THE ASTROPHYSICAL JOURNAL, 518:613–626, 1999 June 20

SECULAR EVOLUTION OF SPIRAL GALAXIES. III. THE HUBBLE SEQUENCE AS A TEMPORAL EVOLUTION SEQUENCE

XIAOLEI ZHANG
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Mail Stop 78, Cambridge, MA 02138
Received 1997 October 22; accepted 1999 January 22

ABSTRACT

In the first two papers in this series, we have presented a dynamical mechanism by which an unstable spiral mode in a galactic disk can remain quasi-stationary on the order of a Hubble time at the expense of a continuous dissipative basic state evolution, manifesting as the secular redistribution of disk matter, as well as the secular heating of disk stars. In the current paper, we consider in detail the astrophysical consequences of this spiral-induced basic state evolution process and show that it could explain, among a spectrum of observational facts, the variation of the disk surface-density profiles across the Hubble types, as well as the well-known age–velocity dispersion relation of the solar-neighborhood stars. We seek to establish that the Hubble sequence for the classification of galaxies, when viewed in the reverse direction, is essentially a temporal evolution sequence. The spiral-induced secular mass redistribution process could lead to the evolution of the Hubble type of a spiral galaxy from late (Sd and Sc) to early (Sb and Sa). Furthermore, since similar radial mass-accretion processes could also be mediated by central stellar bars, or other central nonaxisymmetric structures, this leads to the continued galaxy evolution across Hubble type S0 and into disky elliptical galaxies. The secular evolution mechanism presented in this series of papers, together with other previously proposed evolution mechanisms, provides a framework for understanding the results of *Hubble Space Telescope* Deep Field and Medium Deep Survey, which showed that there is significant evolution of the number statistics of the Hubble types of galaxies with redshift, consistent with an evolution along the Hubble sequence from late to early. It also provides possible solutions to several long-standing problems in the theory and observation of galaxy formation and evolution, among them the cause of the steepening of the galaxy morphology-environment relation; the similarity between the Tully-Fisher and Faber-Jackson relations for spiral and elliptical galaxies, respectively; the existence of a fundamental-plane relation for elliptical galaxies; and the mass–angular momentum anticorrelation observed along the Hubble sequence.

Subject headings: galaxies: evolution — galaxies: fundamental parameters — galaxies: structure

1. INTRODUCTION

In presenting the now well-known classification scheme of galaxies, Hubble (1936) emphasized that the tuning-fork diagram he obtained depicted the systematic variation of the morphological characteristics and that the terms "early type" and "late type" referred only to a galaxy's relative position in this empirical sequence; they did not imply temporal evolution connections. Even so, Hubble could not refrain from mentioning the similarity of his sequence to the one given by the evolution theory of Jeans (1928), which suggested the collapse of spherical nebulae as the mechanism for the formation of disk galaxies and the tidal forcing at two opposite ends of the flattened nebulae as the generating mechanism of spiral structure. It has been said that part of what had originally motivated Hubble to devise a classification scheme was the hope of finding temporal evolution connections among galaxies.

During the past decades, improved observational results and theoretical understandings of the structure and dynamics of galaxies have essentially ruled out the possibility that galaxies evolve along the Hubble sequence from early to late types. A serious consideration of the morphological evolution of spiral galaxies along other possible routes, however, was hampered especially by the well-known result in linear density wave theory, which indicates no secular evolution of the basic state of a galactic disk.

Nevertheless, observational results that had been accumulating during the past few decades pointed with increasing clarity to the likelihood of galaxy evolution along the Hubble sequence in the reverse direction. (1) Early-type spiral galaxies are generally larger in size and more massive than the late-type galaxies (Freeman 1970). Furthermore, from the late to early Hubble types, the total gas content of a galaxy decreases, whereas the molecular to total gas fraction increases (Young 1990). (2) As already noted at Hubble's time, the size of the galactic bulge increases from the late to the early spiral types. Kinematic studies (see, e.g., Gilmore, King, & van der Kruit 1990; Dejonghe & Habing 1993) have also shown that bulges are largely rotationally supported and show a certain degree of flattening (Mihalas & Binney 1981, p. 324), thus appearing to be a continuation from the kinematic properties of the disks. The increase in bulge size, as well as the stratification of the kinematic and chemical properties of bulge stars (Dejonghe & Habing 1993), is most easily explained by a gradual mass-accretion process. (3) As has recently been found in the *Hubble Space Telescope* Medium Deep Survey (MDS) and Hubble Deep Field (HDF) (see, e.g., Couch et al. 1994; Couch et al. 1998; Ellis 1997a), there exists a far larger number of blue late-type galaxies at intermediate to high redshifts than at the present. In fact, there appears to be a smooth and systematic variation of the galaxy number counts from the later to the earlier Hubble types as redshift decreases. An evolutionary scenario is therefore needed to account for this observed change in the statistical distribution of Hubble types with redshifts.

On the theoretical front, we know that, in the widely accepted hierarchical clustering and dissipation scenario for

galaxy formation, any model that forms ellipticals as ellipticals and spirals as spirals can not explain the fact that the early-type objects, which are systematically more massive, all have markedly less specific angular momentum than the late-type objects. This fact is especially difficult to explain away considering the well-known demonstration of Peebles (1969) and Faber (1981) that there is no correlation between the magnitude of the initial density fluctuation and the specific angular momentum of such a clump. A merger scenario for forming the early-type objects, however, has difficulties in producing the smooth spectrum of all Hubble types, including the detailed kinematics of most of the rotationally supported disky elliptical galaxies, which are the most abundant kind among all elliptical galaxies (Bettoni et al. 1997).

Therefore, observational results and theoretical considerations both demand a trend of galaxy evolution from late to early Hubble types, whereby the mass distribution of a galaxy becomes more and more centrally concentrated as a result of the continuous outward angular momentum transport by the galactic spiral structure. Past proposed secular evolution mechanisms (see, e.g., the review articles of Kormendy 1982; Norman 1984; Martinet 1995; Pfenniger 1996) cannot account for all of these secular evolution phenomena. In what follows, we demonstrate that the evolution mechanism presented in the current series of papers could serve as the basis for the possible explanation of many observed secular evolution phenomena.

2. SPIRAL-INDUCED SECULAR EVOLUTION EFFECTS: HEATING OF THE DISK STARS AND REDISTRIBUTION OF DISK MATTER

As we have already mentioned in the previous two papers in this series, the presence of spiral structure serves to accelerate the entropy evolution of the disk galaxy toward configurations of ever-increasing central concentration. This evolution trend coincides with that of an evolution along the Hubble sequence, whereby a gradually greater fraction of a galaxy's mass is concentrated onto the central bulge region. In order to achieve this evolution effect, two basic elements must be present in the relevant evolution mechanism: an element that could lead to the secular redistribution of the disk matter and an element for the heating of the disk stars so they will gradually rise above the galactic plane en route to the galactic central region. We show in the following that these two elements are the natural ingredients of the spiral-induced wave and basic state angular momentum exchange process.

2.1. *Secular Evolution of Disk Surface Density Distribution*

The earlier results of Zhang (1996a, hereafter Paper I; 1998, hereafter Paper II) show that for an open spiral pattern the density and potential spirals generally do not have overlapping spatial distributions but rather are phase shifted from each other. For a spontaneous spiral mode, the sense of this phase shift is such that the potential spiral lags the density spiral inside corotation, and vice versa outside corotation. At the quasi steady state of the wave mode, this phase shift and its associated torque action leads to a secular exchange of energy and angular momentum between the wave and the basic state. The consequence of this exchange process on the wave mode is such that it offsets the spontaneous growth tendency of the mode so that it remains in a quasi steady state despite constantly transporting angular momentum outward (Paper II). The consequence on the basic state stars, however, is such that the stars both inside and outside the corotation radius migrate away from the corotation region, so that with time a more and more centrally concentrated surface density distribution is achieved, together with the buildup of an extended outer envelope.

In Paper II, we have already obtained the rate of spiral-induced orbital change for an average star, which, for a galaxy with a nearly flat rotation curve and for a quasi-stationary spiral wave, is given by

$$\frac{dr_*}{dt} = -\frac{1}{2} F^2 v_c \tan i \sin (m\phi_0) , \tag{1}$$

where i is the pitch angle of the spiral, m the number of arms of the spiral pattern, ϕ_0 the potential and density phase shift, F the fractional amplitude of the spiral, and v_c the circular speed of the galaxy at the relevant radius. This equation has been quantitatively verified by the results of N-body simulations (Paper II).

For a typical two-armed spiral of 20% fractional amplitude and 20° pitch angle, the orbital decay rate is about 0.2 km s^{-1}, or 2 kpc of radial migration in 10^{10} yr. Thus the spiral-induced stellar orbital decay inside corotation should have a significant effect on the evolution of the disk morphology over the time span of a Hubble time. In particular, fast evolution is expected for galaxies that contain large-amplitude and large–pitch-angle spiral patterns.[1]

2.1.1. *A Viscous Accretion Disk Analogy*

We begin by deriving a viscous accretion disk analogy for the spiral disks. This allows us to put our discussion in the general context of the viscous accretion disk theory and to facilitate the comparison with other types of astrophysical accretion disks.

In the general viscous accretion disk theory (Pringle 1981), let $C(r, t)$ be the total z torque applied by the inner disk on the outer disk at radius r and time t (C here is equivalent to the G used by Pringle except for a minus sign). The torque is related to the viscosity parameters through

$$\frac{\partial C}{\partial r} = \frac{-\partial}{\partial r} (2\pi r \nu \Sigma A r) , \tag{2}$$

where $A = r\, d\Omega/dr$ is the shear rate, $\nu \approx l v_T$ is the kinematic viscosity, l is the mean free path, and v_T is the thermal velocity of the disk particles.

To calculate the effective viscosity due to the presence of a quasi-stationary spiral structure, we note that the torque gradient $\partial C/\partial r$ is related to our torque integral $T_1(r) \equiv r \int_0^{2\pi} \Sigma_1 (\partial V_1/\partial \phi)\, d\phi$ and phase shift ϕ_0 through (Appendix A of this paper, as well as the appendices of Paper II)

$$\begin{aligned} -\frac{\partial C}{\partial r} &= 2\pi r \frac{\overline{dL}}{dt} \\ &= -T_1(r) = -r \int_0^{2\pi} \Sigma_1 \frac{\partial V_1}{\partial \phi}\, d\phi \\ &= -\pi F^2 v_c^2 \sin (m\phi_0) \tan i \Sigma_0 r , \end{aligned} \tag{3}$$

[1] The result of the numerical evolution of the Poisson equation shows that the potential/density phase shift generally increases with the pitch angle of the spiral until the pitch angle is around 45°. The phase shift then decreases with further increase of the pitch angle of the spiral until the pitch angle reaches 90°.

where C is the sum of gravitational and advective torque couplings and L is the angular momentum density of the disk matter.

If we compare the two expressions (eqs. [2] and [3]) for $\partial C/\partial r$, we arrive at

$$\frac{\partial}{\partial r}(2\pi v_{\rm eff}\Sigma r^3\Omega') = \pi F^2 v_c^2 \tan i \sin(m\phi_0)\Sigma r , \quad (4)$$

where $\Omega = v_c/r$ and $\Omega' = d\Omega/dr$.

We now try to obtain the expression of $v_{\rm eff}$ for a particular type of spiral disk. If the basic state of the disk has a flat rotation curve and a $1/r$ surface density distribution, we easily obtain from the above that

$$v_{\rm eff} = \tfrac{1}{2} F^2 v_c \tan i \sin(m\phi_0) r . \quad (5)$$

It is obvious from the above expression of $v_{\rm eff}$ that it changes sign at $r_{\rm co}$, where ϕ_0 itself changes sign (Paper I). Therefore the effective viscosity due to a quasi-stationary spiral mode is positive inside corotation and negative outside corotation, which is consistent with the fact that the disk matter accretes inside corotation and excretes outside corotation.

For a 20% amplitude and 20° pitch angle spiral, the effective α parameter (Pringle 1981) produced by the spiral-induced radial mass-accretion process is about 0.01 (Zhang 1996b). Therefore the effective viscosity produced by nonaxisymmetric instabilities is likely to be an important candidate for the long-sought "anomalous viscosity" in many types of astrophysical accretion disks.

2.1.2. *Evolution of the Disk Surface Density Profile*

The surface brightness distribution of the galactic disks is often found to be well fitted by an exponential function of the form $I(r) = I_D e^{-r/r_D}$ (de Vaucouleurs 1959; Freeman 1970), where I_D and r_D are constants representing the extrapolated central brightness and the scale length of the exponential function, respectively. For a constant mass-to-light ratio, this leads to an exponential distribution for the surface density of the disk as well. In the recent years, it becomes clear that most galactic disks have surface density distributions that are better described by a $1/r$ function, which approximates very well the exponential function over a wide range of intermediate galactic radii (Seiden, Schulman, & Elmegreen 1984). We now analyze the consequence of the spiral-induced radial mass-accretion process to the evolution of the surface density profile of the galactic disks.

In general, the evolution of the disk surface density can be calculated by solving the two conservation equations for viscous accretion disks, i.e., the angular momentum conservation equation

$$-\frac{\partial C}{\partial r} = 2\pi r \frac{\partial}{\partial t}(\Sigma r^2\Omega) + 2\pi \frac{\partial}{\partial r}(r\Sigma v_r r^2\Omega) \quad (6)$$

and the mass conservation equation

$$r\frac{\partial \Sigma}{\partial t} + \frac{\partial}{\partial r}(r\Sigma v_r) = 0 \quad (7)$$

for $v_r(r, t)$ and $\Sigma(r, t)$, if we know the distribution of $\partial C(r, t)/\partial r$.

In the case of accretion induced by a galactic spiral structure, if the wave remains quasi-stationary (which means that almost all of the angular momentum deposited by the wave goes to the basic state), we further have that

$$\frac{\partial C}{\partial r} = \pi F^2 v_c^2 \sin(m\phi_0) \tan i\Sigma r , \quad (8)$$

where the wave amplitude F, pitch angle i, and phase shift ϕ_0 are themselves functions of both r and $\Sigma(r, t)$, as well as the distribution of the stellar velocity dispersion. Therefore, it is clear that we have a fully nested nonlinear and global problem, the solution of which could be obtained by an iterative approach.

However, an added level of complication to the above picture is introduced by the fact that the orbital decay rate $dr/dt = v_r$ resulted from the quasi steady state constraint for the wave mode, which leads to equation (1) in the case of a galaxy with a flat rotation curve (Paper II), does not always agree with that derived from the set of equations (6), (7), and (8), which governs the smooth change of basic state properties because of the presence of the wave mode. In cases in which such consistency is lacking, we would not have a true quasi steady state for the wave mode; i.e., the wave mode will be coevolving with the basic state characteristics during the accretion process.

It is easily verified that at least in one case we can obtain a consistent solution that satisfies all four equations. This is the case of a galaxy with a flat rotation curve and a $1/r$ surface density distribution, and with the spiral mode leading to a constant orbital migration rate v_r (which in reality can hold only for limited range in the inner galaxy, since v_r ought to change sign at corotation; for further discussion, see Appendix B). In this case the accretion flux is continuous throughout the galactic radii of concern, and the only place where the surface density is increasing is the central region of the galaxy.

In physical galaxies, we expect that the wave amplitude is nearly steady after a long period of evolution. We also observe that most galaxies do have flat rotation curves and $1/r$ surface density profiles over the intermediate radial range. This indicates that the requirement of global self-consistency during the mass accretion process due to a quasi-stationary spiral structure may be responsible for the creation and the maintenance of the approximate $1/r$ surface density profile of the disk under a flat-rotation-curve galactic potential.

2.2. *Secular Heating of the Disk Stars*

The secular radial migration of mean stellar orbit is not the only effect produced by the wave/basic-state energy and angular momentum exchange. There is also an associated heating effect, because of the fact that the circular speed Ω of stars at a given galactic radius is in general different from the pattern speed Ω_p. Since the basic state stars' energy and angular momentum are related through a factor Ω, and those for the wave are related through a factor Ω_p, the complete angular momentum exchange between the wave and the basic state implies that there will be certain amount of the energy left, which is used to heat the disk stars. In Paper II, we have derived, as well as quantitatively verified through N-body simulations, that the rate of increase of the mean-square random velocity of the basic state stars due to their interaction with a quasi-stationary spiral structure is

given by

$$\frac{d\sigma^2}{dt}(r, t) \equiv D^{(3d)}(r, t) = (\Omega - \Omega_p)F^2 v_c^2 \tan i \sin(m\phi_0) , \tag{9}$$

where σ and D^{3d} are the dispersion and the effective diffusion constant of the random velocity of basic state stars and the rest of the symbols have been explained before. In the following, we show that the above results could explain a number of observational facts pertaining to the magnitude and distribution of the random velocity of stars in the galactic disks.

2.2.1. *The Age–Velocity Dispersion Relation of the Solar Neighborhood Stars*

The number density f_0 of stars at the solar neighborhood was found to follow the Schwarzschild distribution in peculiar velocity $\boldsymbol{v}$:

$$f_0(\boldsymbol{v})\, d^3\boldsymbol{v} = \frac{n_0\, d^3\boldsymbol{v}}{(2\pi)^{3/2}\sigma_r\,\sigma_\phi\,\sigma_z} \exp\left[-\left(\frac{v_R^2}{2\sigma_r^2} + \frac{v_\phi^2}{2\sigma_\phi^2} + \frac{v_z^2}{2\sigma_z^2}\right)\right] \tag{10}$$

(Schwarzschild 1907), where n_0, σ_r, σ_ϕ, and σ_z are constants. Furthermore, it has been found that the velocity dispersion components σ_r, σ_ϕ, and σ_z of the solar-neighborhood stars increase smoothly with the mean age of the stellar group (Gliese 1969), following roughly a $t^{1/2}$ law (Wielen 1977). Starting with the early attempt of Spitzer & Schwarzschild (1951, 1953), a spectrum of mechanisms has been proposed to account for the secular heating of disk stars (Barbanis & Woltjer 1967; Carlberg & Sellwood 1985; Wielen & Fuchs 1990). However, so far none of the proposed mechanisms have been able to satisfy the requirement of being physically realistic on the one hand and being able to account for both the form of the age–velocity dispersion law and the shape of the velocity ellipsoid on the other (Binney & Lacey 1988; Gilmore et al. 1990). As Ivan King pointed out in his Saas Fee lectures, "It is clear that we do not yet fully understand this intriguing problem" (Gilmore et al. 1990, p. 175).

In the past few decades, it has been widely believed that a quasi-stationary spiral structure could not serve to heat the disk stars. However, in Paper I and Paper II, we presented the dynamical mechanism by which a quasi-stationary spiral density wave could interact and exchange energy and angular momentum with the basic state of the galactic disk. We now calculate the form of age–velocity dispersion relation due to the spiral-heating mechanism, as well as the magnitude of this heating effect for the solar neighborhood stars.

Since heating by the spiral structure is due to the action of a local gravitational instability or a collisionless shock (Paper I), the energy injection into the stars is sudden and local. We can thus employ the impulse approximation. This implies that the successive increases of stellar random velocity at each crossing of the spiral arms are independent events and are also independent of the magnitude of the peculiar velocity a given star already has. Furthermore, the velocity increment is also independent of the mass of the star. Therefore, after n crossings of the spiral arms, the total increase of stellar random velocity for an average star is

$$\Delta\sigma \sim \sqrt{n(\delta v)^2} , \tag{11}$$

where δv^2 is the increase in mean-square stellar random velocity per crossing of the arm. Since, over a long time span, $n \propto t$, the age of the star, we see that heating by a quasi-stationary spiral structure can produce a $t^{1/2}$ law for the age–velocity dispersion relation. In general, the stellar velocity dispersion at the current epoch depends also on the average random velocity of stars acquired at birth, so we can write the age–velocity dispersion law in general as

$$\sigma(t) = \sqrt{\sigma_0^2 + D^{(3d)}t} , \tag{12}$$

where σ_0 is the initial velocity dispersion of stars at zero age, and with $D^{(3d)}$ given by equation (9).

A couple of points need to be noted in calculating the numerical values of the coefficients in the age–velocity dispersion relation due to the spiral structure. First of all, the correct values to use in equation (9) for deriving the age–velocity dispersion relation over the life span of the Milky Way are not those for the current epoch, since the secular evolution effect induced by the Galactic spiral structure changes both the properties of the basic state as well as the properties of the wave mode. A further complication is introduced by the secular radial migration of stars during the same evolution process. Therefore, the overall effect is that the exact form of the age–velocity dispersion relation for any given spatial location is expected to be modified from that of equation (12).

However, the result of N-body simulations shows that the actual age–velocity dispersion curve is still quite close to that given by equation (12) despite the above-mentioned complications. In Figure 1, we plot the evolution of the mean-square peculiar velocity of one of the N-body spiral modes obtained in § 5 of Paper II, which is a run with a large number of particles ($N = 300{,}000$), so the contribution of Poisson noise to heating is small and the acceleration of particles witnessed there can be entirely attributed to spiral heating (Paper II). On the same figure, we also plot a line that corresponds to $t^{1/2}$ evolution with an effective diffusion coefficient $D^{1d} = 0.8 \times 10^{-6}$, calculated through equation (9) using parameters appropriated for the spiral mode in that simulation. A simple explanation for the observed approximate $t^{1/2}$ evolution behavior of the N-body heating is that, first of all, the factor $F^2 v_c^2 \tan i \sin(m\phi_0)$ in equa-

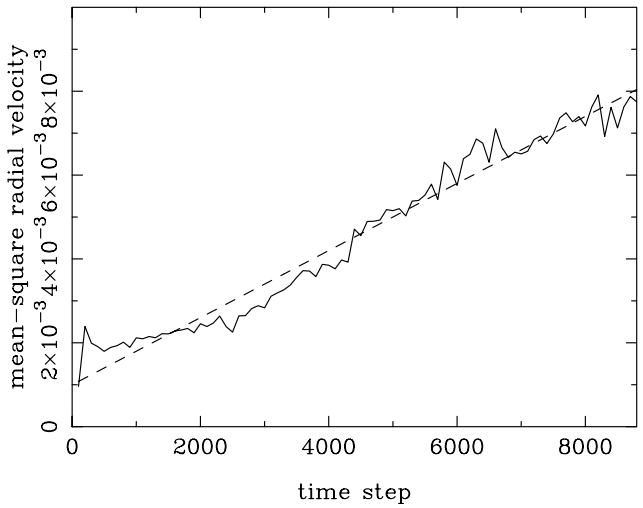

FIG. 1.—Comparison of the N-body evolution of the mean-square velocity of a group of stars inside corotation with the expected linear evolution function.

tion (9) can be shown to be approximately independent of galactic radius for the inner disk region (Appendix B). Furthermore, the factor $\Omega - \Omega_p$ is also approximately constant for the short radial range that this particular group of particles in Figure 1 traverses between time steps 0 and 8000. However, over the time span of a Hubble time, the spiral mode itself would change its properties as the basic state evolves, so in reality we could still expect some small departure from the $t^{1/2}$ law.

Having established the approximate validity of the age–velocity dispersion law (eq. [12]), we now estimate an average value of the apparent diffusion coefficient D^{3d} for the solar-neighborhood stars, appropriate as a mean value for the past 10^{10} yr. For a 20° pitch-angle spiral with $1/r$ density modulation (since we expect that in the past the Galactic spiral pattern was more open than it is now), which results in an averaged phase shift of 3° in the inner disk and an average $\Omega - \Omega_p = 8$ km s^{-1} kpc^{-1}, which is two-thirds of its current value at the solar neighborhood; using for the fractional amplitude of the wave $F = 0.15$ as an average over the past 10^{10} yr, we obtain that the mean velocity diffusion coefficient for the current solar neighborhood stars over the past 10^{10} yr is $\bar{D}^{(3d)} = 5.7 \times 10^{-7}$ (km s^{-1})^2 yr^{-1}, which is quite close to the expected value of $\bar{D}^{(3d)} = 6 \times 10^{-7}$ (km s^{-1})^2 yr^{-1} (Wielen 1977). In Figure 2, we plot the age–velocity dispersion curve with this mean diffusion coefficient, together with some of the published data for the actually stellar velocity dispersion and age group correlation. The agreement between the $t^{1/2}$ curve and the observed data, as was already found by Wielen (1977), was excellent. Note that we have omitted plotting the data beyond the age of the Galaxy, which if done would show the well-known Parenago's discontinuity, as was also confirmed in the recent analysis of *Hipparcos* data by Dehnen & Binney (1998).

2.2.2. *The Cause of Isotropic Velocity Diffusion in Three Dimensions and the Preservation of the Gaussian Velocity Distribution through Time*

Two further aspects of the distribution of stellar random velocities in the solar neighborhood that require explanation are (1) the observed axis ratios of the velocity ellipsoid and (2) the preservation of the Gaussian (Schwarzschild) velocity distribution in time.

From observations of the representative K + M dwarfs in Gliese's catalog, the velocity dispersions in the r, ϕ and z directions are found to be $\sigma_r = 48$ km s^{-1}, $\sigma_\phi = 29$ km s^{-1}, and $\sigma_z = 25$ km s^{-1}, respectively. Wielen (1977) demonstrated that the apparently different rates of the velocity diffusion in the three spatial dimensions is in fact an illusion. The ratios of the apparent diffusion constants C_r, C_ϕ, and C_z can be obtained from the epicycle motion of the stars in the galactic plane, and from the stellar vertical oscillations, using an isotropic one-dimensional diffusion coefficient $D^{(1d)}$.

Both of these aspects can be naturally explained by the spiral heating mechanism. Since the operation of the spiral-arm gravitational instability is sudden and is local, it naturally leads to isotropic velocity diffusion in the three spatial dimensions, which explains the observed correlation between the z-component and the r- and ϕ-components of the stellar random velocities (Wielen 1977). Furthermore, in a general analysis of the stellar diffusion through phase space, one of the most important conclusions drawn by Binney & Lacey (1988) was that the only stellar acceleration mechanism that could preserve the Gaussian distribution of random velocity and also produce a $t^{1/2}$ evolution is one in which the potential perturbation lasts significantly less than an epicycle period, so that the impulse approximation is valid. The spiral heating through a local gravitational instability certainly satisfies this criterion, and it differs from the stochastic spiral heating mechanism in that the effect of heating by a quasi-stationary spiral structure can last persistently and smoothly on the order of a Hubble time, which is needed to explain the smoothness of the age–velocity dispersion curve (Fig. 2) over the past 10^{10} yr.

2.2.3. *The Origin of the Radial Variation of the Stellar Velocity Dispersion with the Galactocentric Distance*

So far there is no complete measurement of the radial variation of the stellar velocity dispersion with the Galactocentric distance for our own Galaxy. However, indirect evidence indicates that stellar velocity dispersion decreases with Galactic radius (Gilmore et al. 1990, p. 194). The observation of external galaxies has further confirmed the radial decrease of the stellar velocity dispersion with galactocentric distance (van der Kruit & Freeman 1986). This radial decrease in stellar velocity dispersion can also be inferred from the observed constant scale height of the old-disk population versus galactic radius (van der Kruit & Searle 1981, 1982; Gilmore et al. 1990, p. 193).

The variation of stellar velocity dispersion with galactic radius has a simple explanation from the current theory. We have shown that the rate of increase of stellar velocity dispersion can be expressed as

$$\sigma \propto \sqrt{(\Omega - \Omega_p) F^2 v_c^2 \tan i \sin (2\phi_0)} .$$

This can be further simplified to $\sigma \propto [\Omega(r) - \Omega_p]^{1/2}$ for the inner-disk stars, since the product of the other factors is approximately constant for the inner-disk region (Appendix B). Therefore, for a nearly constant circular velocity v_c, we can express the radial dependence of the velocity dispersion for the inner disk region as

$$\sigma(r) = \sigma_\odot \sqrt{\frac{(v_c/\Omega_p r) - 1}{(v_c/\Omega_p r_\odot) - 1}} , \tag{13}$$

where the subscript "$\odot$" denotes the solar neighborhood values.

In Paper II we showed that there is heating outside corotation radius as well, and the effective diffusion constant is again described by equation (9). Furthermore, the heating at the corotation radius is not expected to be zero but, rather, smoothly joins the values at the inner and outer disks through the process of conduction (i.e., regular diffusion). Therefore, the overall $\sigma(r)$ profile is expected to resemble a single $r^{-1/2}$ function, i.e.,

$$\sigma(r) = \sigma_\odot \sqrt{\frac{r_\odot}{r}} . \tag{14}$$

In Figure 3, we plot the result of the radial velocity dispersion for the N-body simulation run of Paper II with 300,000 particles, at the beginning ($n = 0$) and the end ($n = 8000$) of the run. Also plotted was a $r^{-1/2}$ fit to the velocity dispersion curve at the end of the run. We see that the $r^{-1/2}$ curve fits very well the $n = 8000$ curve, and the tendency of the variation of the velocity dispersion distribution is such that as the simulation progresses the curve

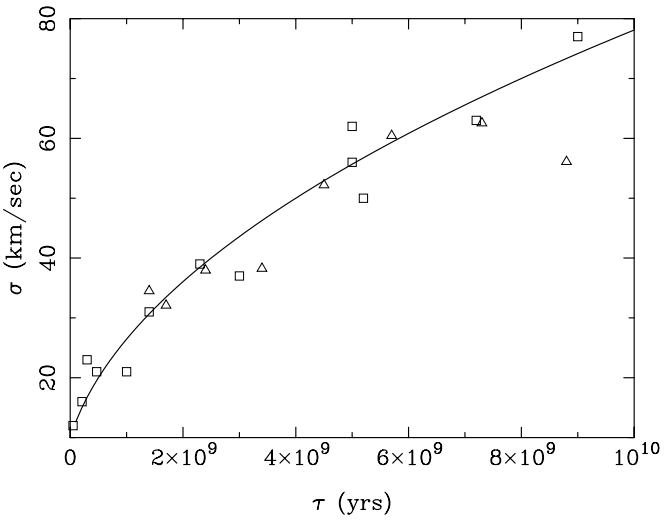

FIG. 2.—Dependence of the stellar velocity dispersion on the mean stellar age. In the figure, the squares are from Wielen (1977) and the triangles are from the Carlberg et al. (1985) tangential velocity measurement and scaled to the space velocities. The solid line is a fit of the form of equation (9) with $\sigma_0 = 10$ km s^{-1}, $D^{(3d)} = 6 \times 10^{-7}$(km s^{-1})^2 yr^{-1}.

looks more and more like a $r^{-1/2}$ curve throughout the galactic radii (i.e., the tail end at larger radii gradually lifts up).

In Figure 4, we plot the *increase* of the velocity dispersion in our previous N-body run, for time steps between 2000 and 5000 when the spiral pattern was found to be of constant amplitude. We see that the velocity dispersion increase was roughly constant except at the inner Lindblad resonance (ILR), located at $r = 10$, where the heating is much more effective than at other places. However, what was unmistakably clear is that heating is NOT confined to ILR, which, for this quasi-stationary spiral mode, is very sharply located at $r = 10$. This is thus an unambiguous piece of evidence that, for the majority of the spiral disk, heating is not due to the broadening of the resonances due to a time-varying spiral structure.

The $r^{-1/2}$ function fits very well the observed radial variation of stellar velocity dispersion for certain external galaxies (van der Kruit & Freeman 1986). Furthermore, if we fit an exponential to an $r^{-1/2}$ function, we obtain a fitted exponential length scale that is twice that for fitting a r^{-1}

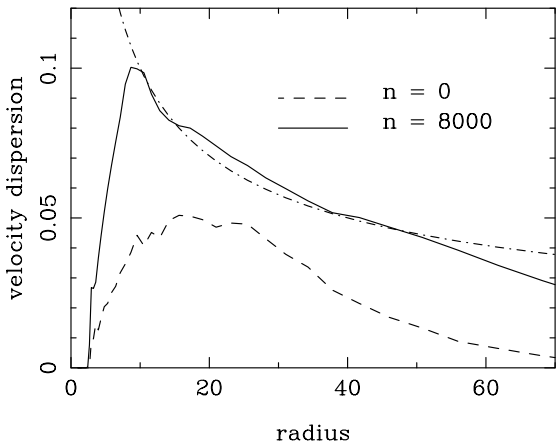

FIG. 3.—Radial velocity dispersion of the disk stars at the beginning and the end of an N-body simulation with $N = 300{,}000$ particles. The dot-dashed line is an $r^{-1/2}$ function fit to the $n = 8000$ curve.

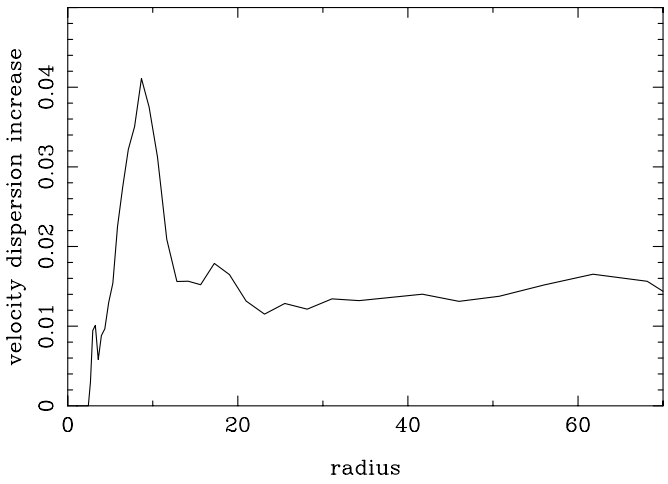

FIG. 4.—Differential heating during the time period from 2000 to 5000 steps, for an N-body spiral mode using $N = 300{,}000$ particles.

function. This explains the well-known observation that the exponential length scale of radial velocity dispersion for many disk galaxies is twice that of the exponential length scale for the surface density (Gilmore et al. 1990).

2.3. *Formation and Evolution of Galactic Bulges*

The proposed formation mechanisms for galactic bulges can be grouped into two main categories: the primordial bulge formation mechanisms and the evolutionary mechanisms (see the review by Wyse, Gilmore, & Franx 1997, as well as the many contributions in Dejonghe & Habing 1993). While the primordial bulge formation mechanisms would have a hard time explaining the observed stratification of ages and abundances of bulge stars, the earlier proposed secular evolutionary mechanisms focused mainly on the radial gas accretion and on the vertical pumping of stars through bar resonances, which usually have only limited radial extent.

In the previous sections, we have shown that, because of the spiral-induced collective dissipation process, an average star inside corotation tends to drift toward the central region of a galaxy. During its radial migration, all three components of the random velocity of a star gradually increase through the very same collective dissipation process. The radial profile of stellar velocity dispersion produced by this heating process has increasing velocity dispersion with decreasing galactic radius. Consequently, when the vertical component of a star's velocity dispersion becomes comparable to its circular velocity, the star will rise out of the galactic plane and become a bulge star.

As an example, we now estimate the rate of mass-accretion for the formation of the bulge of our Galaxy. Assuming that the spiral pattern during the past 10^{10} yr has an average pitch angle of 20° and 20% fractional amplitude around the solar neighborhood. Furthermore, using an average local surface density of $\Sigma \approx 60\ M_\odot$ pc^{-2} (Bahcall 1984; Kuijken & Gilmore 1989), we have that the rate of radial inflow of mass is

$$\frac{dM}{dt} = 2\pi r \frac{dr_*}{dt} \Sigma \approx 0.6\ M_\odot\ \mathrm{yr}^{-1}\ . \qquad (15)$$

This means that in a Hubble time the radial mass accretion across the disk will be able to accumulate an amount of mass toward the center of the Galaxy of the same order as

the mass of the Galactic bulge, which is about $10^{10}\ M_\odot$ (Gilmore et al. 1990, p. 224), especially considering the fact that the bulge was not devoid of matter at the beginning of the Galaxy formation.

The mean metallicity of the Galactic bulge stars is found to be twice the solar value. Some of the Galactic bulge stars are more metal rich than those found anywhere else in the Galaxy (Gilmore et al. 1990, p. 53). This is likely to be because the stars that are currently at the bulge region presumably started their radial migration close to the Galactic center region, and, because the dynamical timescale there is short, the material that currently makes up bulge stars went through more cycles of spiral- or bar-triggered star formation. This is also consistent with the general knowledge that although early-type spiral galaxies and elliptical galaxies contain large populations of old stars, the central regions of these galaxies also contain some of the most metal-rich stars anywhere in the universe (Barbuy & Grenon 1990).

One extra piece of information concerning the shapes of bulges for galaxies of varying Hubble types was noted by Andredakis, Peletier, & Balcells (1995). They found that the gradual variation of the surface density profile of bulges from late to early Hubble types can be fitted very well by the generalize exponential profile of Sersic (1968),

$$\Sigma(r) = \Sigma_0 \exp\left[-(r/r_0)^{1/n}\right], \qquad (16)$$

where, as the Hubble type of a galaxy goes from late to early, the fitted index n in the above expression also increases ($n = 1$ is for a pure exponential, and $n = 4$ is for a de Vaucouleurs law).

We note that our secular evolution theory gives a trend of variation of disk surface density that matches very well that found in the above mentioned result of Andredakis et al. (1995). As a galaxy evolves through the spiral-induced collective dissipation process, its surface density distribution (i.e., disk plus bulge) will become more and more centrally concentrated, together with the buildup of an extended outer envelope. This trend agrees with that of varying the n-index in the above bulge profile as the Hubble types of galaxies change from late to early. In Figure 5, we plot the evolution of surface density for our N-body simulation, which shows such a trend. Note that for this particular N-body mode the evolution speed is low because of the smaller spiral amplitude. A similar plot in Paper I shows a much higher evolution rate.

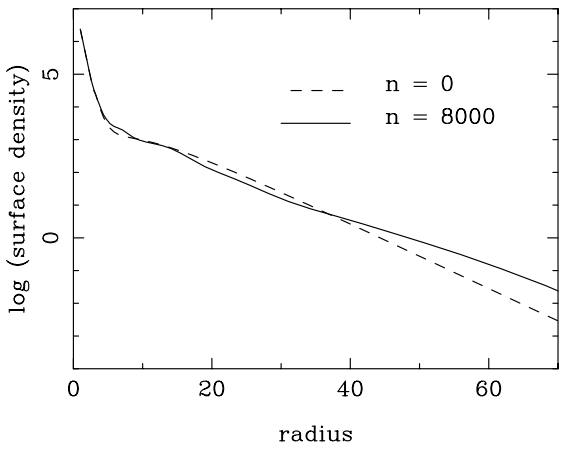

FIG. 5.—Smoothed surface density profiles of the disk plus bulge at the beginning and the end of an N-body simulation using 300,000 particles.

2.4. *What Allowed the Secular Evolution?*

Since the secular evolution effects we discussed in this section are so fundamental, readers of this paper may wonder why they were not borne out from many of the previous analyses on this subject. The reason for this was addressed in detail in Paper I, and the reader is referred to that paper for a thorough analysis. Here we merely repeat a few key arguments: most of the previous analytical solutions of galactic spiral problems were done either in the linear regime or as a non–self-consistent problem, i.e., as the study of the response of stellar orbits to an enforced perturbation potential, whereas, for the secular galactic evolution to be possible, we need to study the behavior of a collection of particles under their self-induced potential, which takes into account the graininess effect of particles at the location of the collective instabilities.

These are essentially the reasons why Lynden-Bell & Kalnajs (1972) did not find that stellar orbits have secular radial migration under a (non–self-consistent) quasi-stationary spiral structure. It is also the reason that Binney & Lacey (1988) could not diffuse stars in the vertical direction even under a transient wave. In these previous analyses, collective effects were ignored; therefore, the density wave and the basic state stars could interact only at the wave-particle resonances.

An objection one might raise to the current analysis is that a physical spiral galaxy is a three-dimensional object while our simulation result was obtained in two dimensions. We comment here that over the past three decades numerous work has been done in this area showing that the main physics of the spiral structure can be obtained in a two-dimensional analysis. Furthermore, the finite grid-cell size and the gravity-smoothing length used in the two-dimensional simulations made the disk under consideration effectively one with finite thickness. In actuality, the physical thickness of the galactic disk mainly serves to modify the kind of spiral mode present on the disk, and after that is determined, the phase shift–closure relation for a finite-thickness spiral disk should be the same as that for a thin disk. We emphasize here that in a collective process the global closure relation, such as the phase shift relation, is satisfied at all costs as long as the quasi steady state is reached; in the eventual self-consistent configuration, the morphology is fully compliant with the dynamics, so in a spiral galaxy the amount of energy dissipation and angular momentum transfer have to be given by the phase shift relation whether it is in two or three dimensions. The general trend of mass redistribution and secular heating that we inferred in this paper has recently been confirmed in a three-dimensional simulation of planetary disk evolution (Harbing 1997).

3. COSMOLOGICAL EVOLUTION OF GALAXIES

3.1. *Hubble Sequence as a Temporal Evolution Sequence*

The dynamical mechanism presented in this series of papers points to a trend of secular evolution of spiral galaxies, whereby on the timescale of a Hubble time the mass of the disk gradually becomes concentrated into a massive central bulge, together with the buildup of a diffused and extended outer envelope. The modal calculations of Bertin et al. (1989a, 1989b) show that the spiral modes supported by disks with massive central bulges are those with tightly wound spiral arms, consistent with the galaxies being of

earlier Hubble type. The secular spiral activity and spiral-triggered star formation also lead to the consumption of gas. Thus we see that all the major characteristics of the Hubble sequence can be naturally explained by a spiral-induced secular evolution of the basic state of the galactic disk.

The trend of morphological evolution of galaxies need not stop at Hubble type Sa, which is the earliest Hubble type in which a prominent spiral pattern is present. A central stellar bar could also induce secular mass accretion, through essentially the same mechanism as that underlying the spiral structure. This is because, when a bar is a true global instability, it is again a case of a dissipative structure. The dissipation in the bar is again achieved through large-scale (mainly collisionless) shocks, which are manifested as the offset dust lanes or the twisted isophotes often observed in galactic central regions. These skewed bars, as in the case of spiral patterns, lead to potential and density phase shift as well as the associated collective dissipation at a quasi steady state. The ram pressure created by the accretion flux in the outer disk could also drive further central accretion of matter.

We would like to stress here that when we state that the secular evolution trend of galaxies is from late to early Hubble types, the reverse argument may not be valid; i.e., not all early-type galaxies are evolved from the late types passing through every stage of the intermediate Hubble types. For example, the cD galaxies in the central regions of clusters are almost without doubt the consequence of mergers. Furthermore, some members of the local dwarf elliptical (dE) population (such as M32), which are the companions of normal galaxies, presumably have not gone through a dissipative collapse to form disks, because of the tidal field disturbance of their massive companions.

Ample observational evidence shows that there is systematic variation of the total mass, the bulge-to-disk ratio, the degree of rotational support, and the details of the light and color distribution across the Hubble sequence. Recent observations have shown (see, e.g., the contributions in Dejonghe & Habing 1993; Arnaboldi, Da Costa, & Saha 1997) that many elliptical galaxies contain faint disks and are partially rotationally supported. The qualitative distinction between bulges and ellipticals could more appropriately be described as a quantitative variation of the bulge-to-disk ratio (Franx 1993). Spectroscopic observations of the spiral galaxy bulges and the elliptical galaxies also show amazing similarities (see, e.g., Mihalas & Binney 1981, p. 364).

One important implication of the spiral-induced secular evolution scenario is that galaxies at different redshifts should show statistical differences in their morphologies. Such a statistical distribution of the morphological types of galaxies was indeed observed in both the so-called "Butcher-Oemler" effect (Butcher & Oemler 1978), which showed that the dominant galaxy type in distant clusters are disk-dominated late-type objects while in similar nearby clusters the dominant galaxy types are mostly S0 and elliptical, as well as in the *Hubble Space Telescope* Hubble Deep Field (HDF) and Medium Deep Survey (MDS), which showed that there is a clear excess of faint blue galaxies at intermediate redshifts (Ellis 1997)

In our current secular evolution picture, the rapid steepening of the morphology-environment relation with time as demonstrated in the contrast between the composition of the Butcher-Oemler clusters and that of the nearby clusters is likely to be a result of the role that galaxy interactions play in determining the rate of secular evolution. As we have seen in Paper II, the intrinsic evolution speed of a quasi-stationary spiral mode in a given basic state could sometimes be too small, because of the smallness of the wave amplitude F and pitch angle i, to give significant evolution effects on the timescale of a Hubble time. However, for galaxies that reside in the dense regions of groups or clusters, the encounters between these group or cluster members usually generate large-amplitude open density waves. Since the radial mass-accretion rate in the inner disk is proportional to the square of the wave amplitude and the square of the spiral pitch angle, during strong tidal interactions the evolution rate of disk galaxies could be enhanced by 1 or 2 orders of magnitude compared with the rate of evolution when they are isolated field galaxies.

The excess late-type objects found by the Hubble Space telescope could not have been the progenitors of the dwarf population in the local region, since the star formation rate indicated by the B-band luminosity of the faint blue galaxies would have made more stars than are contained in a typical dwarf elliptical in a single burst of 2×10^8 yr duration, whereas for the local dwarfs the data indicate that the secondary episodes of star formation took place over a few Gyr (Ferguson & Binggeli 1994). The secular evolution mechanism presented in the current papers indicates that the excess late-type objects found at intermediate redshifts are likely to be the progenitors of the intermediate-type spirals in the nearby universe.

3.2. *Comparison with Other Proposed Mechanisms of the Formation of the Hubble Sequence*

In the past, several other theories had been proposed to account for the formation of the Hubble sequence. In the following, we group them into three major categories and compare each with our current proposal.

3.2.1. *The Primordial Formation Scenarios*

The first major type (see, e.g., the review of Faber 1981) proposes that the Hubble sequence is essentially the result of varying initial conditions of the primordial clumps. There are also two subcategories in this group. One is the so-called "density picture," which is based on the notion that there must be some intrinsic width to the clustering locus on the cooling diagram owing to statistical fluctuations, so the variation of properties of galaxies in the Hubble sequence is a reflection of the variation of the intrinsic properties of the primordial clumps, and galaxies inherit most of their global properties at birth. The second scenario is based on the hierarchical clustering and dissipation theory for galaxy formation (White & Rees 1978), plus angular momentum gain via tidal torques (Fall & Efstathiou 1980). In this picture, the primordial clump of matter collapses because of energy dissipation. The Hubble sequence is then a sequence in the amount of collapse and condensation of the luminous matter relative to nonluminous halos.

The existing data on the structure of galaxies favor the dissipation picture over the density picture for the following two reasons (Faber 1981). (1) One of the more important features of the dissipation picture is its prediction that the fraction of nonluminous matter within the radius of ordinary matter should be greater in the late-type galaxies than in the earlier types. As emphasized by Tinsley (1981),

observed mass-to-light ratios of galaxies along the Hubble sequence do indeed display such a trend. (2) A second prediction of the dissipation picture comes from the dynamical temperature of the luminous matter within the self-gravitating regime. As a result of this process, at constant mass we would expect internal velocities in early-type galaxies to be higher than in late types. The observation of rotation velocities in early Hubble types does indeed show such a trend.

However, there remain several problems to which the dissipation picture could not provide a satisfactory answer and which the secular evolution scenario can naturally explain.

1. The dissipation picture explained away the energy issue but not the angular momentum issue. As is well known, along the Hubble sequence from late to early types, the mass of a galaxy systematically increases, while its specific angular momentum decreases. However, a celebrated result obtained by Peebles (1969) and Faber (1981) showed that the specific angular momentum acquired by the initial perturbation clump via tidal torquing has no correction with the size or mass of the clump. This negative result persists until today as one of the most serious challenges to the primordial scenario for Hubble sequence formation. In the secular evolution scenario, however, the angular momentum of the massive objects, which presumably collapsed to form disks early and started their spiral-induced secular evolution phase early, had already been transported out by the spiral density wave. The merger scenario is another way out of this dilemma.

2. In the dissipation picture, there is also the question of whether the dispersion in the specific angular momentum parameter λ in the tidal torque theory is large enough to account simultaneous for slow rotating ellipticals and rapidly rotating spirals. In the secular evolution picture, however, the spread in the angular momentum distribution reflects only the spread in the degree that the secular evolution has proceeded in a given galaxy, with the evolution speed itself determined by the mass of the galaxy, as well as by the environment in which the galaxy exists.

3. In the dissipation scenario, after the initial collapse phase the morphology of a galaxy stays roughly constant throughout the rest of its lifetime. The major subsequent evolution then is the luminosity evolution (Tinsley 1973). However, observational results show that there are significantly more blue, bright galaxies in the early universe than the luminosity evolution model can explain (see, e.g., Gunn 1981).

3.2.2. *The Merger Scenarios*

The hypothesis that mergers between late-type systems could produce early-type galaxies (Toomre & Toomre 1972; Schweizer 1983) is by now fairly well established. Much less settled, however, is the issue of whether mergers are responsible for producing all or most of the elliptical galaxies, not to say the entire Hubble sequence. Observationally, it is well known that elliptical galaxies show a clear bimodal distribution in many internal properties. The lower luminosity ones ($M_V > -21$) have steep power-law cusps in their inner regions, are disky in their large-scale structure, show relatively rapid rotation, show little evidence of past merger, and are nearly oblate and isotropic. The brighter ellipticals have shallow power-law cusps, are boxy, rotate relatively slowly, and are relatively triaxial and anisotropic (see, e.g., Freeman 1997). The merger picture alone can hardly explain the formation of this bimodal distribution. In the following, we present further evidence that the merger scenario cannot possibly be responsible for the formation of the entire Hubble sequence.

1. As already pointed out earlier by Ostriker (1980) and Tremaine (1981), the merger picture operating alone faces difficulty in explaining the internal spectroscopic distributions in ellipticals, the galaxy core radii changes before and after collision, and the environments where the various types of galaxies are usually found.

2. A merger scenario alone cannot explain the smooth transition of characteristics of all Hubble types, especially not their internal structure and kinematics, including the bulge-disk connection (Courteau, de Jong, & Broeils 1996; Peletier & Balcells 1996); the color and abundance gradient (Strom et al. 1976; Gonzalez & Gorgas 1996, and references therein), and the systematic increase of the specific frequency of the globular clusters from late to early Hubble types (Harris & van den Bergh 1981).

3. As discovered by Bender et al. (1989), there exists strong correlation between the isophote shapes and the radio and X-ray properties of elliptical galaxies. Objects that are radio loud and/or surrounded by gaseous X-ray halos generally have boxy or irregular isophotes. Elliptical galaxies with pointed isophotes (usually interpreted to contain weak edge-on disks) are mostly radio quiet and show no X-ray emission in excess of the discrete source contribution.

4. The simulations of mergers (Heyl, Hernquist, & Spergel 1996, and the references therein) have so far produced merger remnants that are predominantly triaxial and often show large misalignments of the rotation axes and the minor axes of the figures, whereas observationally a large fraction of the elliptical galaxies are found to be oblate rotators (see, e.g., Statler 1996) and most of them have axis misalignment that are small (Franx, Illingworth, & de Zeeuw 1991). Adding gas into the progenitors was found to create its own agonies (Mihos & Hernquist 1994).

Besides major mergers of two equal-sized late-type galaxies as a mechanism for the formation of early-type systems, the accretion of small satellites has also been proposed as a mechanism for the formation and growth of galactic bulges (Walker, Mihos, & Hernquist 1996). While satellite accretion is certainly a process that is going on among many galaxies, it again could not explain the observed radial abundance gradient in bulge stars and the bulge-disk connection, both in color and in kinematics (Courteau et al. 1996; Peletier & Balcells 1996).

3.2.3. *Other Secular Evolution Proposals*

The secular evolution scenario discussed in this series of papers, which was first proposed in Zhang (1992), has also been proposed in similar forms in relation to bar-driven gas inflow (Pfenniger & Norman 1990; Combes et al. 1990; Pfenniger & Friedli 1991; Kormendy 1993; Friedli & Benz 1993; Norman, Sellwood, & Hasan 1996; Courteau et al. 1996), mass inflow driven by perturbers (Byrd et al. 1986; Elmegreen, Elmegreen, & Bellin 1990; Moore et al. 1996), and the evolution of disk morphology due to the presence of certain unknown viscous dissipation mechanisms (Silk & Norman 1981; Lin & Pringle 1987). In the following, we

highlight the relation and the major differences of our current proposal from the previous secular evolution proposals.

1. Many of the secular evolution proposals in the past were based on the results of Lynden-Bell & Kalnajs (1972) that a nonaxisymmetric global pattern transports angular momentum outward through gravitational torques. However, as we have shown in Paper I and Paper II, if a global mode merely transports angular momentum outward and does not couple to the basic state, it leads merely to the modal growth without a significant redistribution of the disk material. This in fact is the reason that, several decades after the work of Linden-Bell & Kalnajs (1972), it is still generally agreed that a spiral disk undergoes no secular evolution. Therefore the true dynamical mechanism responsible for the secular evolution of the basic state of the galactic disk had not been properly identified in these past proposed evolution mechanisms. In our current work, the coupling to the basic state by a quasi-stationary global pattern is achieved through a collisionless shock near the potential minimum of the pattern.

2. Many of the past secular evolution discussions related to the role of gas did not realize that it is not the intrinsic viscosity of the gas which leads to secular evolution. The effective viscosity present in the nonaxisymmetric patterns is collective and gravitational in nature; therefore, it does not depend on whether the underlying mass is composed of gas or stars. The interaction of the wave with both gas and stars requires the operation of collisionless shocks, and it is the scattering of matter in the local gravitational instability on the spiral arms, not the actual collisions of matter, that produces the majority of the viscosity. This is also true for the gas component.

The secular evolution mechanism presented in this series of papers is expected to be one of the most important mechanisms in producing the Hubble sequence. This is both because it is a physical mechanism intrinsic to the individual galaxy and is operating throughout the lifetime of a galaxy—including during and after a merger—irrespective of the environment it is in and because of the recognition of spiral structure in disk galaxies as an example of the general class of "dissipative structures" in nonequilibrium systems (Paper II). The important function of any dissipative structure is to greatly accelerate the speed of entropy evolution of the underlying nonequilibrium system, and this special function is also achieved by the galactic spiral structure; i.e., the emergence of a galactic spiral structure greatly accelerated the speed of entropy evolution of a disk galaxy (Zhang 1992), with the direction of such an evolution along the Hubble sequence from late to early types. This finally answered the decades-old question of Lynden-Bell & Kalnajs (1972): "Why are galaxies spiral at all?" It appears that nature puts a spiral in disk galaxies in order to facilitate their evolution away from the extreme nonequilibrium state.

The above being said, we need to point out, however, that, in addition to the secular evolution process, all of the other mechanisms mentioned in this section, i.e., the variation of the initial density perturbation, the initial dissipative collapse, as well as mergers of spiral galaxies, satellite and gas accretion, etc., are also expected to be operating with varying strengths and contributing to the evolution of the Hubble sequence. Furthermore, some of the amorphous protogalaxies seen in the HDF images maybe a class in their own right; i.e., they are neither the merger product of two late-type disks nor galaxies that are evolving into disk galaxies. Some of these objects may bypass the phase of the formation of a well-defined thin disk and evolve into spheroidal systems with the assistance of some three-dimensional global instability structures through mechanisms of angular momentum transport similar to what we have discussed here for the quasi–two-dimensional disks. This is supported by the observational fact that many of the high-z amorphous galaxies in the HDF images possess twisted isophotes in their luminosity distribution (Fasano, Filippi, & Bertola 1997). A central thread that connects all of the secular evolution scenarios, however, is the principle of increase of entropy for an open and nonequilibrium system.

3.3. *Secular Evolution as the Underlying Mechanism for Producing the Generalized Tully-Fisher/Faber-Jackson and the Fundamental Plane Relations*

There are the well-known scaling laws relating the radii and internal velocities of galaxies to their luminosities: the Tully & Fisher (1977) relation for spirals and Faber & Jackson (1976) relation for ellipticals, which are very similar in form. This similarity in behavior between two families of galaxies having radically different internal structures is a remarkable fact in need of explanation. In the secular evolution picture, we know that during the radial mass-accretion process, approximately one-half of the orbital energy of the accreting stars becomes their random motion energy in the final configuration (the other one-half being transported outward by the spiral density wave), and the final radius of the evolved galaxy is also proportional to its initial radius. Therefore the scaling law for a galaxy in the spiral phase should be mostly preserved as it evolves from late to early type, except for a change of the constant multiplication factor. Observationally, it was indeed found that the combined scaling law for elliptical and spiral galaxies can be written as the following form of a generalized Tully-Fisher/Faber-Jackson relation:

$$L_{\rm galaxy} = K_h \times V_{\rm int}^x \tag{17}$$

(Williams 1997), where $L_{\rm galaxy}$ is the total intrinsic luminosity of a galaxy, $V_{\rm int}$ the internal velocity dispersion of the galaxy, the exponent x depends on the band of observation, and the proportionality constant K_h depends on the Hubble types of galaxies.

The secular evolution scenario provides also a natural explanation for the existence of the fundamental plane relation of elliptical galaxies (Dressler et al. 1987; Djorgovski & Davis 1987). The fundamental plane of elliptical galaxies describes a set of bivariate correlations connecting a set of global observables, such as radii, luminosities, surface brightness, velocity dispersions, masses, densities, etc. In a parameter space defined by at least three independently measured quantities from this set, ellipticals occupy a two-dimensional surface, which is close to a plane if logarithmic quantities are used.

Although physical relations such as the virial theorem appears to underlie the fact that elliptical galaxies of various characteristics can find their quasi-equilibrium configurations along the fundamental plane (see, e.g., Djorgovski,

Pahre, & de Carvalho 1996), a question remains as to why, for a protogalactic clump of a given mass, considerably different equilibrium configurations are chosen. From our secular evolution scenario, it can be easily seen that the reason for the existence of varying configurations for a given mass is that this indicates *the varying degree of advancement of secular angular momentum evolution.* Furthermore, the fact that, for a given mass range, there could be a whole spectrum of the degrees of advancement of secular evolution is likely to be due to the influence of environment on the speed of secular evolution. In other words, in the single independent parameter case, such as the Tully-Fisher or Faber-Jackson relation with a constant proportionality K, the initial condition alone determines how fast a galaxy evolves. Whereas in reality the added factor of the environment comes into play in determining the rate (and the current degree) of secular evolution.

Viewed in this light, we see that the variation of the "constant" K factor with Hubble types in the generalized Tully-Fisher/Faber-Jackson relation (eq. [17]) is consistent with the need for a second independent parameter such as surface brightness in the fundamental plane relation, and both the fundamental plane relation and the generalized Tully-Fisher/Faber-Jackson relation are a reflection of galaxy evolution under the influence of both nature and nurture.

4. CONCLUSIONS

We have shown that a quasi-stationary spiral structure can induce significant secular evolution effects during the lifetime of a spiral galaxy. These effects include the secular redistribution of disk matter and the secular heating of disk stars, both of which lead to the growth of the galactic bulges. We argue that the dynamical mechanism presented in this series of papers, together with the overwhelming observational evidence, points to a trend of galaxy evolution from late to early Hubble types. This evolution trend is consistent with the general trend of entropy evolution in self-gravitating systems. It also provides one of the possible explanations for the observed morphological evolution of the statistical properties of galaxies found in the *Hubble Space Telescope* Medium Deep Survey and Hubble Deep Fields, as well as in ground-based observations conducted in the past two decades.

The author wishes to thank Françoise Combes, Bruce Elmegreen, C. C. Lin, Colin Masson, Frank Shu, Antony Stark, Alan Toomre, and Scott Tremaine, as well as referee Daniel Friedli and editor Greg Bothun, for helpful discussions. This work was partially supported by grants from the Smithsonian Astrophysical Observatory's SMA and AST/RO projects.

APPENDIX A

RATE OF SECULAR EVOLUTION: A CLOSURE RELATION AT THE QUASI STEADY STATE

In Appendix A of Paper II, we have shown that, in the continuum limit, the rate of angular momentum change in an annular ring is related to the total torque coupling integral $C(r)$ through

$$-2\pi r \frac{d\overline{L}}{dt}(r) = \frac{dC(r)}{dr} = \frac{d[C_g(r) + C_a(r)]}{dr}, \tag{18}$$

where $\overline{L}$ is the azimuth-averaged angular momentum density at a particular radius; C_g is the gravitational torque coupling given by

$$C_g(r) = \frac{1}{4\pi G} r \int_{-\infty}^{\infty} \int_0^{2\pi} \frac{\partial \mathscr{V}}{\partial \phi} \frac{\partial \mathscr{V}}{\partial r} d\phi\, dz, \tag{19}$$

where $\mathscr{V}$ is the (smoothed) gravitational potential and G is the constant of gravitation; and C_a is the advective torque coupling, or coupling through Reynolds stress, owing to the large-scale velocity distribution produced by the spiral structure, given by the expression

$$C_a(r) = r^2 \int_0^{2\pi} \Sigma V_r V_\phi\, d\phi, \tag{20}$$

where Σ is the surface density, and where V_r and V_ϕ are, respectively, the radial and azimuthal velocity perturbations relative to the circular velocity.

We want to ask now that at the quasi steady state of the wave mode, how can we relate the above rate of angular momentum change with the parameters of the spiral mode under concern? In another words, what is dC/dr equal to in terms of spiral parameters?

In Paper I and Paper II we have introduced an alternative expression of a torque integral (which is different from the torque coupling integral C),

$$T_1(r) = r \int_0^{2\pi} \Sigma_1 \frac{\partial \mathscr{V}_1}{\partial \phi} d\phi, \tag{21}$$

which expresses the total torque experienced by the basic state stars in an annular ring due to the torquing of the perturbation spiral potential. We now want to know how $T_1(r)$ and $C = C_a + C_g$ are related in the different regimes of the wave mode.

We have argued in Paper II that, although in the linear regime $T_1(r) = dC_g/dr$, this equality does not hold away from the linear regime, after the spiral shock forms, because of the fact that one of the relations needed in the proof of this equality, i.e., the differential form of the Poisson equation, is no longer valid between the gravitational potential at an arbitrary position and the averaged volume density of the disk at the corresponding location.

The results of N-body simulations (Paper II) indicate that at the quasi steady state of the wave mode, the appropriate closure relation governing the wave and basic state energy and angular momentum exchange is in fact

$$dC/dr = d(C_a + C_g)/dr = T_1(r) \; ; \tag{22}$$

we now give a more rigorous proof of this relation at the quasi steady state.

We begin by observing that in the case of a particle disk, the Reynolds stress indicated by C_a is also gravitational in nature, since in a collisionless disk the momentum transfer contained in the Reynolds stress expression is mediated through the short- to intermediate-range gravitational field of the particles constituting the spiral gravitational instability. Therefore, if we also take into account the particle discreteness effect and its contribution to the potential, the net torque experienced by particles in an annular ring should be purely gravitational in nature and there is not a separate advective contribution.

Thus we can write the total torque experienced by particles in a unit-width annular ring, which is equal to the rate of the angular momentum change, as

$$2\pi r \frac{d\bar{L}}{dt}(r) = \sum_i \boldsymbol{r}_i \times \boldsymbol{F}^i = \int_{\text{annular ring}} \sum_i \boldsymbol{r}_i \times \delta^2(\boldsymbol{r} - \boldsymbol{r}_i) m^i \left(- \frac{\partial \tilde{\psi}_i}{\partial \boldsymbol{r}} \right) d^2 r = \int_{\text{annular ring}} \boldsymbol{r} \times \Sigma(\boldsymbol{r}) \left(- \frac{\partial \tilde{\psi}_i}{\partial \boldsymbol{r}} \right) d^2 r \; , \tag{23}$$

where the summation i is for all the particles in the annulus, $\tilde{\psi}_i$ is the potential at the position $\boldsymbol{r}_i$ of the ith particle due to the rest of the particles in the disk, the integral is over all $\boldsymbol{r} = (x, y, 0)$ in the annulus, and $\Sigma \equiv \sum_i (\boldsymbol{r}) m^i \delta^2(\boldsymbol{r} - \boldsymbol{r}_i)$.

Taking the continuum limit, and using arguments similar to that leading from equation (1) to equation (4) in Paper I, we have thus shown that

$$2\pi r \frac{d\bar{L}}{dt}(r) = \sum_i \boldsymbol{r}_i \times \boldsymbol{F}^i = \boldsymbol{r} \cdot \int \Sigma_1(r, \phi) \frac{\partial \mathscr{V}_1}{\partial \phi}(r, \phi) d\phi \equiv - T_1(r) \; ; \tag{24}$$

i.e., $- T_1(r)$ gives the total torque that the particles in a unit width annular ring experiences.

At the quasi steady state, since this interaction is irreversible in nature, all the orbital angular momentum lost by this group of particles is permanently lost. Therefore, in terms of what the basic state stars experience at the quasi steady state, it is much more natural for us to adopt a Lagrangian point of view; i.e., $T_1(r)$ is the total loss/gain of basic state angular momentum of this group of stars in the annulus at a particular instant, and as a result of the action of $T_1(r)$ the stars will stream out of this annulus at the next instant; however, for the wave, on the other hand, since it usually consists of different particles and since it is at a quasi steady state, it is more natural for us to adopt a Eulerian point of view; i.e., we consider how much angular momentum the wave leaves onto each annular ring by calculating the gradient of $C_a(r) + C_g(r)$. At the quasi steady state, by definition all the angular momentum deposited by the wave is deposited onto the basic state matter, so we come to the inevitable conclusion that $d[C_a(r) + C_g(r)]/dr = T_1(r)$ at the quasi steady state.

APPENDIX B

RELATION BETWEEN THE RADIAL PROFILE OF TOTAL TORQUE-COUPLING AND THE SECULAR CHANGE OF DISK SURFACE DENSITY

We have shown in Paper II that

$$- \frac{1}{2\pi r} \frac{dC(r)}{dr} = \frac{d\bar{L}}{dt} \; , \tag{25}$$

where $\bar{L}$ is the azimuthally averaged local angular momentum density and C is the total torque coupling integral. In the quasi steady state of the wave, the (positive or negative) angular momentum deposited by the density wave is left entirely onto the basic state stars, which leads to the secular change of the mean stellar orbit,

$$\frac{d\bar{L}}{dt} = v_c \Sigma \frac{dr_*}{dt} \; , \tag{26}$$

where v_c is the circular velocity, Σ the surface density of the basic state of the disk, and dr_*/dt the rate of change of mean stellar radius.

Using equation (25) in equation (26), we have for the rate of orbital change

$$\frac{dr_*}{dt} = - \frac{1}{2\pi r v_c \Sigma} \frac{dC}{dr} \; . \tag{27}$$

Now the equation of continuity (eq. [7] in the main text) gives

$$2\pi r \frac{\partial \Sigma}{\partial t} = -\frac{\partial}{\partial r}\left(\frac{dr_*}{dt}\Sigma 2\pi r\right). \tag{28}$$

Using equation (27) in equation (28), we obtain

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{2\pi r}\frac{d}{dr}\left(\frac{1}{v_c}\frac{dC}{dr}\right). \tag{29}$$

For nearly constant v_c with r, equation (29) is solved by

$$\Sigma(r, t) = \frac{v_c}{2\pi r}\frac{d^2C(r)}{dr^2}\, t + \Sigma_0(r)\ . \tag{30}$$

In Paper II, we showed that the total torque coupling C is generally of an asymmetric bell shape, with the top of the bell located at the corotation radius of the relevant spiral mode. Such a shape can be described by the following composite Gaussian functions:

$$C(r) = \begin{cases} C_0\, e^{-(r-r_{\rm co})^2/r_1^2} & \text{for } r \le r_{\rm co} \\ C_0\, e^{-(r-r_{\rm co})^2/r_2^2} & \text{for } r \ge r_{\rm co}\ , \end{cases} \tag{31}$$

where $r_{\rm co}$ is the corotation radius.

We therefore have

$$\frac{d^2C}{dr^2} = \begin{cases} \left[\dfrac{-2C_0}{r_1^2} + \dfrac{4C_0(r-r_{\rm co})^2}{r_1^4}\right] e^{-(r-r_{\rm co})^2/r_1^2} & \text{for } r \le r_{\rm co} \\ \left[\dfrac{-2C_0}{r_2^2} + \dfrac{4C_0(r-r_{\rm co})^2}{r_2^4}\right] e^{-(r-r_{\rm co})^2/r_2^2} & \text{for } r \ge r_{\rm co}\ . \end{cases} \tag{32}$$

We thus see that d^2C/dr is positive for $r < r_{\rm co} - r_1/\sqrt{2}$ and $r > r_{\rm co} + r_2/\sqrt{2}$ and that it is negative for $r_{\rm co} - r_1/\sqrt{2} < r < r_{\rm co} + r_2/\sqrt{2}$. Referring back to equation (30), we see that this sign distribution of d^2C/dr^2 leads to an increase of Σ with time for $r < r_{\rm co} - r_1/\sqrt{2}$ and $r > r_{\rm co} + r_2/\sqrt{2}$ and a decrease of Σ with time for $r_{\rm co} - r_1/\sqrt{2} < r < r_{\rm co} + r_2/\sqrt{2}$. This indicates that both the central region and the outer skirt of the disk will accumulate matter with time and the surface density of the region near the corotation radius will decrease with time. Such a trend of surface density evolution was indeed that observed in our N-body simulation result shown in Figure 5.

The above derivation assumed that C_0, as well as $r_{\rm co}$, r_1, and r_2, is a constant. Over a longer time span, the evolution of the basic state is expected to alter the kind of spiral mode present also; therefore, we expect that all of these "constants" will change slowly with time. In particular, $r_{\rm co}$ and r_2 are expected to increase with time; r_1 will also increase slightly, but it cannot exceed $r_{\rm co}$ itself.

We now derive an approximate result that will be used in the main text. For a galaxy with a flat rotation curve and with $\Sigma(r) \propto 1/r$, we have $dr_*/dt \propto dC/dr$, from equation (27). For a certain radial range of the inner disk region, we can write approximately $dC/dr \sim$ constant, which leads to $dr_*/dt \sim$ constant independent of radius. Now, since we also have

$$\frac{dr_*}{dt} = -\frac{1}{2}F^2 v_c \tan i \sin(m\phi_0)\ ,$$

which was equation (1) in the main text, we see that for such a galaxy the product of $F^2v_c^2 \tan i \sin(m\phi_0)$ is also approximately constant, a result we will use in § 2 of the main text.

REFERENCES

Andredakis, Y. C., Peletier, R. F., & Balcells, M. 1995, MNRAS, 275, 874
Arnaboldi, M., Da Costa, G. S., & Saha, P. 1997, ASP Conf. Ser. 116, The Second Stromlo Symposium: The Nature of Elliptical Galaxies, ed. M. Arnaboldi, G. S. Da Costa, & P. Saha (San Francisco: ASP)
Bahcall, J. N. 1984, ApJ, 287, 926
Barbanis, B., & Woltjer, L. 1967, ApJ, 150, 461
Barbuy & Grenon 1990, in ESO Conf. Proc. 35, Bulges of Galaxies, ed. B. J. Jarvis & D. M. Jerndrup (Garching: ESO), 83
Bender, R., Surma, P., Dobereiner, S., Mollenhoff, C., & Madejsky, R. 1989, A&A, 217, 35
Bertin, G., Lin, C. C., Lowe, S. A., & Thurstans, R. P. 1989a, ApJ, 338, 78
———. 1989b, ApJ, 338, 104
Bettoni, D., Fasano, G., Kjaergaard, P., & Moles, M. 1997, in ASP Conf. Ser. 116, The Second Stromlo Symposium: The Nature of Elliptical Galaxies, ed. M. Arnaboldi, G. S. Da Costa, & P. Saha (San Francisco: ASP), 71
Binney, J., & Lacey, C. 1988, MNRAS, 230, 597
Butcher, H., & Oemler, A., Jr. 1978, ApJ, 219, 18
Byrd, G. G., Valtonen, M. J., Sundelius, B., & Valtaoja, L. 1986, A&A, 166, 75
Carlberg, R. G., Dawson, P. C., Hsu, T., & Vandenberg, D. A. 1985, ApJ, 294, 674
Carlberg, R. G., & Sellwood, J. A. 1985, ApJ, 292, 79
Combes, F., Debbash, F., Friedli, D., & Pfenniger, D. 1990, A&A, 233, 82
Couch, W. J., Barger, A. J., Smail, I., Ellis, R., & Sharples, R. M. 1998, ApJ, 497, 188
Couch, W. J., Ellis, R. S., Sharples, R. M., & Smail, I. 1994, ApJ, 430, 121
Courteau, S., de Jong, R. S., & Broeils, A. H. 1996, ApJ, 457, L73
Dehnen, W., & Binney, J. J. 1998, MNRAS, 298, 387
Dejonghe, H., & Habing, H. J. 1993, IAU Symp. 153, Galactic Bulges (Dordrecht: Kluwer)
de Vaucouleurs, G. 1959, in Handbuch der Physik, vol. 53, ed. S. Flugge (Berlin: Springer), 275
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Djorgovski, S. G., Pahre, M. A., & de Carvalho, R. R. 1996, in ASP Conf. Ser. 86, Fresh Views of Elliptical Galaxies, ed. A. Buzzoni, A. Renzini, & A. Serrano (San Francisco: ASP), 225
Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R. J., & Wegner, G. 1987, ApJ, 313, 42
Ellis, R. 1997, ARA&A, 35, 389

Elmegreen, D. M., Elmegreen, B. G., & Bellin, A. D. 1990, ApJ, 364, 415
Faber, S. M. 1981, in Astrophysical Cosmology, ed. H. A. Bruck, G. V. Coyne, & M. S. Longair (Vatican City: Pontificia Acad. Sci.), 191
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Fall, S. M., & Efstathiou, G. 1980, MNRAS, 193, 189
Fasano, G., Filippi, M., & Bertola, F. 1997, in ASP Conf. Ser. 116, The Second Stromlo Symposium: The Nature of Elliptical Galaxies, ed. M. Arnaboldi, G. S. Da Costa, & P. Saha (San Francisco: ASP), 557
Ferguson, H. C., & Binggeli, B. 1994, ARA&A, 6, 67
Franx, M. 1993, in IAU Symp. 153, Galactic Bulges, ed. H. Dejonghe & H. J. Habing (Dordrecht: Kluwer), 243
Franx, M., Illingworth, G., & de Zeeuw, T. 1991, ApJ, 383, 112
Freeman, K. C. 1970, ApJ, 160, 811
———. 1997, in ASP Conf. Ser. 116, The Second Stromlo Symposium: The Nature of Elliptical Galaxies, ed. M. Arnaboldi, G. S. Da Costa, & P. Saha (San Francisco: ASP), 1
Friedli, D., & Benz, W. 1993, A&A, 268, 65
Gilmore, G., King, I. R., & van der Kruit, P. C. 1990, The Milky Way as a Galaxy, Saas-Fee Advanced Course Lecture Notes 19, (Mill Valley, CA: University Science)
Gliese, W. 1969, Veroff. Astron. Rechen-Inst. Heidelberg, 22
Gonzalez, J. J., & Gorgas, J. 1996, in ASP Conf. Ser. 86, Fresh Views of Elliptical Galaxies, ed. A. Buzzoni, A. Renzini, & A. Serrano (San Francisco: ASP), 225
Gunn, J. 1981, in Astrophysical Cosmology, ed. H. A. Bruck, G. V. Coyne, & M. S. Longair (Vatican City: Pontificia Acad. Sci.), 233
Harbing, H. 1997, unpublished presentation given at IAU General Assembly, Kyoto, 1997 August 17–30
Harris, W. E., & van den Bergh, S. 1981, AJ, 86, 1627
Heyl, J. S., Hernquist, L., & Spergel, D. N. 1996, ApJ 463, 69
Hubble, E. 1936, Realm of the Nebulae (New Haven: Yale Univ. Press)
Jeans, J. H. 1928, Astronomy and Cosmology (Cambridge: Cambridge Univ. Press)
Kormendy, J. J. 1982, in Morphology and Dynamics of Galaxies, 12th Advanced Course of the SSAA, ed. L. Martinet & M. Mayor (Geneva: Geneva Obs.), 113.
Kuijken, K., & Gilmore, G. 1989, MNRAS, 239, 605
Lin, D. N. C., & Pringle, J. E. 1987, ApJ, 320, L87
Lynden-Bell, D., & Kalnajs, A. J. 1972, MNRAS, 157, 1
Martinet, L. 1995, in Fundam. Cosmic Phys., 15, 341
Mihalas, D., & Binney, J. 1981, Galactic Astronomy (New York: Freeman)
Mihos, J. C., & Hernquist, L. 1994, ApJ, 437, L47
Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A., Jr. 1996, Nature, 379, 613
Norman, C. A. 1984, in Formation and Evolution of Galaxies and Large Structures in the Universe (Dordrecht: Reidel), 327
Norman, C. A., Sellwood, J. A., & Hasan, H. 1996, ApJ 462, 114
Ostriker, J. P. 1980, Comm. Astrophys., 8, 177
Peebles, P. J. E. 1969, ApJ, 155, 393
Peletier, R. F., & Balcells, M. 1996, AJ, 111, 2238
Pfenniger, D. 1996, in Barred Galaxies and Circumnuclear Activity, Nobel Symp. 98, Lecture Notes in Physics, vol. 474, ed. Aa. Sandqvist, & P. O. Lindblad (New York: Springer), 91
Pfenniger, D., & Friedli, D. 1991, A&A, 252, 75
Pfenniger, D., & Norman, C. 1990, ApJ 363, 391
Pringle, J. E. 1981, ARA&A, 19, 137
Schwarzschild, K. 1907, Gottingen Nachr., 614
Schweizer, F. 1983, in IAU Symp. 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), 319
Seiden, P. E., Schulman, L. S., & Elmegreen, B. G. 1984, ApJ, 282, 95
Sersic, J. L. 1968, Atlas de Galaxias Australes (Cordoba: Obs. Astron.)
Silk, J., & Norman, C. 1981, ApJ, 247, 59
Spitzer, L., & Schwarzschild, M. 1951, ApJ, 114, 385
———. 1953, ApJ, 118, 106
Statler, T. S. 1996, in ASP Conf. Ser. 87, Fresh Views of Elliptical Galaxies, ed. A. Buzzoni, A. Renzini, & A. Serrano (San Francisco: ASP), 27
Strom, S. E., Strom, K. M., Goad, J. W., Vrba, F. J., & Rice, W. 1976, ApJ, 204, 684
Tinsley, B. M. 1973, ApJ, 186, 35
———. 1981, ApJ, 250, 758
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Tremaine, S. 1981, in The Structure and Evolution of Normal Galaxies, ed. S. M. Fall & D. Lynden-Bell (Cambridge: Cambridge Univ. Press), 67
Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661
van der Kruit, P. C., & Freeman, K. C. 1986, ApJ, 303, 556
van der Kruit, P. C., & Searle, L. 1981, A&A, 95, 105
———. 1982, A&A, 110, 61
Walker, I. R., Mihos, J. C., & Hernquist, L. 1996, ApJ, 460, 121
White, S. M., & Rees, M. J. 1978, MNRAS, 183, 341
Wielen, R. 1977, A&A, 60, 263
Wielen, R., & Fuchs, B. 1990, in Dynamics and Interactions of Galaxies, ed. R. Wielen (Berlin: Springer), 318
Williams, D. 1997, unpublished presentation given at IAU 23rd General Assembly, Kyoto, 1997 August 17–30
Wyse, F. G., Gilmore, G., & Franx, M. 1997, ARA&A, 35, 637
Young, J. S. 1990, in The ISM in External Galaxies, The Second Wyoming Conf., ed. H. Thornson & M. Shull (Washington, DC: NASA), 67
Zhang, X. 1992, Ph.D. thesis, Univ. California, Berkeley
———. 1996a, ApJ, 457, 125 (Paper I)
———. 1996b, in ASP Conf. Ser. 121, Accretion Phenomena and Related Outflows, IAU Colloq. 163, ed. D. Wickramasigne, G. Birknell, & L. Ferrario (San Francisco: ASP), 840
———. 1998, ApJ, 499, 93 (Paper II)