxy develops a bar, some fraction of its gas accretes rapidly into the center (e.g., Roberts, Huntley, and van Albada 1979; van Albada and Roberts 1981; Schwarz 1981; Athanassoula 1989), though the effect of star formation has been often neglected. In addition, it has been suggested many times that the observed activity in the nuclei of disk galaxies is correlated with the presence of a bar or oval distortion (Adams 1977; Simkin, Su, and Schwarz 1980; Shlosman, Frank, and Begelman 1989; Phinney 1990), and the presumed reason for this is the build-up and fueling of a large mass concentration in the nucleus as a consequence of the overall gas inflow. NGC 1068 is a relevant example here with its central active nucleus, its inner bar, and its population of orbiting molecular gas clouds in the environment of the bar (Scoville *et al.* 1988).

We have undertaken to study a system with the typical properties described above. Our aim is to understand the dynamics of the central regions of a barred galaxy with a central mass concentration such as a black hole. Clouds are modeled as orbiting points with some degree of dissipation and drag. This latter assumption requires some justification. Modeling the interstellar gas is not a trivial task (Combes and Gerin 1985). Considering its different phases (Spitzer 1978; McKee and Ostriker 1977; Norman and Ikeuchi 1989), which imply widely different time scales, its ability to transform itself into stars and vice versa, its clumpiness and turbulent behavior (Scalo 1987), a continuous representation of the gas entailed by using the Navier-Stokes equations is not obviously more justified than a discrete representation using sticky particles. In order to show which physics is relevant, it is necessary to make simplified models mimicking the physical properties that we wish to elucidate.

..363..391P

1990ApJ.

First, we can reasonably assume that the gas is not normally gravitationally dominant, which allows us to neglect selfgravitation in most cases. Second, the dissipation time scale has to be generally long with respect to the local orbital rotation time; otherwise, gas would not be observed to move for many turns on circular orbits, as in disk galaxies, or on oval orbits, as in SB galaxies with rings. Third, a further simplifying aspect is that the clumpiness of the interstellar gas suggests a representation by a collection of weakly interacting particles, instead of a continuum. Inferred cloud lifetimes are short, of the order of 10⁷ yr (Vogel, Kulkarni, and Scoville 1988), and therefore when we speak about a cloud here orbiting many turns, it is a succession of hypothetical clouds that have been formed and destroyed many times but that follow the orbit of a canonical nonevolving cloud having the properties of a slightly dissipative, ballistically orbiting particle. In summary, the physical conditions in the interstellar medium are such that the total energy density of all the energy types susceptible to interact is generally dominated by the gravitational and rotational kinetic energy; the Navier-Stokes equations can be viewed then as a dissipative perturbation of the equations of collisionless stellar dynamics.

Therefore, as an initial approach to study the essential physics, we assume that the problem can be reduced to describing the motion of weakly dissipative test particles, with various friction laws, in different barred galaxy potentials. This kind of system belongs to the class of dissipative systems in the limit of vanishingly small dissipation, close to the well-studied systems obeying purely Hamiltonian mechanics. In this way, we can examine the primary physical factors which control how and why gas accretes in bars, without being distracted by further complications such as collective or self-gravitational effects, or star formation. Also, we can check if the specific way in which dissipation occurs is crucial or not.

We show how this generic combination of possibly chaotic orbits with dissipation leads to two important related effects. First, the inward flow of gas is greatly enhanced in the regions associated with broad resonance regions formed by the central mass concentration and bar interaction. Second, the central mass concentration broadens the region over which orbital energy can be pumped into the stellar orbits in the vertical region. The dissipation allows particles initially located in the disk to explore a large volume of phase space and to penetrate into resonant regions. In this way, significant heating of the disk up into the bulge can occur.

In this paper it is useful to comment on some general properties of weakly dissipative, almost Hamiltonian systems (§ II) showing how the effects of a weak dissipation may be amplified in a resonance region. In § III the orbit analysis is done for a singly barred potential and in § IV for a doubly barred potential. In § V we discuss in detail the astrophysical implications and give a summary and conclusions in § VI.

II. WEAKLY DISSIPATIVE SYSTEMS

In this section we consider the general properties of slightly dissipative, almost Hamiltonian systems. For a perceptive review, see Schmidt (1987). This limiting case is important in general because natural systems frequently have some weak dissipation that limits the time over which the Hamiltonian system description is applicable. Galaxies are in a first approximation dissipationless over many rotation periods; however, over a Hubble time, dissipation, like the effect of water dripping on a stone, must be considered.

a) General Properties of Dissipative Systems

In the broadest meaning, dissipation in dynamical systems corresponds to the contraction of phase space volumes with time. Energy dissipation is a particular but common case. In general coordinates, q_i , \dot{q}_i , i = 1, ..., N, where N is the number of degrees of freedom of the system, the forces in the system can be split into conservative and dissipative forces. The conservative forces derive from a potential V, and can be inserted into a Lagrangian $L(q_i, \dot{q}_i) = T(q_i, \dot{q}_i) - V(q_i, \dot{q}_i)$, where T is the kinetic energy. We describe the nonconservative forces by $F_j(q_i, \dot{q}_i), j = 1, ... N$. In the next sections, they will be supposed to be proportional to a small and positive constant γ . The corresponding Lagrange's equations of motion read (Goldstein 1971, §§ I-V),

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = F_i , \quad i = 1, \dots N .$$
(1)

By the Legendre transformation,

$$H(q_{i}, p_{i}) = \sum_{i=1}^{N} p_{i} \dot{q}_{i} - L(q_{i}, \dot{q}_{i}),$$

where $p_i = \partial L / \partial \dot{q}_i$, we obtain the Hamiltonian H. The equations of motion in canonical coordinates, q_i , p_i , become,

$$\dot{q}_{i} = \frac{\partial H}{\partial p_{i}}, \qquad i = 1, \dots N . \qquad (2)$$
$$\dot{p}_{i} = -\frac{\partial H}{\partial q_{i}} + F_{i},$$

We recognize Hamilton's equation with the additional dissipative terms F_i . By differentiating H with respect to the time, and using equation (2), we obtain that if all the friction forces F_i have a sign opposed to the respective velocities \dot{q}_i , then H is a monotonically decreasing function of time,

$$\dot{H}(q_i, p_i) = \sum_{i=1}^{N} \left(\frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i \right) = \sum_{i=1}^{N} F_i \dot{q}_i < 0.$$
(3)

This inequality is not necessarily true when H is an explicit function of time, $H = H(q_i, p_i, t)$, as in § IV.

Linearizing equation (2) around a given solution $q_i(t)$, $p_i(t)$, we get the linear differential equations for the variations δq_i , δp_i ,

$$\begin{split} \delta \dot{q}_{i} &= \sum_{j=1}^{N} \left(\frac{\partial^{2} H}{\partial p_{i} \partial q_{j}} \, \delta q_{j} + \frac{\partial^{2} H}{\partial p_{i} \partial p_{j}} \, \delta p_{j} \right), \quad i = 1, \dots, N \\ \delta \dot{p}_{i} &= -\sum_{j=1}^{N} \left(\frac{\partial^{2} H}{\partial q_{i} \partial q_{j}} \, \delta q_{j} + \frac{\partial^{2} H}{\partial q_{i} \partial p_{j}} \, \delta p_{j} \right) \\ &+ \sum_{j=1}^{N} \left(\frac{\partial F_{i}}{\partial q_{j}} \, \delta q_{j} + \frac{\partial F_{i}}{\partial p_{j}} \, \partial p_{j} \right), \quad i = 1, \dots, N \; . \end{split}$$
(4)

Noting the vector of the variations $Z \equiv (\delta q_i, \delta p_i)$, we can rewrite this equation as

$$\dot{\boldsymbol{Z}} = \boldsymbol{A}[\boldsymbol{q}_i(t), \, \boldsymbol{p}_i(t)]\boldsymbol{Z} \,, \tag{5}$$

where A is the $2N \times 2N$ matrix containing the derivatives of the right-hand side of equation (4). The Wronskian $W[q_i(t), p_i(t)]$ of this linear system of differential equations, the determinant of A, expresses the evolution of a small phase space volume around the solution q(t), p(t) (Arnold 1974, § 28).

Its rate of change is the trace of A. We obtain,

$$\frac{1}{W}\frac{dW}{dt} = \sum_{i=1}^{N} \frac{\partial F_i}{\partial p_i}.$$
(6)

As expected, the conservative part containing H has disappeared. If the F_i are proportional to a constant γ , then the rate of contraction of phase space volumes is also proportional to γ . A general property is that the frictional forces F_i have to depend on the respective momenta p_i for leading to dissipational effects. For the particular case that each F_i depends only linearly on p_i , then the rate of contraction $d \ln W/dt$ is a constant along any orbit. Liouville's theorem is a particular case for which the rate of contraction vanishes. Such dissipative terms that are linear in the momenta are rather special, and without additional physical motivation one should adopt more general forms in order to get generic results.

b) Phase Space Structure

The phase space structure of dissipative systems is radically different from that of Hamiltonian systems, which is known to be made up principally by quasi-periodic and chaotic orbits. As soon as some dissipation is introduced, most of the quasiperiodic and chaotic orbits of a Hamiltonian system are destroyed. In the long time limit, the trajectories of bounded dissipative systems approach a small number of asymptotic orbits called attractors that have only three generic types. In phase space, attractors may be either fixed points, limit cycles corresponding to periodic motion, or strange attractors which are particular invariant chaotic orbits with a fractal dimensionality (see, e.g., Lichtenberg and Lieberman 1983, hereafter LL; Schmidt 1987). Fixed points in the weakly dissipative case correspond to stable fixed points in the Hamiltonian limit, whereas limit cycles correspond to particular degenerate stable quasi-periodic orbits (as, e.g., in Pfenniger 1985). The precise Hamiltonian limit of strange attractors has been considered only recently (Chen, Györgyi, and Schmidt 1986, 1987; Schmidt 1987). Apparently, by a cascade of bifurcations, these attractors dissolve into an infinity of pieces as the Hamiltonian limit is reached.¹ Strange attractors need at least three phase space variables. So, in the long time limit, sufficiently simple systems, i.e., systems which can be decomposed locally into subsystems of dimension two or less, tend to either fixed points or limit cycles, or eventually combinations of these. Orbits in dissipative systems without strange attractors exhibit a regular behavior after sufficiently long time. In some cases, one can rule out the existence of strange attractors because the dissipative process monotonically decreases the energy toward a simple minimum, that is, typically a fixed point or a periodic orbit.

Many nearly integrable physical systems perturbed by a weak dissipation tend to spend a long time in the quasi-

¹ If $J_n < 1$ is the value of the Jacobian of the transformation matrix associated with the motion (a Poincaré surface of section) at which a strange attractor made up by 2^n pieces bifurcate into a 2^{n+1} -piece attractor, then, as dissipation is decreased, the J_n accumulate onto the Hamiltonian value $J_{\infty} = 1$, and a universal law holds,

$$\lim_{n \to \infty} \frac{J_n - J_{n-1}}{J_{n+1} - J_n} = 2$$

(Chen, Györgyi, and Schmidt 1986, 1987). The eigenvalue 2 is a universal number, distinct from the Feigenbaum numbers 4.67 ... and 8.72

periodic or regular regime and a relatively short time in the chaotic regime. This has been beautifully illustrated by studies of binary stars (Savonije and Papaloizou 1984) or planetary satellites (Wisdom 1987) undergoing tidal friction. The passage through resonances, where motion is chaotic, is rapid, and most of the time is spent in the quasi-periodic state.

Our understanding of the physics of the dissipative orbits considered in this paper is based on two observations. First, a chaotic region is essentially filled with many resonance regions (frequently of high order) threaded by chaotic orbits (see LL, Fig. 3.5). Second, studies of dissipative systems in one space dimension, which are integrable, show that when crossing a resonance, the relevant action decreases with a discontinuous jump (Parson 1986). In terms of the action, a quantity which is well defined only in the regular regions, such a resonance crossing can be characterized by an increase in the rate of action decay. In simple dissipative oscillators, the equivalent statement is that energy is lost faster at resonances. The same is true for other actions. This resonance jumping phenomenon implies that less time is needed to cross chaotic multiresonant regions of phase space similar to those encountered in § III. Particles will spend relatively more time in regular regions. This is the main reason that dissipative orbits in nearly integrable systems are seen more often in regular regions; it is not merely that in such systems chaotic regions are rarer than the regular ones.

General systems are not integrable, and dissipation laws have rarely the simplifying property to be linear in the momenta. However, we expect that in the weakly dissipative limit, most of the behavior will remain similar to the Hamiltonian case for a limited time. The time scales to compare are the time to cross a resonance of a given width and the orbital time. If, in one turn, an orbit dissipates sufficiently to cross a resonance, clearly the resonance will have no effect. The higher the order of the resonance is, the narrower it is, and the weaker the dissipation has to be in order that an orbit can feel the resonance. Typically, around a resonance, phase space, as viewed in a Poincaré surface of section, is made up by a series of regular islands separated by a chaotic layer. If a slight dissipation is then applied, an initial orbit in the chaotic region is either trapped by islands, in which case the motion is ultimately periodic or remains chaotic by being trapped by a strange attractor. Numerical experiments show (e.g., Schmidt 1987) that when dissipation is sufficiently weak, strange attractors catch exceedingly few orbits, since these attractors dissolve before the Hamiltonian limit is achieved. Also, the larger the islands, the larger is their probability to attract trajectories. High-order resonances of small width are wiped out. Assuming a uniform rate of contraction of phase space across a resonance, and noting that the width of the chaotic region at the resonance is reduced by the presence of the islands, the phase space flow crossing the resonance must proceed faster across the gaps left between the islands. One can picture the resonance crossing by a fluid in a pipe which is forced to pass through a region occupied by slightly porous stones of various sizes. The velocity of the flow far from the stone region represents the rate of dissipation far from the resonance. By continuity, the flow has to speed up around the stones. The larger the dissipation rate or the smaller the islands, the more porous the islands become, up to the point that they do not block the flow at all. A given dissipation rate can erase resonances of higher order, while the lower order ones continue to act to increase locally the rate of action loss.

c) One-dimensional Time-dependent Dissipative Pendulum

We now give a concrete illustrative example of the phenomenon of resonance jumping. The crossing of a resonance due to a weak dissipation is studied here by a simple model of a one-dimensional pendulum. This model is relevant here because the phase space of the pendulum is known to characterize typical resonances (see LL, § 1.3). The phase space of a conservative pendulum is generated by the equations of motion,

$$\begin{aligned} x &= v , \\ \dot{v} &= -\sin 2\pi x , \end{aligned} \tag{7}$$

and is everywhere regular because it is time-independent. Phase space is periodic in x with an interval $\Delta x = 1$. Figure 1 shows a surface of section (x, v) at constant time intervals $\Delta t = 1$. It represents a small region of the typical phase space around a high-order resonance inside a much larger island, for which the curvature of the chain of small islands is negligible. The chain of islands can then be unfolded and the topology of the pendulum's phase space is obtained. The width of the resonance is given by the amplitude in v which keeps the motion oscillating ($\Delta v \approx \pm 0.8$ in Fig. 1).

Normally, typical resonances have a chaotic layer around them. Chaotic motion is introduced by a time-dependent modulation of the natural frequency of the pendulum. Time dependence extends phase space by two dimensions (LL, \S 1.2b). In a usual pendulum, this could be achieved, e.g., by modulating either the length of the pendulum or the strength of the gravity field. Choosing a simple modulation law of period 1, the equations of motion become

$$\dot{x} = v ,$$

$$\dot{v} = -(1 + \epsilon \cos 2\pi t) \sin 2\pi x .$$
(8)

For a small ϵ a stochastic layer develops around the separatrix between oscillating and librating motion (Fig. 2).

Dissipation is introduced by a dependence on v in the equation of motion. In order to simulate a resonance crossing, the



FIG. 1.—Surface of section (x, v) of the unperturbed pendulum showing a regular phase space separated into a librating (*top* and *bottom*) and oscillating (*middle*) regions. The limit curve between the two types of motion is the separatrix.



FIG. 2.—Surface of section (x, v) of the pendulum with a time-dependent perturbation ($\epsilon = 0.1$). A stochastic layer develops around the separatrix, and new islands appear at x = 0, $v = \pm 1$, due to the time-dependent perturbation.

friction should not vanish when v = 0. We have retained the following nonlinear friction law:

$$F_{\rm fric} = -\gamma |v - v_0| (v - v_0) , \qquad (9)$$

where v_0 and γ are constants. The constant v_0 corresponds to the velocity v at which friction vanishes and has the effect, when negative, of tending to pull orbits starting at large positive v across the resonance at v = 0. The positive constant γ is the friction coefficient. The final equation of motion reads

 $\dot{x} = v$,

$$\dot{v} = -(1 + \epsilon \cos 2\pi t) \sin 2\pi x - \gamma |v - v_0| (v - v_0)$$
. (10)

The rate of contraction of phase space volume reads (eq. [4] and [6]),

$$\frac{\dot{W}}{W} = -2\gamma \left| v - v_0 \right|, \qquad (11)$$

so is proportional to y and is negative, but vanishes at $v = v_0$.

We have run many orbits starting at v > 1 and crossing the resonance at v = 0 with different initial conditions and sets of parameters. A typical orbit representing a weakly dissipative and slightly nonintegrable case is shown in Figure 3. This orbit starts at the top of the surface of section (x, v) and slowly sinks to the bottom, crossing the resonance region. The big regular island region is clearly avoided by the orbit. The time-averaged velocity $\langle v \rangle$, averaged here over 50 periods, is closely related to the action outside the resonance, in the region of rotation. Figure 4 shows that the velocity drops abruptly at the resonance.

In more general systems, such as those discussed below, major resonances are surrounded by numerous higher order resonances. Each crossing of a high-order resonance will be associated with similar jumps, the amplitudes of which are proportional to the corresponding resonance widths. The rate of action loss will be significantly accelerated in nonintegrable systems which have a high density of resonances of various orders.

However, if the system is sufficiently chaotic and dissipative, orbits may be caught by a strange attractor, in which case the resonance is never crossed, the orbits jumps back and forth



No. 2, 1990

FIG. 3.—Surface of section (x, v) of the pendulum with both time-dependent and dissipative perturbations ($\epsilon = 0.001$, $\gamma = 0.00001$, $v_0 = -10$). Contrary to the two previous figures, only one orbit is plotted here, which, due to dissipation, explores phase space. It starts at the top ($x_0 = 0.4$, $v_0 = 1.5$) and ends at the bottom, avoiding the large central island. Note the two comet-like series of points at $v = \pm 1$; they are produced by new small bent islands coming from the time-dependent and dissipative perturbations.

across the resonance erratically. Such a case is shown in Fig. 5. In § III strange attractors are ruled out because H decreases monotonically (cf. eq. [3]), while in § IV, they are *a priori* possible, and indeed found.

III. DISSIPATION IN A BARRED POTENTIAL

a) Bar Model and Orbits

We illustrate the preceding general considerations by computing the orbital decay rate of slightly dissipative single particles, representing idealized gas clouds, onto a central mass, in the three-dimensional barred galaxy model used in earlier studies (P84; Pfenniger 1984b, 1985). The potential is made up of an axisymmetric potential $\Phi_{MN}(R, z)$ (Miyamoto and Nagai



FIG. 4.—Decay of the average velocity $\langle v \rangle$ of the orbit in Fig. 3 as a function of time t. Note the abrupt jump through the resonance at $\langle v \rangle = 0$. The two bumps at $\langle v \rangle = \pm 1$ are due to the new small bent islands visible in Fig. 3. The velocity average is taken over 50 time units, which somewhat decreases the effective slope of the jump around $t \approx 890$.

FIG. 5.—Decay of the average velocity of an orbit in a strongly chaotic regime ($\epsilon = 1$). Instead of jumping through the resonance as in Fig. 4, it is apparently locked by a strange attractor.

1975; Binney and Tremaine 1987, hereafter BT, p. 43) representing a disk-spheroid component of horizontal scale length A + B, vertical scale-height B and mass M_D ,

$$\Phi_{\rm MN}(R, z) = -\frac{GM_D}{\sqrt{R^2 + (A + \sqrt{B^2 + z^2})^2}},$$
 (12)

with A = 3, B = 1, $GM_D = 0.9$, $R^2 = x^2 + y^2$, and a triaxial n = 2 Ferrers bar (Ferrers 1877; P84) of semiaxes a, b, and c, and mass M_B , the density of which is

$$\rho(x, y, z) = \frac{105M_B}{32\pi abc} (1 - m^2)^2 , \text{ for } m < 1 ,$$

= 0 , otherwise , (13)

where

$$m^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$
, and $a = 6, b = 1.5, c = 0.6, M_B = 0.1$.
(14)

The bar corotation radius R_{CR} is located at the bar end ($R_{CR} = a$). The associated bar rotation frequency $\Omega_b = 0.05471$ corresponds to a bar rotation period of about 115. In this section, a small Plummer sphere of mass M_{bh} and radius $R_{bh} = 0.01$, representing a dense mass condensation is also put at the center. The length unit can be chosen conveniently to be the kpc, and the time unit, the Myr.

The method is straightforward. We integrate numerically orbits starting at rest in the rotating frame of the bar close to the Lagrangian points L_1 (near the end of the bar) for various dissipation laws. We use a th-8th order Runge-Kutta-Fehlberg integrator with adaptative time step and 16 digit arithmetic, which ensures that numerical errors are kept to a minimum, contrary to what is feasible, for instance, in N-body type codes. The value of the Hamiltonian at L_1 and at zero velocity corresponds to the highest value for which the orbits are necessarily bounded and decay toward the center. The initial position in z is chosen to be slightly above the galactic plane (20 pc) in order to observe the effects of vertical instabilities. As long as dissipation is weak, the orbits slowly visit the basin of attraction made by the bar potential and are influenced successively by the different orbit families described in the conservative case (P84). We deliberately choose a dissipation rate that is much lower than any plausible one in order to let the particles explore in detail the basin of attraction. The integration is stopped either when the time t amounts to 125,000 (about 10 Hubble times), or when the orbit is confined inside the central Plummer sphere. The orbit coordinates are sampled at regular time intervals $\Delta t = 25$, which has been fixed by the available mass storage capacity.

The dissipation laws that have been considered include dynamical friction (BT, eq. [7-17]), friction laws proportional to v^n , $n \ge 1$, where v is the velocity in the rotating frame, and more complicated laws depending on the projection of the velocity vector parallel and/or perpendicular to the effective potential. All these laws dissipate kinetic energy in the rotating frame, without special constraints, such as, for example, the conservation of angular momentum, since such quantities are not conserved in barred potentials anyway. Since the conditions of inequality (3) are fulfilled, H is a decreasing function of time. As a consequence, in all the cases considered the orbits converge toward the only attractor existing in this potential, namely the center.² The results have been verified to be qualitatively independent of the dissipation law, as long as dissipation is everywhere weak, i.e., much slower than the local orbital period. This is not surprising since weak dissipation does not change either the attractors of the system or substantially alter the structure of the basins of attraction. In the following, we therefore display only the orbits having a friction force $F_{\rm fric}$ tangent to v and proportional to v^2 ,

$$\boldsymbol{F}_{\rm fric} = -\gamma \boldsymbol{v} \boldsymbol{v} \,, \tag{15}$$

where v is the velocity in the rotating frame of the bar, and γ a small constant. In terms of canonical coordinates p_x , p_y , and p_z , we have, $v = (p_x + \Omega_b y, p_y - \Omega_b x, p_z)$. The rate of contraction of phase space volume becomes (eq. [6])

$$\frac{\dot{W}}{W} = -4\gamma v . \qquad (16)$$

If γ were to be a function of the total local density, then the assumption of weak dissipation would be invalid in the central regions, not only because the density is high there, but also because the local orbital period is short. On the contrary, as is easily checked, the velocity range does not vary much in most of the region that we are examining.

b) Unperturbed Case

We first consider the unperturbed case without a central mass concentration $(M_{bh} = 0)$. The orbital decay time is directly dependent on the actual assumed dissipation parameter γ , which is presently far from being known accurately. Figure 6 shows that the decay of H as a function of γt is nearly the same for different γ . The total decay time is therefore not particularly physically relevant, in particular whilst a satisfying representation of γ is not available. However, the ratios of the



FIG. 6.—Initial decay of H as a function of γt in the potential of § III, without central mass concentration, but different γ 's. The average slopes of $H(\gamma t)$ are globally similar. The curves with the larger oscillations correspond to the larger γ 's. Note the similarity of the curves, a factor of 4 of magnification of γ produces similar but 4 times larger oscillations.

orbital decay times of the perturbed and unperturbed problems indicate the relative increase of the inflow rate when the central mass is increased. We have retained a very weak dissipation relative to the galactic time-scale ($\gamma = 0.0005$), allowing the orbits sufficient time to explore in detail the basin of attraction. Without the central mass concentration, the dissipation time-scale needed to reach the very central region amounts to typically 10 Hubble times, or 1000 bar rotation periods.

In Figures 7a, 7b, and 7c (top frames), the radius R, the height |z|, and the Hamiltonian value H for the case without a central mass are plotted as a function of time t. The positions of major axisymmetric horizontal and vertical resonances for direct circular orbits are indicated,

$$m = \frac{\kappa}{\Omega - \Omega_b}, \quad n = \frac{\nu}{\Omega - \Omega_b}, \quad n, m = 2, 3, \dots, \quad (17)$$

where Ω is the circular orbit frequency, κ and v the radial and vertical epicyclic frequencies. The frequencies Ω , κ , and ν are computed in the azimuthally averaged potential. The axisymmetric resonances correspond only approximately to the actual orbit resonances in the barred potential, because the bar perturbation is substantial, and some periodic orbit families are sometimes far from circular. For m or n small integers, a faster radial decay, or rapid vertical excursion respectively is apparent in Figure 7. At the beginning of the integration, the orbit is temporarily attracted toward a family of stable periodic orbits, which, just inside the corotation radius, is the 4/1family (family O in P84). It follows this family in the sense of decreasing H until the family ceases to exist at the next large resonance, the m = 4 one. It enters then a chaotic region associated with this and other resonances, where it decays faster, both in the R(t) and H(t) curves. This is especially visible in the H(t) curve, which presents a corner at the transition. Since H decays monotonically, strange attractors are ruled out. As expected from the discussion in § II, the decay is especially fast in the region of resonances. Figure 8 shows a magnification of the curve H(t) in this chaotic region. At sufficient magnification, it is apparent that the curve is not smooth, though monotonically decreasing, but made of numerous small steps

391P

363.

L990ApJ

² In fact the Lagrangian points L_4 and L_5 can also be attractors, if stable, for particles with an initial *H* larger than the value of *H* at $L_{4,5}$ at zero velocity. However, due to their associated high energy, these points do not belong to the part of phase space which is bounded by the Jacobi integral, and, as suggested by *N*-body simulations, in self-consistent barred galaxies these points could well be only marginally stable in the conservative case, i.e. no longer attractors in the dissipative case (Pfenniger 1990).



 $\ensuremath{\textcircled{O}}$ American Astronomical Society $\ \bullet$ Provided by the NASA Astrophysics Data System



398

FIG. 8.—Magnification of the curve H(t) as the particle leaves the regular 4/1 oval orbit family on the left and enters, at t > 8400, into a large chaotic region. The curve in the chaotic region is steeper than in the regular region, and shows numerous steps corresponding to high order resonances.

or oscillations (see also Fig. 6), corresponding to the crossing of high-order resonances. This explains why the orbit moves quickly through this chaotic region. After it reaches the zone of the regular quasi-periodic orbits supporting the bar (family B in P84), its orbital radius decays much more slowly. When the orbital period is comparable to the sampling interval one observes a slowing down of the oscillations in Figure 7, due to beats between the orbital frequency and the sampling frequency. Such an aliasing is physically irrelevant; only the average radius is interesting. The motion in the vertical direction is slightly amplified at $t \approx 10,000$ due to the presence of vertical instabilities of this barred system associated with the n = 4 ultraharmonic resonance (UHR) and higher order resonances, as mentioned in P84.

c) Effect of a Central Mass Concentration

Next, we examine how the accretion rate is modified by different central point masses $M_{\rm bh}$. In general, a peak of density at the center produces both horizontal and vertical resonances in the epicyclic motions. This is illustrated in Figure 9 which shows that even if the barred system without a central singularity has no ILR, a central point mass creates ILRs and higher order resonances, both horizontally and vertically. The radii of resonances increase as the central point mass grows. Figure 10 shows the radii of the horizontal and vertical resonances in different Miyamoto-Nagai disks with a point mass at the center and different scale heights. As the central mass increases, the resonance radii increase especially rapidly at small masses. The second effect is that the flatter the disk, the closer to the center are the vertical resonances. This is crucial for the understanding of the growth of bulgelike structures from flat disks, as discussed in more detail in § V. In the Appendix is shown that very general constraints such as the positivity of mass imply the existence of vertical resonances, both for direct and retrograde circular orbits in galactic disks subject to small nonaxisymmetric deformations. The width of resonances is principally determined by the amplitude of the bar perturbation. Since this perturbation is strong, the resonance widths are broad. Although it is difficult to estimate these widths analytically, because they depend not only on the potential derivatives, but also on the orbit shapes, the fact is that the combination of a central mass concentration and the bar give large regions of both horizontal and/or vertical unstable motion associated with the broad resonance regions (see, e.g., Athanassoula et al. 1983, showing the effect of a singular center on the orbits at the horizontal ILR).

Figure 7 displays coordinates of orbits with identical initial conditions to the previously discussed unperturbed case, showing how their dissipation rate is modified by a central point mass. Again, the orbits spend most of the time in the regions where stable periodic orbits can be found in the nondissipative case. Due to the presence of new resonances, the



FIG. 9.—Resonance diagrams of the azimuthally averaged barred potential of § III, without (solid line) and with (dashed line) a central mass concentration of $M_{\rm bh} = 0.001$. Figure (a) shows the usual horizontal resonances curves $\Omega - \kappa/m$, $m = 2, 3, \dots 8$ and Ω as a function of galactocentric radius R, where Ω is the circular orbit frequency and κ the radial epicyclic frequency. Figure (b) shows the vertical resonances curves $\Omega - \nu/n$, $n = 2, 3, \dots 8$ and Ω as a function of R, where v is the vertical epicyclic frequency. The divergence of the central curves implying low order resonances is a general behavior (see the Appendix) due to the transition from a spherical potential at the center to a flattened one at large radii.



No. 2, 1990

FIG. 10.—Dependence of the resonance radii on the horizontal (solid) and vertical (dashed) resonances in a Miyamoto-Nagai disk (A = 3) on the central point mass M_{bh} , as the scale height B of the disk is varied, and for a corotation radius $R_{cr} = 6$ typical for a bar. The respective m and n values are indicated at the top of the curves. Top, B = 1; middle, B = 0.5; bottom, B = 0.2. The flatter is the disk, the closer to the center are the low order vertical resonances.

plunging toward the center is dramatically accelerated by the presence of the point mass. As argued above, the fast contraction of the orbital radius occurs inside resonances, mainly the lower order ones. The approximate radii of resonances in the barred potential is marked in Figure 7 by the radii of the corresponding axisymmetric resonances of the azimuthally averaged model.

The height above the plane |z| is often temporarily amplified by vertical resonances and the interaction with the central mass concentration as illustrated in Figure 7b. Figure 11 shows that the orbits have a very flat initial shape in the disk, due to the initial conditions, which subsequently becomes a more or less spheroidal shape for a time depending on $M_{\rm bh}$ and the dissipation rate. Larger central mass concentrations produce larger excursions in z. This can be understood simply by the fact that, as shown in Figure 10, the larger the central mass concentration, the farther out in radius are the vertical resonances, so that they are encountered when the orbits have a higher energy. Once a particle enters the chaotic sea, its motion is bounded principally by the three-dimensional energy surface corresponding to its present energy, shown in Figure 12. This figure exhibits interesting similarities with real bulges, as we will discuss in § V. This implies that substantial heating of the disk particles will occur in this region. These excursions in z are only temporary here because, unlike real gas clouds potentially forming stars, here the particles dissipate constantly and the final attractor is the center.

Therefore, for a given dissipation rate in a given barred potential, clouds are accreting faster inside resonances, principally the ILR, when a central mass concentration exists already. Since every mass concentration at the center of a rotating bar produces at least one ILR, the accretion onto the center is self-amplifying, unless some modifications of the scenario occur. Once the clouds have crossed the ILR, they can accumulate onto the nearly regular circular orbits around the central mass and form eventually a new disk. As soon as this small disk is massive enough, it becomes gravitationally independent of the large bar and at its own scale may develop its own bar instability (Norman 1988).

IV. DISSIPATION IN A POTENTIAL WITH AN INNER AND OUTER BAR

a) Double Bar Model

Bars inside bars have been frequently observed (e.g., NGC 1291, 1543 in Jarvis et al. 1988), and we now explore this more complicated system. If a small inner bar develops on its own due to a sufficiently strong self-gravity, it is more likely that it rotates at a completely different speed than that of the large bar. The large bar feels only the averaged rapidly rotating potential of the small one, and the small bar, if dense enough, feels only a slowly varying perturbation of its potential by the large one. In order to reduce the chaos produced by the major resonances, a favorable configuration is to locate the ILR of the large bar at the corotation radius of the small bar. Such a situation is still difficult to simulate by fully consistent N-body simulations, as the ranges in masses, lengths, and time scales between the bars are large (more than an order of magnitude). However, it is obvious that for sufficiently small, but dense inner bars, it must be possible to construct such systems, at least for some reasonably long time-scales.

A small bar provides a mechanism for regulating the accretion of gas into the inner regions. From hydrodynamical simulations (e.g., Sanders and Tubbs 1980), a rotating bar repels gas outside its corotation radius. Basically, it does so because the effective potential, like the rim of a crater, is maximum at corotation. In contrast to this, Shlosman, Frank, and Begelman (1989) assume that a small bar increases the accretion rate through the corotation radius. The problem is to cross corotation and to enter the region where the bar can sweep the gas inward. This may occur when sufficient mass is accumulated outside corotation, it becomes progressively more selfgravitating, and our simplified treatment of the dynamics less valid. At some point, the gas self-gravitation can dominate the dynamics and allow crossing of corotation. This may lead to bursts of activity instead of a continuous fueling.

We consider now the same barred potential as in the previous section, except that the central Plummer sphere is now replaced by a second Ferrers bar, much smaller than the large one, and rotating in the same direction. The corotation of the inner bar is close to the ILR of the large bar. The mass of the small bar and all its linear dimensions amount to 10% of those of the large bar. The small bar rotation frequency is about 5.86 times higher than that of the main bar; thus, they are not commensurable. The small bar introduces a periodic time dependence in the reference frame of the main bar which destroys the strict Jacobi integral *H*.

b) Attractors

Again we consider the same kind of weakly dissipative orbits and look for their attractors, starting, as before, at rest near the Lagrangian points L_1 of the main bar. The initial phase differ-





© American Astronomical Society • Provided by the NASA Astrophysics Data System



FIG. 12.—Zero-velocity contour lines, or the effective potential $\Phi_{eff}(R, z) = \Phi(R, z) - \frac{1}{2}\Omega_b^2 R^2$ (solid lines) corresponding to the barred galaxy potential of § III, in the meridional plane (x, z) at y = 0. In other meridional planes, the figure would look nearly the same. Orbits are confined only by the oval zero-velocity surfaces passing inside the corotation radius around x = 6.12, z = 0. Note the similarities of these curves with real bulges. Outside corotation, the zero-velocity curves do not confine orbits, but the bar perturbation decreases, and at least near the galactic plane the potential tends to be separable. The lines orthogonal to the zero-velocity curves (*dashed lines*) suggest the effect of a new quasi integral of motion there. The border between the bulge and thick disk would be determined by the bar corotation radius.

ence of the two bars is set arbitrarily to zero. This has no consequence for the orbits, as their starting point, near the Lagrangian point L_1 of the large bar, is far from the small bar, and the arrival time in the region of influence of the small bar is practically random. In such a time-dependent potential, the center is no longer necessarily the only attractor, but limit cycles or strange attractors are the known *a priori* additional possibilities. Due to the time dependence, the small bar can provide energy to the particle which eventually compensates the frictional losses.

In Figure 13, the orbital radius R, the height |z|, and the Hamiltonian value H are plotted as a function of time t, for different friction coefficients y. In this experiment, the mass of the central component is not varied, and consequently the system can maintain the near coincidence of the radii of the outer bar ILR and the inner bar corotation radius. The variation of γ allows us to investigate the conflict between the tendency to accrete inwards and the repelling effect of the inner bar. For small y's, R and |z| do not decrease the zero for very long times, and H fluctuates in an erratic manner. This is typical for a strange attractor. For $\gamma = 0.002$, we obtain a limit cycle in the plane z = 0 with an oscillating R and nearly constant H. This oval ring is a retrograde periodic orbit, approximately 1.5-2 times larger than the small bar, but inside the large bar and perpendicular to its major axis. In a surface of section (y, \dot{y}) at $x = 0, \dot{x} < 0$, the orbit spirals down toward the limit cycle, which is a point in the surface of section (Fig. 14a). In a frame at rest, the orbital period of this limit cycle turns out to be identical with the inner bar rotation period. For a limited range around $\gamma = 0.002$, the limit cycle can coexist with strange attractors, both in three dimensions and in the plane z = 0. Figure 14b shows a surface of section of a strange attractor confined in the plane z = 0 coexisting with the limit cycle.

After having computed many orbits, the conclusions are as

follows. If the dissipation is sufficiently strong (or the bar strength sufficiently weak) to overcome the repulsion due to the small bar, the attractor is in most cases still the center. For a given bar strength, there exists a dissipation rate below which the repulsion of the small bar is effective enough to prevent most orbits from crossing the small bar's corotation radius. They are attracted toward limit cycles or strange attractors. The energy and angular momentum lost by dissipation are supplied by the small bar. As in the single bar potential, the orbits can reach sufficiently large heights above the galactic plane (1-2 kpc) so that they fill an essentially threedimensional volume, which has morphological similarities to the typical galactic bulges. In Figure 15, the particle positions are plotted at regular time intervals, showing the large excursions outside the galactic plane, the spheroidal shape of the strange attractors, and the oval shape of the limit cycle at $\gamma = 0.002.$

As we have not explored fully the phase space and parameter space of this time-dependent system, attractors other than those described here can exist. We do not expect further fixed points, but other limit cycles or strange attractors are likely. These qualitative results are exploratory and would need further investigations. With the adopted bar mass and axis ratios, the longest dissipation time scale $\tau_{diss} \sim \gamma^{-1}$ necessary to overcome the inner bar repulsion amounts to about 10³, or of order of 9 outer bar rotation periods. If the actual dissipation time scale of the galactic gas is larger or smaller than this critical value, its behavior in the central regions would be very different.

V. ASTROPHYSICAL IMPLICATIONS

From the previous orbit calculations, two principal implications for the evolution of the central regions of galaxies have been found: (1) the radial inflow of mass to the central regions and (2) the vertical disk heating phenomenon that may lead to bulge building. The basic physical phenomenon responsible for the behavior described here is the association of resonances and weak dissipation. The generality of the physical laws considered means that many more situations of astrophysical interest could be relevant. For example, in any axisymmetric galaxy with a central peak of density, nearly all nonaxisymmetric perturbations will broaden the resonances near the center. Bars are especially strong and frequent perturbers, but other factors such as spiral arms or satellites should also contribute to similar evolutionary effects as described below. A different example could be a triaxial elliptical galaxy which would have a nearly integrable potential. A central peak of density, as examined in Pfenniger and de Zeeuw (1989), would also be a resonance broadening perturbation and act to enhance the effect of dissipation.

a) Mass Inflow

First, a rapid inflow of gas clouds in systems with bars and central mass concentration can occur. This follows from the unavoidable presence of low-order resonances throughout the region inside the bar. These resonances are caused by the presence of the central mass concentration and are broadened by the strong bar perturbation. As seen in the calculations of §§ III and IV, the radial accretion rate is boosted inside the resonances. This process tends to build up the central mass concentration and increases the radial extent of the resonance region. This will increase the mass that can rapidly accrete and indicates a self-amplifying process, which may lead possibly to a



FIG. 13.—(a) Radius R, (b) height |z|, and (c) Hamiltonian value H of $\langle n \rangle$ orbit in the potential with outer and inner bars, for different friction coefficients γ . In contradistinction to Fig. 7, we observe that in many cases ($\gamma < 0.0015$), the orbits do not decay to the center, but its coordinates fluctuate in a quasi-stationary fashion, tracing a strange attractor. In the case $\gamma = 0.002$, the coordinates tend toward a regular behavior corresponding to an oval limit cycle.





starburst or AGN phase in the central regions. The process can end when the gas supply is exhausted due to star formation or expelled from the central regions due to energy input from the starburst or AGN.

This analysis is based on studying orbits of clouds which can characterize the flow of a multicomponent and turbulent

medium such as the ISM. More realistic models of the gas flow incorporate the possibility of shocks that are clearly seen in gas-rich galaxies and simulations (Roberts, Huntley, and van Albada 1979; van Albada and Sanders 1982; Athanassoula 1989). The viewpoint taken here is that such shocks merely increase the dissipation in the domain where the clouds



FIG. 14.—(a) Two-dimensional projection (y, \dot{y}) of a Poincaré section at x = 0, $\dot{x} < 0$ of an orbit converging towards a limit cycle, a point. (b) Surface of section (y, \dot{y}) at x = 0, $\dot{x} < 0$, z = 0, $\dot{z} = 0$ of the limit cycle (*large dot*) coexisting with a strange attractor (*cloud of points*).





404

© American Astronomical Society • Provided by the NASA Astrophysics Data System





Fig. 15—Continued

405

Vol. 363

encounter them. As the mean-free path shortens and pressure forces begin to dominate the motion, our simplified description will break down. The only way to analyze this is with realistic simulations of the gas which are beyond the scope of this paper. Clearly, however, this approximation still holds while clouds have slightly dragged orbits between shocks. A review of such previously calculated simulations (e.g., Roberts, Huntley, and van Albada 1979; van Albada and Sanders 1982) shows that the fluid elements still approximately follow purely gravitational orbits between shocks, essentially because the local energy density associated with the gas pressure is still an order of magnitude smaller than the gravitational energy. Another refinement is to consider realistic equations of state, including an accurate representation of the cooling and heating mechanisms. But such complications, which in any case can not yet be properly simulated because the collapse time scale associated with the cold gas component is much too short with respect to the orbital time, will not help us to elucidate the primary factors which determine the global gas flow. A conclusion of Hernquist (1989), using a sophisticated hydrodynamical code, is that his results concerning the accretion of gas in a spiral galaxy are in good agreement with the ones of Carlberg and Freedman (1985), who used a much simpler discrete cloud method. This confirms that the global gasdynamics in galaxies is, in first order, only weakly sensitive to its detailed dissipative properties, essentially because the energy density in the gas pressure is lower than the gravitational and kinetic energy densities, typically by an order of magnitude.

We have seen that an inner bar may repel clouds moving in from the outer bar to form a ring structure outside its corotation radius. This is consistent with previous simulations of gas flow in bars, which repel gas outside corotation, and is also consistent with the sense of angular momentum transfer by nonaxisymmetric perturbations (Lynden-Bell and Kalnajs 1972; Schwarz 1981). A massive central ring or torus around the inner bar may be formed in this way. As the gas dissipates, the inner bar clearly must supply energy and gas angular momentum to the torus. As the mass of this torus builds up, its own self-gravity will begin to dominate the local dynamics, and our simplified description will break down.

b) Disk Heating and Bulge Formation

The second principal astrophysical implication is that at least a significant component of bulges can form from the disk heating due to vertical resonances broadened by nonaxisymmetric perturbations. In the presence of a central mass concentration, these vertical resonances are unavoidable in the central regions as shown in Figures 9 and 10 and in the Appendix. The broadening of the vertical resonances by a bar or an oval distortion is typically of the same importance as the broadening of the horizontal resonances (see P84). Central nonaxisymmetric distortions are frequent, since about $\frac{2}{3}$ of disk galaxies are detected to be SB or SAB, and even the bulge of some galaxies classified S has been inferred to be triaxial (Gerhard, Vietri, and Kent 1989). In the extreme case that a disk galaxy could form without any bulge and be very flat (as suggested by the Sm, Sd galaxies), some slight dissipation would bring some mass to the center and create horizontal and vertical resonances. Dissipation would be also favorable for cooling the disk and leading to a bar instability, broadening the resonances. As either the central mass grows or the pattern speed of the bar decreases, the resonances sweep the disk outward and can affect a substantial fraction of it. Particles inside a vertical resonance region have access to the whole three-dimensional stochastic region, the envelope of which is delineated by the zero-velocity surface at a given energy as visible in Figures 11, 12, and 15. This heating is purely a dynamical effect and is not inhibited by the weak dissipation effects discussed here, essentially because the growth rate of the vertically unstable regions are larger than the dissipation rate.

The vertical disk heating mechanism might seem to counteract the radial accretion enhancement discussed before. Clouds lifted above the galactic plane, by accessing a new region of space, should see their effective collision rate strongly decreased, lowering their dissipation and their radial inflow. However, the gaseous part of clouds should still suffer enhanced shocks at each crossing of the galactic plane, because they all have to crowd near the plane periodically. This can contribute to limit the amplitude of their motion in z but not their radial inflow. On the contrary, stars formed in clouds, once lifted above the galactic plane, should keep their high amplitude of oscillation above the plane and be confined in their radial and vertical motions only by the Jacobi integral. Therefore, the disk heating in z should principally concern stars moving close to the plane, presumably young stars initially formed in the clouds. In a fraction of an orbital period, say of order 10⁸ yr, only stars of type A or later should remain observable. In recent self-consistent N-body simulations with gas component, Katz and Hernquist (1989) find that a bulge can grow spontaneously from stars formed in a gaseous disk.

The time scale for the bulge building process is relatively short, of the order of the growth rates of the vertical instability strips, typically 10 bar rotation periods. On the other hand, it is a known property of gas in bar models that it can be depleted also in about 10 rotation periods (e.g., Roberts, Huntley, and van Albada 1979). It is also a known property of many galaxies, including ours, to exhibit a hole of molecular gas inside about 5 kpc (Young 1987), corresponding to the size of a typical bar. So the overall aspect of a gas-rich (Sc, Sd galaxy) forming a bar will have changed substantially after a few billion years, leading to evolution along the Hubble sequence toward earlier types. This is consistent with the well-known correlation that the size of bulges is larger in early galaxy types than in late ones (e.g., Kormendy 1982). Another correlation is that the rotational velocity of early galaxy types is higher than in late ones (e.g., Rubin 1983). In the sequence Sc-Sa galaxy types, the typical rotation velocities pass from 100 to 300 km s^{-1} . This can be understood, in first order, by the virial theorem of purely self-gravitating systems,

$$E_{\rm tot} = E_{\rm grav} + E_{\rm kin} = -E_{\rm kin} , \qquad (18)$$

where E_{tot} is the total energy, E_{grav} is the gravitational energy, and E_{kin} is the kinetic energy. If the initial conditions of galaxies have a vanishingly small total energy (the mass of a protogalaxy is supposed to be extended and nearly at rest), the present kinetic energy amounts to the dissipated one, which is minus the present total energy. Clearly, the general effect of dissipation is to let disk galaxies evolve toward higher rotation velocities and central binding energies.

In the case where an inner bar is present, a time-dependent energy input is added to the vertical heating of the outer bar. A strange attractor can develop since the energy lost by dissipation can be balanced by the time-dependent potential of the inner bar. This shows also that the effect of repulsion of a bar persists in three dimensions. Provided that the density of clouds is not too high, forcing them to stay in the plane by frequent collisions near the galactic plane, a quasi-equilibrium cloud distribution might take a significant vertical extent.

406

391P

No. 2, 1990

[990ApJ...363..391P

This leads to the question why are not all galaxies with bulges barred? The answer might lie in the same dissipational property which causes or helps the formation of disks, then bars, central mass concentrations and bulges. It would be also responsible for the destruction of bars by a central mass concentration. Such a mechanism for the bar destruction has been studied recently by Norman and Hasan (1990) and Hasan and Norman (1990), who calculated how much mass has to be put at the center of a bar to destroy the main family of orbits supporting the bar. Their conclusion is that if the central region has accreted sufficient mass to produce an ILR radius comparable with the bar minor axis, then the elongated orbits supporting the bar do not exist anymore and the bar should dissolve, leaving a dynamically hot central part.

This scenario has a number of significantly different predictions for stellar populations and kinematics in the bulge component of disk galaxies. The metallicity and stellar population content of bulges reflect that of their associated disks at the epoch when the bar heating mechanism was in operation. The kinematics of bulges and spheroids are more closely associated with rotationally supported systems than for ellipticals (Kormendy 1982). We suspect that merging and interactions can also play a significant role in building bulges and spheroids. A relatively clean test of our models is to find barred Sc galaxies with metal-rich youngish populations in small rotationally supported bulges that have been formed by this process on a time scale of order a billion years or so. Such bulges could be dominated by A- or F-type stellar populations that have been formed at a high star formation rate in the disk and resonantly heated to form the bulge. While this scenario radically alters the conventional models of bulges forming first and consisting of old metal-deficient populations (Eggen, Lynden-Bell, and Sandage 1962), we have been unable to find any convincing evidence that rules out our model. On the contrary, with the discovery that the inner Galactic bulge is metal-rich ([Fe/H] $\approx 0 \pm 1$) (Freeman 1987; Frogel 1988; Rich 1989), that thick disks have a continuity of properties with bulges (Gilmore, Wyse, and Kuijken 1988), that stellar bars have the colors of their bulge (Kormendy 1982), and that the length of bars is roughly twice the size of their bulge (Athanassoula and Martinet 1980), this vertical heating mechanism appears quite natural. The bulges of M31 and our Galaxy seem to be metal-rich ([Fe/H] > -2) rotationally supported flattened systems (Mould 1986), consistent with disk heating at intermediate times (after globular cluster formation). In addition, the thick disks of M31 and our Galaxy have similar properties to their respective bulges, and in fact, thick disks and bulges have more in common with each other in terms of both populations and kinematical properties than they do with the more extended, metal deficient spheroidal component (Wyse and Gilmore 1988). That a general disk heating occurs in the central regions of our Galaxy is clear from the work of Lewis and Freeman (1989) who find a significantly increased velocity dispersion toward the Galactic center in their K-giant sample. However, for the more extended and more massive spheroidal component, such as the extended globular cluster component, bar heating would seem to be energetically insufficient.

VI. SUMMARY AND CONCLUSIONS

We have taken a highly simplified model of the gas dynamics of the interstellar medium and analyzed the basic effects of a

weak dissipation in barred galaxies with central mass concentrations, either a central point mass or a second small inner bar. The model assumes that in first order the behavior of the gas can be modeled by gas clouds moving on ballistic orbits with a dissipative drag force. The physics of this weakly dissipative, nearly Hamiltonian system has proved to be not only fascinating in its richness of behavior, such as exhibiting strange attractors etc., but has demonstrated the possibility of two major astrophysical consequences. The first is the rapid radial inflow of gas in broad resonant regions created by the dynamical interaction between a bar and a central mass concentration. This is important for the fueling of active galaxies and starbursts. The second effect is that substantial vertical heating of the disk can enhance the bulge component of disk galaxies leading to such components having a significant content of relatively young, metal-rich, and rotationally supported stellar populations.

In § II we have given a general discussion of weakly dissipative systems showing that dissipation can be amplified by resonances and their associated stochastic regions. In the limit of weak dissipation, the results have been found to be qualitatively independent of the dissipation law. The phenomenon of resonance jumping was discussed and illustrated using the generic model of a pendulum. A specific barred galaxy model is studied in § III, where it is shown that a central mass concentration in a barred galaxy creates horizontal and vertical resonances that are significantly broadened by the strong bar. As the large chaotic regions grow, this implies that the radial inflow of dissipative clouds is greatly enhanced. This is clearly evident in the numerical simulations (Fig. 7) and is particularly effective near lower order resonances such as the ILR. The strong radial inflow could form a new inner disk and inner bar. Many such possibilities can arise, and we have analyzed here the case of an inner bar and its interaction with an outer, larger bar. It is shown how clouds can be trapped in a ring, or in reality a torus, inside the large bar but outside the corotation point of the inner bar. A fascinating result here is that a strange attractor can form for slightly dissipative orbits of clouds in the time-dependent potential of the inner bar, which one cannot, in general, transform away in the rotating frame of the outer bar. In summary, radial gas flows in barred galaxies can generate a positive feedback mechanism where the growth of the central mass concentration enhances the inflow rate (due to the regions of broad resonances and enhanced stochasticity) further enhancing the central mass concentration. The end of this runaway process, which we associated with starbursts and AGNs, may occur due to the dissolution of the orbits sustaining the bar, leaving a kinematically hot inner spheroid, that will eventually become weakly triaxial.

In all cases, vertical resonances in bars are strong enough to heat perhaps the gas but certainly the stars to heights comparable with the size of bulges. The time scale for this process to develop once the bar is formed is rapid, of order 10 bar rotation periods, or of order one-tenth of a Hubble time. Clearly a mechanism for switching off or saturating this process must occur, and among the various possibilities suggested here we include (1) dissolution of the bar by the growth of a sufficiently large central mass concentration, or (2) the exhaustion of the gas supply due to star formation, inward and outward sweeping of the gas by the bar, or (3) mass expulsion and winds due to the energy input from supernovae or an active galactic nucleus. In summary, the disk heating mechanism to form the bulge can be visualized as stars diffusing through the regions of vertical resonances, loosing all previous constraints on their motion (except the Jacobi integral) which trapped them on islands of regular motion in phase space. Once free of the islands, which confined them to the disk, they can wander through the stochastic sea, exploring the available phase space, the spatial envelope of which has an approximately spheroidal shape. It is a pleasure to thank James Binney, Ken Freeman, Hashima Hasan, Lars Hernquist, Tim Heckman, Louis Martinet, Jim Pringle, and Rosemary Wyse for useful, enjoyable, and stimulating discussions on this subject. D. P. acknowledges the hospitality of the Space Telescope Science Institute. This work has been also partially supported by the Swiss National Science Foundation.

APPENDIX

RESONANCES IN TYPICAL DISK POTENTIALS

Some general constraints on the existence of radial and vertical resonances can be deduced from simple arguments. Let $\Phi(R, z)$ be an axisymmetric gravitational potential, and Ω the rotation frequency of circular orbits in the plane z = 0, κ the horizontal epicyclic frequency of circular orbits, and ν the vertical epicyclic frequency,

$$\Omega^{2} \equiv \frac{\Phi_{R}}{R}\Big|_{z=0}, \qquad \kappa^{2} \equiv \left(\Phi_{RR} + \frac{3}{R}\Phi_{R}\right)\Big|_{z=0}, \qquad \nu^{2} \equiv \Phi_{zz}\Big|_{z=0}.$$
(A1)

For stable circular orbits, Ω^2 , κ^2 , and ν^2 have to be positive. Since the density ρ has to be positive too, by Poisson's equation we have in the plane z = 0,

$$0 < \nabla^2 \Phi(R, 0) = \left(\Phi_{RR} + \frac{\Phi_R}{R} + \Phi_{zz} \right) \Big|_{z=0} = \kappa^2 + \nu^2 - 2\Omega^2 , \qquad (A2)$$

or

$$2\Omega^2 < \kappa^2 + \nu^2 . \tag{A3}$$

Considering the convexity of galactic potentials, let us assume that the axisymmetric potential can be approximated at z = 0 locally by a function of the elliptic distance d,

$$\Phi = \Phi(d)$$
, where $d^2 = R^2 + \frac{z^2}{c^2}$, (A4)

and c is the local z/R potential axis ratio. By differentiation one deduces,

$$\Omega^2(R) = c^2 \Phi_{zz}(R, 0), \quad \text{or} \quad \Omega = cv.$$
(A5)

The inequality (A3) becomes

$$\kappa^2 > (2c^2 - 1)v^2$$
 (A6)

In a spherical potential c = 1, so $\Omega = v$, and the property $\kappa > v$ holds. Now, an arbitrarily flat disk has an arbitrarily large v, i.e., $v > \kappa$. It follows that a potential which passes from a spherical shape at some radius to a sufficiently flat shape at another radius must have, by continuity, $\kappa = v$ at one intermediate radius at least. In some range of radii, κ and v are comparable, and their associated resonances too. From inequality (A6) one deduces also that the equality $\kappa = v$ can occur only in oblate potentials ($c \le 1$), and that κ^2 can not be negative in potentials less flattened than $c = 1/2^{1/2}$. Further insight can be gained by assuming that the axisymmetric potential $\Phi(d)$ is *locally* a power law of d,

$$\Phi(d) = C \frac{d^p}{p}, \quad \text{for } p \neq 0,$$

= $C \ln d$, for $p = 0$, (A7)

where C is a positive constant. For such potentials, we get,

$$\Omega^{2} = CR^{p-2}, \qquad \kappa^{2} = C(p+2)R^{p-2}, \qquad v^{2} = \frac{C}{c^{2}}R^{p-2}, \qquad \nabla^{2}\Phi \bigg|_{z=0} = C(p+c^{-2})R^{p-2}.$$
(A8)

The positivity of κ^2 implies p > -2, and the positivity of mass implies $p > -c^{-2}$. If one assumes that the radial density gradient ρ_R is negative, we get, p < 2. If furthermore one assumes that the density at constant R is maximum at z = 0, then $(\nabla^2 \Phi)_{zz} < 0$, which gives p < 2, and $p > 2 - 3c^{-2}$. In short, reasonable assumptions constrain the power p and c by

$$2 > p > \max(-2, -c^{-2}, 2 - 3c^{-2}).$$
 (A9)

Figure 16 shows the regions in the plane (p, c) which are excluded by the positivity of ρ , and κ^2 , and the negativity of the derivatives ρ_R and ρ_{zz} .

© American Astronomical Society • Provided by the NASA Astrophysics Data System

391P

Jaboan

No. 2, 1990

1990ApJ...363..391P

For a nonaxisymmetric perturbation of corotation R_{CR} and pattern speed Ω_b , normalizing C by $C \equiv \Omega_b^2/R_{CR}^{p-2}$, the conditions of major resonances for *direct* circular orbits are

$$m = \frac{\kappa}{\Omega - \Omega_b} = \frac{\sqrt{p+2}}{1 - (R/R_{CR})^{1-p/2}},$$

$$n = \frac{\nu}{\Omega - \Omega_b} = \frac{1/c}{1 - (R/R_{CR})^{1-p/2}},$$

$$m, n = 0, \pm 1, \pm 2, \dots.$$
(A10)

The denominators behave both as $f(x) = 1 - x^{1-p/2}$, with $x = R/R_{CR}$. This function is plotted in Figure 17 for $-2 \le p \le 2$ (solid). It is monotonically decreasing for each p, except for the degenerate case for the potential of an harmonic oscillator (p = 2), so m and n are increasing with R. This property implies that typically between the center and corotation m and n increase from their minimum value at the center, $m_0 = (p + 2)^{1/2}$ and $n_0 = c^{-1}$ respectively, shown in Figure 16, to $+\infty$ at corotation, and from corotation outward, m and n increase from $-\infty$ to 0.

The conditions of major resonances for retrograde circular orbits are

$$m = \frac{\kappa}{\Omega + \Omega_b} = \frac{\sqrt{p+2}}{1 + (R/R_{CR})^{1-p/2}},$$

$$n = \frac{\nu}{\Omega + \Omega_b} = \frac{1/c}{1 + (R/R_{CR})^{1-p/2}},$$

$$m, n = 0, \pm 1, \pm 2, \dots.$$
(A11)

The denominators behave now both as $g(x) = 1 + x^{1-p/2}$, i.e., they are monotonically increasing functions. This function is plotted in Figure 17 for $-2 \le p \le 2$ (dash). A similar analysis as above shows that only resonances m, n of lower order than $m_0 = (p + 2)^{1/2}$ and $n_0 = c^{-1}$, respectively, occur at some radii. If p < 2, then m < 2, and only the m = 1 resonance occurs. In many cases, c > 0.5, then similarly n < 2 and only the n = 1 resonance occurs. This later resonance corresponds to Binney's instability strip (Binney 1981).



FIG. 16.—Regions in the plane (p, c) of possible gravitational spheroidal potentials which can be approximated locally by a power p of the ellipsoidal radius of axis ratio c. The shaded regions are excluded by the indicated constraints. The curve $\kappa = v$ is also indicated. The dashed lines $m_0 = 1$, 2 and $n_0 = 1$, 2, ... are the resonant values extrapolated at R = 0, i.e., they correspond to $m_0 = (p + 2)^{1/2}$ and $n_0 = c^{-1}$. The potentials of harmonic oscillators, along the line p = 2, are especially degenerate, because the resonance lines $n_0 > 0$ exist only on the left side. Kepler's potential is located at p = -1, c = 1, at the intersection of the curves $\rho = 0$, $\kappa = v$, $\rho_{zz} = 0$, $m_0 = 1$, and $n_0 = 1$. Typical disk galaxy potentials are often located in the region $p \approx 0$, c < 1 which have flat rotation curves. The flatter the disks are ($c \ll 1$), the more crowded are the vertical resonances.

FIG. 17.—Functions $f(x) = 1 - x^{1-p/2}$ (solid) and $g(x) = 1 + x^{1-p/2}$ (dash) for different values of p, p = -2, -1.5, ... 2. The case p = 2, the potential of a harmonic oscillator, is clearly degenerate.

PFENNIGER AND NORMAN

..363..391P

1990ApJ.

- Adams, T. F. 1977, Ap. J. Suppl., 33, 19. Arnold, V. 1974, Equations différentielles ordinaires (Moscow: Mir). Athanassoula, E. 1989, in Dynamics of Astrophysical Discs, ed. J. A. Sellwood
- (Cambridge: Cambridge University Press), p. 145. Athanassoula, E., and Martinet, L. 1980, Astr. Ap., 87, L10.
- Athanassoula, L., Bienaymé, O., Martinet, L., and Pfenniger, D. 1983, Astr. Ap., 127, 349.
- Binney, J. 1981, M.N.R.A.S., 196, 455.
 Binney, J., and Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton University Press) (BT).
- Carlberg, R. G., and Freedman, W. L. 1985, A.J., 298, 486.
- Chen, C., Györgyi, G., and Schmidt, G. 1986, Phys. Rev. A, 34, 2568.
- . 1987, Phys. Rev. A, 35, 2660.
- Combes, F., and Gerin, M. 1985, Astr. Ap., 150, 327.
- Combes, F., Debbasch, F., Friedli, D., and Pfenniger, D. 1990, Astr. Ap., 233, 82. Contopoulos, G. 1983, Ap. J., 275, 511. Contopoulos, G., and Papayannopoulos, T. 1980, Astr. Ap., 92, 33. de Vaucouleurs, G. 1963, Ap. J. Suppl., 8, 31. Eggen, O. J., Lynden-Bell, D., and Sandage, A. 1962, Ap. J., 136, 748. Ferrers, N. M. 1877, Quart. J. Pure Appl. Math., 14, 1. Freeman, K. C. 1987, Ann. Rev. Astr. Ap., 25, 603. Frogel, J. A. 1988, Ann. Rev. Astr. Ap., 26, 51. Gerhard, O., Vietri, M., and Kent, S. M. 1989, Ap. J. (Letters), 345, L33. Gilmore, G., Wyse, R. F. G., and Kuijken, K. 1989, Ann. Rev. Astr. Ap., 27, 555. Goldstein, H. 1971, Classical Mechanics (Reading, Mass.: Addison-Wesley). Hasan, H., and Norman, C. A. 1990, Ap. J., 361, 69. Hernquist, L. 1989, Nature, 340, 687. Jarvis, B. J., Dubath, P., Martinet, L., and Bacon, R. 1988, Astr. Ap. Suppl., 74, 82

- Jarvis, B. J., Dubath, P., Martinet, L., and Bacon, R. 1988, Astr. Ap. Suppl., 74, 513.
- Katz, N., and Hernquist, L. 1989, in *The Epoch of Galaxy Formation*, ed. C. S. Frenk, R. S. Ellis, T. Shanks, A. F. Heavens, and J. A. Peacock (Dordrecht: Kluwer), p. 433.
- Kormendy, J. 1982, in Morphology and Dynamics of Galaxies, Twelfth Saas-Fee Lecture, ed. L. Martinet and M. Mayor (Geneva: Geneva Observatory), p. 113
- p. 115.
 Lewis, J. R., and Freeman, K. C. 1989, A. J., 97, 139.
 Lichtenberg, A. J., and Lieberman, M. A. 1983, *Regular and Stochastic Motion* (New York: Springer) (LL).
 Lynden-Bell, D., and Kalnajs, A. J. 1972, M.N.R.A.S., 157, 1.
 McKee, C. F., and Ostriker, J. P. 1977, Ap. J., 218, 148.
 Miyamoto, M., and Nagai, R. 1975, *Pub. Astr. Soc. Japan*, 27, 533.
 Mould, J. 1986, in *Stellar Populations*, ed. C. A. Norman, A. Renzini, and M. Toci (Combridge University Press), p. 0.

- Tosi (Cambridge: Cambridge University Press), p. 9.

- REFERENCES
 - Norman, C. A. 1988, in Lecture Notes in Physics, Vol. 297, Comets to Cosmology, ed. A. Lawrence (Berlin: Springer), p. 177. Norman, C. A., and Ikeuchi, S. 1989, Ap. J., 372, 383.

 - Norman, C. A., and Hasan, H. 1990, in Lecture Notes in Physics, Heidelberg Conference on Dynamics and Interactions of Galaxies, ed. R. Wielen (Berlin:

 - ed. D. Merritt (Cambridge: Cambridge University Press), p. 81. Pfenniger, D., and Friedli, D. 1991, Astr. Ap., submitted.

 - Phinney, S. 1990, Heidelberg Conference on Dynamics and Interactions of
 - Galaxies, unpublished talk.
 Prendergast, K. H. 1983, in *IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Dordrecht: Reidel), p. 215.
 Rich, R. M. 1989, *A.J.*, 95, 828.

 - Roberts, W. W., Huntley, J. M., and van Albada, G. D. 1979, Ap. J., 233, 67. Rubin, V. C. 1983, in IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), p. 3.

 - b) Galdares, ed. E. Athanassona (Dordrecht, Reicht, p. 2.)
 Sanders, R. H., and Tubbs, A. D. 1980, Ap. J., 235, 803.
 Savonije, G. J., and Papaloizou, J. C. B. 1984, M.N.R.A.S., 207, 685.
 Scalo, J. M. 1987, in Interstellar Processes, ed. D. J. Hollenbach and H. A. Thronson (Dordrecht: Reidel), p. 349.
 - Schmidt, G. 1987, in *Chaotic Phenomena in Astrophysics*, ed. J. R. Buchler and H. Eichhorn (New York: New York Academy of Sciences), p. 97. Schwarz, M. P. 1981, *Ap. J.*, **247**, 77.

 - Scoville, N. Z., Matthews, K., Carico, D. P., and Sanders, D. B. 1988, Ap. J. (Letters), 327, L61.
 - Shlosman, I., Frank, J., and Begelman, M. C. 1989, *Nature*, **338**, 45. Simkin, S., Su, H., and Schwarz, M. P. 1980, *Ap. J.*, **237**, 404. Sparke, L., and Sellwood, J. A. 1987, *M.N.R.A.S.*, **225**, 653.

 - Spitzer, L. 1978, Physical Processes in the Interstellar Medium (New York: Wiley)

 - Wiley). Teuben, P. J., and Sanders, R. H. 1985, M.N.R.A.S., **212**, 257. van Albada, G. D., and Roberts, W. W. 1981, Ap. J., **246**, 740. van Albada, T. S., and Sanders, R. H. 1982, M.N.R.A.S., **201**, 303. Vogel, S. N., Kulkarni, S. R., and Scoville, N. Z. 1988, Nature, **334**, 402. Wisdom, J. 1987, *Icarus*, **72**, 241. Wyse, R. F. G., and Gilmore, G. 1988, A.J., **95**, 1404. Young, J. S. 1987, in *IAU Symposium 115*, Star Forming Regions, ed. M. Pairnbert and L. Lurghy, Worderschit, Paidel, p. 557. Peimbert and J. Jugaku (Dordrecht: Reidel), p. 557.

COLIN NORMAN: Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218

DANIEL PFENNIGER: Geneva Observatory, CH-1290 Sauverny, Switzerland