DISSIPATIONAL GALAXY FORMATION: CONFRONTATION WITH OBSERVATIONS

JOSEPH SILK

Department of Astronomy, University of California, Berkeley

AND

COLIN NORMAN

Huygens Laboratory, University of Leiden

Received 1980 January 2; accepted 1980 December 12

ABSTRACT

We explore the hypothesis that a protogalaxy consists of an aggregate of interacting gas clouds, and undergoes mergers with neighboring systems as envisaged in both hierarchical clustering and fragmentation schemes for galaxy formation. Such a system of bound gas clouds behaves like test particles in an N -body simulation, while also undergoing such dissipative processes as internal star formation and inelastic cloud collisions. Both gaseous dissipation and violent relaxation play fundamental roles in our galaxy formation model: such a hybrid approach seems necessary in order to be able to account for the diverse structural and dynamical properties of spheroidal galaxies, ranging from velocity anisotropy to metallicity gradients. Protogalaxy mergers fuel the cores of the merged systems with gas clouds, resulting in the development of centrally condensed cores that are limited in binding energy by the velocity-dependent inelasticity of the colliding clouds. Isolated systems preferentially develop rotationally supported bulges and exponential disks, whereas much of the enriched gas is stripped during galaxy collisions in clusters. Protogalaxy mergers during the initial stages of galaxy clustering can account for the observed spatial distribution of spiral, S0, and elliptical galaxies. While many of the arguments developed here are equally valid for mergers between spiral galaxies, the need for efficiently converting large amounts of gas into stars in order to generate metallicity gradients and centrally condensed cores of spheroidal galaxies leads us to favor protogalaxy mergers. Galaxy formation can occur slowly and at late epochs, the time scale for disks to form being $\lesssim 10^{10}$ year. Spheroidal systems and S0's acquire their morphology at the epoch of galaxy clustering, $z = 5-10$, the epoch of protogalaxy formation, when these systems undergo final bursts of star formation, being extended, as a result of successive mergers, over a characteristic cluster crossing time of $\sim 10^9$ years.

Subject headings: galaxies: formation — galaxies: structure

I. INTRODUCTION

Over the past decade, observations of nearby galaxies have revealed a number of underlying correlations and structural properties that still await a convincing explanation. Dissipation is generally considered to be a dominant process in the formation of disk galaxies, although there are varying views on its relevance to elliptical galaxy formation. It is generally argued that stellar dynamical interactions determine the structure of the spheroidal components of galaxies (cf. Freeman 1975). It seems clear to us, however, that dissipational processes must also be important in elliptical galaxy formation for reasons that we enumerate in detail below. Our aim here is to provide a unified theory of galaxy formation in which many of the observed properties of elliptical and disk galaxies can be understood if both dynamical and dissipative processes play a comparable role.

The high degree of central concentration of luminous matter indicates that, at least relative to the spatial distribution of dark halo material, dissipation must have occurred. Pioneering studies of dissipative galaxy formation have been performed by Larson (1974, 1976).

However, his models of elliptical galaxies fail to account for the flattening and low observed rotation of ellipticals. Dissipationless infall theories involving violent stellar relaxation (Gott 1975) meet with a similar difficulty.

Recent attempts to account for the low observed rotation of ellipticals have emphasized the role of galaxy mergers (White 1979a). The possible importance of mergers of spiral galaxies was suggested by Toomre (1977), although Silk (1977) has argued that protogalaxy mergers of predominantly gaseous systems are probably more relevant in accounting for the metallicities and luminosities of ellipticals. N -body simulations of galaxy clustering are found to lead to low impact parameter mergers and small induced tidal torques, satisfactorily accounting for the low rotation of elliptical galaxies (Aarseth and Fall 1980). However, galaxy mergers encounter notable objections in attempting to explain several other fundamental properties of ellipticals (cf. Ostriker 1980).

It seems clear that a hybrid approach to elliptical galaxy formation is necessary, in which both dissipation and violent relaxation play a role in preserving velocity anisotropy and accounting for various other features of

elliptical galaxies. The model that we shall pursue envisages a protogalaxy to consist of a collection of bound gas clouds that suffer inelastic collisions. Such a state may be attained by the collapse and fragmentation of a single large protogalactic cloud or by the hierarchical clustering and merging of many smaller primordial clouds and protogalaxies. Consideration of mergers and interactions of predominantly gaseous systems enables us to understand many properties of elliptical and disk galaxies. We commence in § II by summarizing the key observed properties of elliptical and disk galaxies. In the subsequent sections, the fundamental model of a protogalaxy is described (§ III), and we evaluate various characteristics of our protogalaxy model for elliptical galaxies (§ IV) and disk galaxies (§ V). In § VI, we discuss how our galaxy formation model is modified by gravitational clustering. A final section summarizes our results and proposes various observational tests.

II. OBSERVED PROPERTIES

We list in Table 1 nine important properties of elliptical galaxies that viable models must address, along with some principal references. The relation (E1) between luminosity and core velocity dispersion is well established. The correlation originally found by Fish (1964) between luminosity and galaxy radius ($L \propto R^2$) is listed

(E2) as a correlation with the de Vaucouleurs radius $L \propto R^{1.7}$ (Strom and Strom 1978) or $L \propto R^{0.7}$ (Kormendy 1980). The luminosity-metallicity relation (E3) has been calibrated by Sandage and Visvanathan (1978), based on magnitudes and colors for a wide range of elliptical galaxies. Important structural properties include metallicity gradients (E4) and the universal radial surface brightness profile (E5), described by either a Hubble or a de Vaucouleurs law. Structural properties that display a wide variation are the ratio of peak velocity to central velocity dispersion (E6), the degree of flattening (E7), and the dependence of ellipticity ϵ on excess metallicity relative to the luminosity-metallicity relation (E8). The limiting binding energy of luminous material (E9), as inferred from the central velocity dispersions of ellipticals, is also noteworthy.

Key parameters of galactic disks that models must incorporate are the angular momentum (S1), the exponential disk surface brightness distribution (S2), the disk metallicity gradients (S3), and the correlation between infrared luminosity and disk rotational velocity (S4). The bulges of disk galaxies deserve special mention, and we cite the bulge luminosity-velocity dispersion correlation (Bu 1), bulge luminosity-radius relation (Bu 2), peak rotational velocity (Bu 3), surface brightness profile (Bu 4), and degree of flattening (Bu 5). No theory of

TABLE 1

FUNDAMENTAL OBSERVED PROPERTIES OF GALAXIES

Type	Relation ^a	Reference ^b
Ellipticals:		
E1	$L = L_0(\sigma/\sigma_0)^4$	Faber and Jackson 1976; Sargent <i>et al.</i> 1977; Schechter and Gunn 1979
E2	$L \propto R^{1.7}$	Fish 1964; Strom and Strom 1978
E3	$L \propto R^{0.7}$	Kormendy 1980
E4	$I^{0.2-0.3} \propto Z$	Sandage and Visvanathan 1978; Tinsley 1978
E5	$\partial(Z/y)/\partial \ln r \approx -1$	Strom <i>et al.</i> 1976; Faber 1977
E6	$I \propto r^{-n} (1.7 \lesssim n \lesssim 2.1)$	Kormendy 1977
E7	$v_p/\sigma \approx 0.3 < (v_p/\sigma)_{\text{obl}}$	Illingworth 1977
E8	$\lambda \approx 0.06$	Schechter and Gunn 1979
E9	Flattening	Illingworth 1977; Schechter and Gunn 1979
E10	$\langle \epsilon \rangle \approx E6$	Sandage, Freeman, and Stokes 1970; Binney and de Vaucouleurs 1981
E11	$dZ/d\epsilon \lesssim 0$	van den Bergh 1979
E12	$\sigma \lesssim 400 \text{ km s}^{-1}$	
Spirals:		
S1	$\lambda \approx 0.03$	Bosma 1977
S2	$\mu = \mu_0 \exp(-ar)$	Freeman 1970
S3	$\partial(Z/y)/\partial \ln r \approx -1$	Shields and Searle 1978; Pagel 1979
S4	$L_{\text{ir}} \propto v_p^4$	Aaronson <i>et al.</i> 1979
Bulges:		
Bu 1	$L_B = L_0(\sigma_B/\sigma_0)^4$	Whitmore <i>et al.</i> 1979
Bu 2	$L_B \propto R^n$	Kormendy 1980
Bu 3	$v_p/\sigma \sim (v_p/\sigma)_{\text{obl}}$	Kormendy and Illingworth 1981
Bu 4	$I \propto r^{-n} (1.8 \lesssim n \lesssim 3.8)$	Kormendy 1980
Bu 5	Flattening: oblate, $\langle \epsilon \rangle \gtrsim E4$	de Vaucouleurs 1974
Bu 6	Bulge-to-disk ratio	Freeman 1970; Burstein 1979
Environment:		
	Fraction of SO, Sp, E	Melnick and Sargent 1977; Dressler 1980

^a Symbols are defined in text.

^b Cited references are not meant to be comprehensive or necessarily primary sources, but rather a guide to the recent literature.

galaxy formation is complete without consideration of variations along the Hubble sequence, where a key parameter is the variation of bulge size relative to the disk (Bu 6). Finally, the distribution of morphological types of galaxies depends on environment, and we specifically mention the quantitative estimates of S0 and E galaxy fractions in galaxy clusters (Dressler 1980; Melnick and Sargent 1977). Our aim will be to make a preliminary and tentative attempt to understand many, if not all, of these observed properties in terms of galaxy formation theory.

III. THE PROTOGALAXY MODEL

The idea that protogalaxies consist of a collection of interacting gas clouds is not new. It has formed a central role in models such as those proposed by Brosche (1970), Larson (1974), and Gott and Thuan (1976). Our aim here is to extend this approach into the more general context of the theory of the origin of galaxy structure and morphology.

There are many different evolutionary scenarios in which a gaseous protogalaxy can develop an inhomogeneous gas-rich structure. To demonstrate that clumpiness is a general feature of protogalaxy models, we consider two extreme viewpoints that constitute the principal alternative schemes of galaxy formation (cf. Gott 1977).

a) Hierarchical Clustering of Primordial Isothermal Density Fluctuations

Extrapolation of the spectrum of density fluctuations at recombination inferred from covariance function analyses and N -body simulations of galaxy clustering (Efstathiou 1979) indicates that, at recombination,

$$\delta\rho/\rho \propto (M/M_0)^{-\alpha}, \quad (1)$$

where $\alpha \approx \frac{1}{2}$ (if the density parameter $\Omega \approx 1$) or $\alpha \approx \frac{1}{3}$ (if $\Omega \approx 0.1$), and $M_0 \approx 10^{14.9-3/\alpha} \Omega^{1-1/\alpha} h^{-1}$ (with $h \equiv H_0/50 \text{ km}^{-1} \text{ s}^{-1} \text{ Mpc}^{-1}$). It follows that $M_0 \approx 10^8 M_\odot$ ($\Omega \approx 0.1$)– $10^9 M_\odot$ ($\Omega = 1$). A plausible inference is that a primordial spectrum of isothermal fluctuations in the early universe could have led to the formation of bound clouds of mass M_0 at $z \sim 1000$. White and Rees (1978) have assumed that these clouds immediately fragmented into dark remnants that subsequently clustered dissipationlessly, forming the potential wells out of which galaxies eventually developed at a much later epoch.

A different approach is taken by Peebles and Dicke (1968), who argue that Jeans mass clouds at this epoch (of mass $10^{5.8} \Omega^{-1/2} M_\odot$ and radius $\sim 8.9 \Omega^{-1/2} \text{ pc}$) survived as gravitationally bound gas clouds until galaxies formed, identifying the surviving remnants of a small fraction of these clouds with the globular clusters. In the Peebles-Dicke theory, protogalaxies formed from the clustering and disruption of the globular cluster-sized gas clouds. These clouds must have avoided cooling and gravitational collapse, and survived for a period $\gtrsim 10^8 \text{ yr}$ by means of internal energy generation from ongoing nondisruptive cycles of star formation. We consider below the plausibility of the key assumption that the gas clouds can survive as bound entities.

b) Fragmentation of Primordial Adiabatic Fluctuations

An alternative scenario for galaxy formation utilizes primordial adiabatic fluctuations. Surviving adiabatic fluctuations have mass in excess of $\sim 10^{13} M_\odot$. Such massive gas clouds must recollapse at a relatively late epoch. General considerations suggest that the collapse will be highly anisotropic and result in thermal instability and extensive fragmentation. Characteristic fragment masses are estimated to be 10^6 – $10^9 M_\odot$ (Doroshkevich, Shandarin, and Saar 1978; Sunyaev and Zel'dovich 1972). Provided that these gaseous fragments are stabilized by internal star formation, the subsequent evolution of a galaxy involves the inelastic interactions of such gaseous subunits. Thus our basic protogalaxy model emerges either in the hierarchical or in the fragmentation theories.

c) Survival

The survival of gas clouds for time scales $\gtrsim 10^9 \text{ yr}$ (long compared to their collapse time scales) is an essential requirement of our model for dissipative galaxy formation. That such a phenomenon, namely the survival of gas-rich clouds over long time-scales, occurs may be inferred from the existence of numerous gas-rich dwarf galaxies (Silk and Norman 1979), which appear to have only undergone active star formation within the past 10^9 years. Within our own Galaxy, it is clear that many molecular clouds ($\gtrsim 10^5 M_\odot$) may have lifetimes $\gtrsim 10^8$ years, significantly longer than their collapse times. This can be understood if ongoing star formation by low mass stars provides an internal source of energy (Norman and Silk 1980a). We now describe in more detail how primordial gas clouds may survive for 10^8 – 10^{10} yr .

In the absence of heavy elements, such clouds will radiatively cool to $T \sim 10^4 \text{ K}$, and we may express the characteristic radius and density of a self-gravitating gas cloud as

$$R = 10^{1.82} M_6 T_4^{-1} \text{ pc}$$

and

$$n = 10^{1.37} T_4^3 M_6^{-2} \text{ cm}^{-3}, \quad (2)$$

where the virial temperature $T_4 \equiv T/10^4 \text{ K}$ and the cloud mass is $M_6 \equiv M/10^6 M_\odot$. The virial temperature can exceed the gas temperature because of turbulent energy input by stellar winds and supernovae. It is necessary to show that a primordial cloud will not be disrupted by such processes. Once we establish this, we shall then argue that the energy input associated with star formation and evolution adjusts itself to maintain the cloud in appropriate virial equilibrium, until the gas supply is exhausted and the ongoing star formation rate abruptly decreases.

The primary mode of energy input is likely to be via H II regions and supernovae, since massive star formation is presumed to occur in pregalactic enriched gas clouds (cf. Silk 1977). To illustrate our argument, we shall consider the role of supernova remnants, believed to be the primary source of energy input into the local interstellar medium, in heating and disrupting the cloud. Other modes of stellar energy input, including stellar winds as

well as H II regions, may also be important, but are unlikely to affect the gross energetics.

A supernova remnant will attain a maximum radius determined by pressure balance with the ambient cloud gas, given by

$$R_{\text{SNR}} = 10^{1.25} E_{51}^{0.32} M_6^{0.72} T_4^{-1.28} \text{ pc}, \quad (3)$$

where the energy in the initial explosion $E_{51} \equiv E/10^{51}$ ergs (cf. Chevalier 1974). For the cloud to survive supernova explosions, two conditions are required. First, the cloud radius must exceed R_{SNR} , where

$$M_6 > 10^{-2.07} E_{51}^{1.14} T_4^{-1}. \quad (4)$$

Second, we require that successive supernovae do not result in the intersection of supernova remnants prior to shell breakup and mixing of the hot interior with ambient cool gas. Otherwise, the coherent interaction of supernova remnants will drive a wind and disrupt the cloud. Since shell breakup occurs at the limiting remnant radius R_{SNR} , this constraint implies that the mean time t_{SN} between supernovae must satisfy

$$t_{\text{SN}} > t_{\text{crit}} \equiv R_{\text{SNR}}^4 (R^3 v_s)^{-1} \\ = 10^{4.62} E_{51}^{1.28} M_6^{-0.12} T_4^{-2.52} \text{ yr}. \quad (5)$$

Here, v_s is the sound velocity in the cloud.

To relate this restriction on t_{SN} to the cloud lifetime, we assume that for each supernova, a mass M_* of stars forms. In the local interstellar medium, the observed star formation rate ($\sim 3 M_\odot \text{ yr}^{-1}$) and mean supernova rate ($\sim 0.03 \text{ yr}^{-1}$) imply that $M_* \sim 10 M_\odot$. However, the mass-to-luminosity ratio of our Galaxy (interior to the Sun) is ~ 10 ; if substantial amounts of low mass stars form in the primordial gas clouds, M_* could be as large as $100 M_\odot$ (since the mass-to-luminosity ratio of the entire galaxy may be ~ 100). Despite this uncertainty in M_* , we can express, in order of magnitude, the cloud lifetime against star formation (neglecting recycling of gas by stars, an effect which will increase the cloud lifetime) as $(M/M_*)t_{\text{SN}}$. Since $t_{\text{SN}} > t_{\text{crit}}$ according to equation (5), the lower limit on t_{SN} yields the constraint that $t_{\text{cl}} \equiv (M/M_*)t_{\text{SN}} > (M/M_*)t_{\text{crit}}$, whence

$$M_6 < 10^{0.43} E_{51}^{-1.45} T_4^{2.86} \\ \times [(t_{\text{cl}}/10^9 \text{ yr})(M_*/100 M_\odot)]^{1.14}, \quad (6)$$

where t_{cl} is the cloud lifetime. For a specified cloud lifetime, more massive clouds will be disrupted by the effects of a high star formation and supernova rate. Comparing equations (4) and (6), we infer that clouds in the mass range

$$10^{-2.07} E_{51}^{1.14} T_4^{-1} < M_6 < 10^{0.43} E_{51}^{-1.45} T_4^{2.86} \\ \times [(t_{\text{cl}}/10^9 \text{ yr})(M_*/100 M_\odot)]^{1.14} \quad (7)$$

can have lifetimes of $\sim 10^9$ yr, well in excess of the collapse time of a protogalaxy.

Are such lifetimes actually attainable?

The following argument seems plausible, although it is far from rigorous. Suppose that when a cloud initially

cools and contracts, the rate of massive star formation is sufficient to provide sufficient energy input to reverse the collapse. This will in turn lead to a reduction in the star formation rate, which can evidently adjust itself to maintain the cloud in a more or less quasi-equilibrium state, until the gas supply is exhausted. The obvious weakness in this reasoning is that the initial burst of star formation could be so catastrophic as to blow the cloud apart. For disruption by supernovae, this requires condition (5) to be violated. Using a conventional supernova rate $\sim 0.03 \text{ yr}^{-1} (10^9 M_\odot)^{-1}$ for the present interstellar medium, we infer that $t_{\text{SN}} \sim 3 \times 10^4 M_6^{-1} \text{ yr}$, comparable with t_{crit} in equation (5). Hence only if the efficiency of massive star formation is much greater than at present will the cloud be disrupted. We assume that this is not the case.

Since we expect the clouds to initially be at $T_4 \sim 1$, it indeed seems likely that a gas cloud of mass 10^6 – $10^7 M_\odot$ can avoid disruption by supernova energy input and star formation for 10^9 – 10^{10} yr. However, in the absence of sufficient heat input, a cloud will cool and collapse. For a cloud to be maintained in virial equilibrium, we must evidently require the supernova rate to occur at the critical rate $t_{\text{SN}} = t_{\text{crit}}$. Thus the actual cloud lifetime must be $(M/M_*)t_{\text{crit}}$, or

$$t_{\text{cl}} = 10^{8.62} M_6^{0.88} T_4^{-1.52} E_{51}^{1.28} (100 M_\odot/M_*) \text{ yr}. \quad (8)$$

Hence gas clouds of mass 10^6 – $10^7 M_\odot$ can survive for times long compared to their collapse time scales ($\sim 10^{6.9} M_6 T_4^{-3/2} \text{ yr}$). In an isolated environment, it seems not implausible that such clouds could exist, continually forming stars, for up to $\sim 10^{10}$ yr.

d) Nondissipational Cloud Clustering

N-body simulations of clustering have demonstrated that substructure is ultimately erased as larger structures form (White 1976; Aarseth, Gott, and Turner 1979). However, it is important to realise that the basic units from which all structures are built, namely the first generation of bound systems to form after decoupling, need not necessarily be destroyed as a result of the merging. This is because the first systems may have radiated a considerable amount of internal energy and thereby increased their binding energy.

The criterion for survival of bound gas clouds to cosmological epoch t can be expressed as the requirement that cloud collisions do not occur, or $t_{\text{coll}} = (N_{\text{cl}} \sigma_{\text{cl}} V_{\text{cl}})^{-1} \gtrsim t$, where N_{cl} , σ_{cl} , and V_{cl} are the cloud number density, cross section, and mean random velocity at time t . Using equation (1) with $\alpha \equiv (n+3)/6$ and $\delta\rho/\rho \propto t^{2/3}$, this condition can be written

$$\frac{t_{\text{coll}}}{t} > \left(\frac{M_0}{\rho_d \sigma_{\text{cl}} v_0 t_d} \right) \left(\frac{t}{t_d} \right)^{4(n+2)/3(n+3)}, \quad (9)$$

where σ_{cl} is assumed fixed and ρ_d is the density at the decoupling epoch t_d . We have assumed that the random velocity of a cloud is comparable to the virialized velocity in a newly formed aggregate of clouds of approximately unit density contrast relative to the expanding back-

ground. At decoupling, the internal velocity dispersion of a cloud is v_0 . Since the observed limits on α indicate that $n > -2$, the exponent in equation (9) is undoubtedly positive. Condition (9) will therefore be satisfied at all times, provided it is satisfied initially at t_d , that is, provided $M_0(\rho_d \sigma_d v_0 t_d)^{-1} > 1$. This condition is equivalent to requiring that the newly formed clouds acquire a moderate degree of central concentration: dissipation ensures that this should occur. We conclude that, in hierarchical clustering, cloud collisions will not occur while the emerging structures remain self-similar.

A simple cloud clustering model must also apply in the fragmentation model of galaxy formation. Direct evidence for the universality of gravitational clustering is inferred both from studies of the galaxy correlation functions and from the calculations and N -body simulations that have successfully reproduced them (Peebles 1974; Davis and Peebles 1977; Gott, Turner, and Aarseth 1979). According to Doroshkevich, Shandarin, and Saar (1978), gaseous fragmentation of massive gas clouds results in the formation of small clouds at a late epoch ($z \lesssim 10$), subsequently forming galaxies by hierarchical clustering. In this scenario, protogalaxies are largely gaseous at the epoch of cluster formation, and individually consist of many smaller clouds.

Evidently both the fragmentation and clustering models may result in protogalaxies that remain largely gaseous at the epoch of cluster formation. We emphasize again that the observed properties of galaxies apparently require such a gas-rich protogalactic phase when galaxy interactions are occurring. The fact that this situation arises "naturally" in the fragmentation model is a favorable aspect of this theory.

In considering the details of cloud interactions, we may distinguish two possible regimes, depending on whether or not the interactions involve strong dissipation of kinetic energy of bulk motion. The dark material that contributes the bulk of the matter in galaxy halos and galaxy clusters will presumably cluster dissipationlessly. We shall argue that the gas clouds that form protogalaxies must interact dissipatively. Dissipation leads to the formation of the condensed luminous regions of galaxies. However, the dark material, or the precursor matter from which it originates, must have been dynamically significant when galaxy clusters and halos formed.

According to one recent scheme for galaxy formation (White and Rees 1978), the dark matter formed at an early epoch and clustered gravitationally, developing the potential wells into which gas infall subsequently occurred followed by dissipative galaxy formation. If we commence with basic units of mass M_i at epoch t_i , the growth of density fluctuations on larger scales can be described by a spectrum of the form

$$\frac{\delta\rho}{\rho} = \left(\frac{M}{M_i}\right)^{-1/2-n/6} \left(\frac{t}{t_i}\right)^{2/3}, \quad (10)$$

where subscript i refers to the initial epoch at which clustering is first described by a growth rate proportional

to $t^{2/3}$. When $\delta\rho/\rho$ exceeds unity on a given scale, that corresponding mass scale freezes out as a gravitationally bound unit (Press and Schechter 1974). In an Einstein-de Sitter universe ($\Omega = 1$), the amount of matter M that clusters on scale r without any dissipation is accordingly given by

$$M/M_i \sim (R/R_i)^{6/(n+5)}. \quad (11)$$

Here $M_i = \frac{4}{3}\pi\rho_i R_i^3$, ρ_i is the mean cosmological density at epoch t_i , and the parameter n and the initial clustering scale R_i when density fluctuations first recollapse at epoch t_i can in principle be inferred from the observed galaxy correlation function $\xi(R)$. For example, adopting

$$\xi(R) = (R/R_0)^{-1.8}, \quad R_0 = 8h^{-1} \text{ Mpc}, \quad (12)$$

we infer from equation (11) that

$$n = 0, \quad M_i \approx 5 \times 10^{14} (1 + z_i)^{-2} M_\odot,$$

and

$$R_i \approx R_0 (1 + z_i)^{-5/3}.$$

This result is applicable to the hierarchical theory if we identify z_i with the recombination epoch, and it applies equally to the fragmentation theory if we identify z_i with the epoch of galaxy clustering (i.e., at a redshift $\lesssim 10$). Observational evidence suggests that the dark matter satisfies $M \propto R$ (Ostriker, Peebles, and Yahil 1974; Einasto, Kaasik, and Saar 1974). This is consistent with $n \approx 0$ as inferred from the N -body simulations in an Einstein-de Sitter universe, although inconsistent with the value $n \approx -1$ inferred in a low density cosmological model (Gott, Turner, and Aarseth 1979).

e) Dissipational Clustering

In fact, the self-similar nature of hierarchical clustering changes once a system containing many clouds has become gravitationally bound. As indicated by the N -body simulations, a core-halo structure develops in the cloud distribution, resembling a Hubble law (White 1979a; Roos 1981). The density contrast between core and outer halo may be crudely estimated by the ratio of the de Vaucouleurs to King core radii, which is typically 10 to 1 (Kormendy 1977). Consequently, the cloud collision time has a value in the halo of about 10^3 times that in the core. Since cloud collisions are likely to trigger star formation, the core-halo structure should result in the formation of a stellar core, surrounded predominantly by a cloudy gaseous halo.

The dimension of the core region depends on the dynamical history of the protogalaxy. If it evolves in isolation, dissipative collisions and associated star formation will be spread out over a relatively long period of time, $\sim 10^9$ – 10^{10} yr (§ V). In denser environments, mergers between neighboring protogalaxies will occur. Violent relaxation associated with mergers will inject many gas clouds into the core of the resulting galaxy. These will collide frequently over t_{coll} and form stars. The merging of the protogalaxies can sustain the injection process for a few crossing times until merging is complete. Since cloud collisions will rapidly convert the newly injected clouds

into stars, we can identify the core radius of the merged galaxy by the condition $t_{\text{coll}} \sim t_{\text{dyn}}$. This condition plays an important role in determining the resulting core properties, as we now describe.

IV. FORMATION OF SPHEROIDAL SYSTEMS

We consider dissipative cloud interactions, which, as we shall argue, determine the structure of the cores of galaxies. It is immaterial for the present discussion whether the dark matter is produced prior to, or as a consequence of, the cloud interactions.

a) The Maximum Binding Energy of Galaxies

Colliding bound gas clouds will merge provided that the shock cooling length-scale is less than the size of the final virialized system. Consider two clouds of mass M_0 and internal velocity dispersion σ_0 ($\sim 10 \text{ km s}^{-1}$) colliding at a relative velocity appropriate to that of a larger virialized system with velocity dispersion σ_g . The condition that a merger occurs can be expressed as an upper bound on σ_g . If σ_g is too great, the collision will be disruptive. We require that the shock width, expressed as the cooling length-scale, be less than the cloud radius. Making use of equation (2), we infer that

$$\sigma_g \leq 600(f/M_0)^{1/2}(\sigma_0/10 \text{ km s}^{-1})^2 \text{ km s}^{-1}, \quad (13)$$

where $f(\sigma, Z) (\geq 1)$ is the ratio of the actual cooling rate to the free-free cooling rate in the shocked cloud material.

This gas-dynamical constraint may provide a clue to understanding the origin of the characteristic luminosity L_* in the luminosity function of galaxies (Schechter 1976): specifically, it suggests that galaxies have a limiting binding energy determined by the disruption of the gaseous subunits out of which they have formed. More tightly bound systems would undergo less star formation and only attain low surface brightness: possible examples would be the extended envelopes found in clusters of galaxies often dominated by a cD galaxy. It has previously been noted that simple considerations of gas-dynamical processes in the fragmentation of self-gravitating gas clouds lead to characteristic masses and radii for the surviving protogalaxies (Rees and Ostriker 1977; Silk 1977). We have shown in equation (13) that a third parameter, the maximum binding energy for the luminous matter in galaxies, can be expressed in terms of the initial Jeans length of the basic gaseous subunits. That galaxies indeed possess a limiting binding energy may be inferred from the luminosity dependence of σ_g for elliptical galaxies ($\sigma_g \propto L^{1/4}$) and the approximately exponential cutoff in the luminosity function at $L > L_*$.

The preceding argument only delineates an upper limit to the attainable binding energy: mergers may cease before this limit is actually attained. We shall argue below that it is the degree of interaction of a protogalaxy with its environment that determines the supply of gaseous subunits during its early evolution.

b) Luminosity Profile

The stellar component in clouds will be removed during a collision. Moreover, the collision-induced shock

may trigger further star formation. Thus the radiative energy losses associated with the collision result in a loss of orbital energy of the clouds, which therefore form a centrally condensed core. Two independent calculations suggest that the resulting profile follows a Hubble law.

i) The clouds can be regarded as particles in an N -body simulation, and they violently relax where collisions are unimportant. The merged system is characterized by the development of a core-halo structure, and only in the core do cloud collisions play a significant role. Outside the core, where collisions and coagulation are unimportant, the collisionless Vlasov equation describes the evolution of the system by

$$\frac{\partial N}{\partial t} + \mathbf{v} \cdot \nabla N + \nabla \phi \cdot \frac{\partial N}{\partial \mathbf{v}} = 0. \quad (14)$$

The N -body simulations of merging galaxies will provide an appropriate model over time scales up to approximately t_{coll} . Merging, for example, takes several crossing times; and if $t_{\text{coll}} \gtrsim 3t_{\text{dyn}}$, relaxation will have occurred before any collisional effects take place. Consequently, the resulting radial distribution should be approximately a Hubble law in mass density, as found in the numerical N -body simulations (Roos 1981; White 1979a). We envisage that this violently relaxed state is reached outside the cores of merged systems.

ii) Another approach is to model the dissipative effect of collisions in a global hydrodynamic approximation. Larson's (1974) early models of elliptical galaxy formation were characterized by this approach, and he demonstrated that, with suitable assumptions about the density dependence of the star formation rate and the dissipation rate, a Hubble law could be derived. We have developed an alternative hydrodynamical description for cloud interactions based on a coagulating self-gravitating system of gas clouds. We show in the Appendix that, provided a small but fixed fraction ϵ of the orbital kinetic energy is lost per collision, a simple power law solution for the stationary mass density in clouds $N(r)$ is inferred:

$$N(r) \propto \{1 + (r/r_1)^7\}^{1/[2(\epsilon+1)]} \propto r^{-3.5} \quad \text{for } M_{\text{dyn}}(r) = \text{constant} \quad (15)$$

and

$$N(r) \propto \{1 + (r/r_2)^6\}^{1/[2(\epsilon+1)]} \propto r^{-3} \quad \text{for } M_{\text{dyn}}(r) \propto r, \quad (16)$$

where r_1 and r_2 are constants, we take ϵ to be small compared to unity, and the effective dynamical mass in the protogalaxy is denoted by M_{dyn} . Evidently, if M/L is approximately constant for stars formed in the interacting clouds, a Hubble luminosity profile is again obtained.

c) The M - σ - R Relations

In any self-gravitating system we can express the total mass interior to radius r as

$$M(r) \sim 3\sigma^4(4\pi G^2 \rho r)^{-1}, \quad (17)$$

where ρ is the mean density interior to r , and σ is the central velocity dispersion. Suppose that the entire system consists of an ensemble of identical clouds. If μ is the characteristic surface density of the interacting clouds, we can express the mean cloud-cloud collision time as $(n_{cl} \pi R^2 2^{1/2} \sigma)^{-1} = 2^{1/2} \mu (\rho \sigma)^{-1}$, where n_{cl} is the cloud number density at radius r and R is the cloud radius. Since the dynamical time scale of the system is r/σ , we can write

$$M(r) = \frac{3}{2\pi} \frac{\sigma^4}{G^2 \mu} \left[\frac{t_{coll}(r)}{t_{dyn}(r)} \right]. \quad (18)$$

We now apply the conclusion that we drew from our discussion of mergers in § III, namely that the core mass is determined by the condition that the cloud collision time be of the order of the dynamical time-scale, since core growth can only be achieved by injection of clouds during a merger (or in a similarly catastrophic process over a dynamical time scale), whence

$$M_{core} \sim \sigma^4 (G^2 \mu)^{-1}. \quad (19)$$

This yields an estimate of the core mass, since the core is effectively truncated at a radius where cloud collisions become ineffective. This truncation hypothesis is the key point in our model: it delineates the extent of the luminous bulges, and leads to an understanding of the core parameters. The basic physical reasoning is that star formation is enhanced in the cloud collision-dominated regime. Obviously, dissipation and infall can also occur over a much longer time scale: we argue below that this occurs over a viscous (and relatively slow) time scale. Since the interacting clouds have been injected from a region where few collisions have occurred, we may consider them to have retained the characteristic parameters with which they were formed. Specifically, in the hierarchical clustering model, we infer for Jeans mass clouds produced at decoupling that $\mu \sim v_s/Gt \sim 10^4 \Omega^{1/2} M_\odot \text{pc}^{-2}$. A spheroidal galaxy will form provided that the clouds are efficiently converted into stars in the galaxy core with $M/L \sim 10$. If the galaxy luminosity is proportional to the core mass, relation (19) accounts for the observed L - σ relation (E1).

A consequence of the above model is that Fish's law (1964) is inferred, since we have

$$\frac{M}{R^2} = \mu \left[\frac{t_{dyn}(R)}{t_{coll}(R)} \right]. \quad (20)$$

If we identify R with the de Vaucouleurs radius and assume that M/L has a universal value (for the matter within R), we conclude that the matter which has undergone cloud collisions and dissipation must satisfy $L \propto R^2$. The normalization is consistent with observed values for the core-averaged surface luminosity in M87 and the bulge of M31 provided that $\mu \sim (3 \times 10^3 - 10^4) (M/L) M_\odot \text{pc}^{-2}$ (cf. Schweizer 1979). While Fish's law ($L \propto R^2$) has been essentially confirmed by Strom and Strom (1978), who find $L \propto R^{1.7}$, a discordant relationship ($L \propto R^{0.7}$) has been reported by Kormendy (1980). Until the observational situation is resolved, it may be

premature to place much significance in the preceding discussion. While relation (19) is consistent with the Stroms' data if $\mu \approx \text{constant}$, it obviously could be reconciled with Kormendy's luminosity-radius relation by introducing a new parameter, namely a dependence of μ on M_{core} . It is only worth noting for the present that the $L \propto \sigma^4$ relation can be reconciled with an arbitrary (L, R) relation by means of this added degree of complexity: equations (18) and (20) yield

$$L/R^2 \approx (t_{dyn}/t_{coll})^2 \approx \mu^2 G^2 (L/\sigma^4).$$

From the viewpoint of galaxy formation theory, a systematic variation in μ could undoubtedly be explained. However, such a resolution would certainly detract from the appeal of the present model.

A weakness in the preceding argument is that it omits any consideration of the galaxy mass function. A single unit of some characteristic mean initial mass can make either a few very luminous ellipticals or very many low luminosity ellipticals. Fragmentation must be incorporated into the model, so that if a low- σ system forms first, the rest of the matter in the initial unit out of which it formed does not stay gaseous, but forms additional galaxies. Even with fragmentation and mergers playing a role, the truncation hypothesis is crucial to understanding the (L, σ, R) relations. Successive mergers will lead to formation of higher σ systems and consequently systems of increasing luminosity.

d) Angular Momentum

White (1979a, b) has shown that mergers between initially unbound stellar systems lead to elliptical galaxies that rotate too rapidly to be consistent with the observed rotation of ellipticals. If angular momentum is conserved in the merger with impact parameter p_∞ and velocity V_∞ at infinity, then an estimate of the ratio of peak rotational velocity V_p to core velocity dispersion σ is

$$\frac{V_p}{\sigma} \sim \frac{1}{4} \left(\frac{p_\infty}{V_c} \right) \left(\frac{V_\infty}{\sigma} \right), \quad (21)$$

where, according to White (1979b), simulations give $V_\infty/\sigma \sim 1$ and $p_\infty \approx 2r_c$, yielding $V_p/\sigma \sim 0.5$. Resolution of this difficulty may have been achieved by numerical simulations of galaxy mergers in an expanding universe (Aarseth and Fall 1980), which show that most mergers occur between marginally bound two-body orbits of low angular momentum. The inferred mean value of V_p/σ is 0.23, with a flat distribution ($0 \lesssim V_p/\sigma \lesssim 0.5$).

These expanding universe simulations represent galaxies by point masses. Computational limitations currently prevent more realistic simulations, but we can speculate that the effect of an extended halo will result in mergers occurring at larger impact parameters $p_\infty \lesssim 2r_c$. Furthermore, the cores of the merging components can orbitally decay in the flat gravitational potential if a dark halo is present. Provided that the cores are initially orbiting in the background potential of the dark halo material and eventually merge into a self-gravitating,

centrally condensed core, the resulting value of V_p/σ will be enhanced over what it would have been in the absence of the halo. It seems plausible that the resulting value of V_p/σ will be correlated with the degree of coherence of the dark halo material.

e) *Flattening and Velocity Anisotropy*

Dynamical studies of elliptical galaxies (Illingworth 1977; Schechter and Gunn 1979) indicate that the observed flattening is not generally attributable to rotation, but must rather be interpreted in terms of an anisotropic velocity distribution. We wish to describe a simple mechanism that may lead to such a distribution as a consequence of protogalaxy mergers. During the actual merger, the orbital momentum loss is primarily in the orbital plane, according to the impulsive approximation. This momentum transfer will lead to preferential heating in the orbital plane. A similar concept has been explored by Sanders and van Albada (1979) in their investigation of bar dissolution, resulting in anisotropic velocity distributions.

For a crude estimate of this effect, suppose that all of the internal energy produced in the merger goes into increasing the velocity dispersion of particles in the orbital plane. In this case, the energy increase is estimated for a merger of two galaxies with masses m_1 and m_2 and the same core radius, according to the impulsive approximation model of Marchant and Shapiro (1977):

$$\frac{\Delta E}{E} \sim \left(\frac{\sigma}{V_\infty}\right)^2 \left(\frac{M_1}{M_2}\right)^2 \left[\frac{7}{1 + 9(p_\infty/r_c)^4} \right]. \quad (22)$$

The anisotropy parameter Q , defined by Binney (1978) as $Q \equiv \Delta E(M_{\text{dyn}} V_{\text{rot}}^2)^{-1}$, can then be written as

$$Q = 2 \left(\frac{\sigma^4}{V_\infty^2 V_p^2} \right) \left(\frac{M_1}{M_2} \right)^2 \left[\frac{7}{1 + 9(p_\infty/r_c)^4} \right]. \quad (23)$$

We know that to fit the observed data on elliptical rotations, we need Q to be of order unity. A good fit is actually found (Binney 1978) for oblate spheroids with $Q \sim 2$. This will occur for impact velocity $V_\infty \sim \sigma$, $M_1 \sim M_2/3$, and small impact parameters $p_\infty \lesssim r_c$. A more precise estimate awaits $N \gtrsim 10^4$ N -body simulations to determine the detailed structure of the merged systems.

The above arguments are strictly valid only in the absence of collisions. However, we note that when cloud-cloud collisions occur, the stellar component will be ejected from the clouds while the clouds will undergo dissipation. Further star formation will occur, these stars themselves being ejected in subsequent collisions. Therefore we expect that the anisotropy in the ejected stellar component will be preserved for a time on the order of a few collision times.

f) *Color-Magnitude Relation*

We have argued that the violent relaxation of gas clouds during a merger will result in the injection of some clouds into the core of the merged galaxy, where these clouds will rapidly undergo collisions. The resulting star

formation has two effects. It will increase the central surface brightness of the merged system relative to the mean mass distribution. Moreover, if the efficiency of star formation increases with the relative velocity of the colliding clouds, more massive galaxies will acquire higher degrees of enrichment. One can understand such a velocity-dependent efficiency of star formation in cloud-cloud collisions because the mass of compressed gas will increase with shock Mach number, provided that the shock is radiative. This effect can lead to a quantitative relation between color and magnitude, as first suggested by Silk (1977) and subsequently explored in more detail by Tinsley and Larson (1979).

To illustrate this, assume that the star formation efficiency $\eta \propto \sigma^n$, where σ is the internal velocity dispersion of the merged galaxy, mass M_g . Using the relation $L \propto \sigma^4$, we have $\eta \propto L^{n/4}$, and following the approach of Tinsley and Larson (1979) we obtain

$$Z \propto M_g^{n/(n+4)} \propto \sigma^{4n/(n+4)} (M_g/L)^{n/(n+4)}.$$

Observations are consistent with $n \sim 1$ and $Z \propto M_g^{0.2}$; consequently, we infer that $Z \propto \sigma^{0.8} (M_g/L)^{0.2}$. Thus multiple mergers will result in enhanced metallicity being correlated with luminosity: the actual slope of the correlation constrains the (otherwise arbitrary) star formation efficiency. One notable implication is that, because mergers will in general occur with random orientation, multiple mergers result in rounder systems. However, the broad dispersion in luminosities means that one should not expect to find any flattening-luminosity correlation. These rounder systems are systematically more metal rich, as the star formation efficiency increases because of the increased merger rate. Thus the systems with enhanced metallicity are also expected to be of lower eccentricity. This is only a second order effect, the principal result being that mergers promote metallicity, since mergers between unequal systems will tend to produce a final system with the shape of the more massive parent. However, it is of interest to note that observations of elliptical galaxies may provide evidence for such a trend (van den Bergh 1979).

g) *Metallicity Gradients*

As will be demonstrated in § V, dissipative cloud-cloud collisions lead to a radial infall of material relative to the stellar component that is ejected from the colliding clouds. This naturally gives rise to a metallicity gradient, which can be written (Tinsley and Larson 1978) as

$$\frac{\partial(Z/y)}{\partial \ln r} \sim \frac{3\dot{M}_*(r)}{\dot{M}_i(r)}, \quad (24)$$

where y is the fractional yield in heavy elements per generation of star formation, $\dot{M}_*(r)$ is the rate of star formation at radius r , and $\dot{M}_i(r)$ is the mass inflow rate at radius r . Therefore, to obtain a typical observed gradient where $\partial(Z/y)/\partial \ln r \sim 1$ (Faber 1977), we require $\dot{M}_* \sim \frac{1}{3}\dot{M}_i$. We note that if the stars remained in the clouds and enrichment was purely internal to the clouds, which evolved as isolated systems, no gradient would

develop by this mechanism. It is necessary that the clouds fall radially relative to the ejected stars: dissipation of orbital kinetic energy provides a natural mechanism for ensuring that this happens. The key assumption (justified previously) is that individual clouds can survive several collisions without being totally disrupted by internal star formation or by collision-induced shocks.

V. FORMATION OF DISK SYSTEMS

There is strong observational evidence that luminous spiral galaxies possess massive halos with $M \propto R$ extending out to at least two Holmberg radii. The halo is therefore the dominant mass constituent of a typical spiral galaxy. We have argued that one implication is that mergers between protogalaxies will tend to result in the formation of galaxies with rotationally supported cores. We now consider how a disk may form. It is clear that mergers will tend to disrupt preexisting disks, the implications of which we explore in the following section. However, an isolated protogalaxy will inevitably develop a disk as cloud-cloud collisions systematically dissipate energy. We address two questions below: (1) What is the angular momentum of the disk? (2) What is its structure?

a) Disk Angular Momentum

A convenient measure of the angular momentum of a galaxy with angular momentum J , mass M , and total energy E is given by the dimensionless parameter $\lambda \equiv JE^{1/2}G^{-1}M^{-5/2}$. For the disk of our Galaxy, $\lambda \approx 0.4$, whereas for elliptical galaxies $\lambda \approx 0.06$. According to the gravitational instability model of galaxy formation by hierarchical clustering, rotation is induced by tidal interactions between merging protogalaxies. Numerical simulations of this process (Efstathiou and Jones 1979) yield a median value of $\lambda = 0.07 \pm 0.02$, in general agreement with the observed properties of elliptical galaxies. However, this low value of λ leads to a difficulty in understanding the formation of spiral disks, since a self-gravitating centrifugally supported disk, even in the absence of viscous decay, would necessarily lead to collapse by an unacceptably large factor in radius (Efstathiou and Jones 1980; Fall and Efstathiou 1980). However, as noted by these authors, the presence of a massive halo drastically modifies this argument. The gas clouds that develop into the disk and form stars are not initially a self-gravitating system. In the absence of viscous effects, the final radius of the self-gravitating disk (mass M_d) is of order $\lambda^2(M_h/M_d)R_h$, where M_h and R_h are the halo mass and radius, indicating that collapse by a factor $\sim \lambda^{-1}$ is required in a preexisting isothermal halo.

b) Disk Structure

The preceding estimate of the disk radius is modified by the viscous drag associated with cloud-cloud collisions. The viscosity will tend to increase the degree of central concentration as angular momentum is transferred outward and further radial inflow occurs. A characteristic radius r_0 at which viscous effects become dominant can be obtained by setting the cooling time scale $t_c \sim l_0(\bar{v}\eta)^{-1}$ due to energy loss via cloud-cloud collisions equal to the

viscous time scale $t_v \sim r^2/\nu$, where $\nu \sim \frac{1}{3}l_0\bar{v}$, η is the fraction of the cloud kinetic energy radiated in a collision, $\bar{v}$ is the cloud random velocity, and l_0 is the cloud mean free path at some initial time. This condition yields $r_0 \sim \eta^{-1/2}l_0 < R_d$, and we may identify r_0 with the centrally condensed region of the disk.

We now examine in more detail the structure of the disk. In fact, the system will evolve with time as energy is lost and angular momentum is transferred outward in such a way that the cloud mean free path must vary with position. The physical situation is analogous to that encountered in accretion disks, and maintenance of viscous inflow requires that the angular momentum transfer and cooling rates must be comparable. Therefore we assume that the viscosity adjusts itself such that $t_c \sim t_v$ at any radius r , and $\nu = v_0(r/r_0)$. For a flat rotation curve determined by the halo potential, this viscous accretion disk will evolve in a self-similar way (Lynden Bell and Pringle 1974) to give a surface density law

$$\Sigma = \Sigma_0 \left(\frac{r_0}{r} \right) \left(\frac{r_0^2}{v_0 t} \right)^2 \exp \left(- \frac{rr_0}{v_0 t} \right). \quad (25)$$

A similar exponential disk law is also found for a Keplerian rotation curve. It is noteworthy that viscous disk calculations for a flat rotation curve naturally lead to an exponential disk mass distribution, together with an inner bulge that is intermediate in radial profile between an isothermal disk and a Hubble law density gradient. The actual spheroidal components of disk systems are presumed to collapse on a time scale $\lesssim t_c$; hence the viscous disk solution suggests that the bulge may continue to grow, accreting low angular momentum material from the disk.

The characteristic scale of these disks is determined by the star formation rate, since viscosity cannot act on the stellar component. Once the disk becomes predominantly stellar, its structure is subsequently frozen. Thus the characteristic scale size α^{-1} is

$$\alpha^{-1} \approx vt_* r_0^{-1} \\ \approx 3(\bar{v}/30 \text{ km s}^{-1})(\eta/0.1)^{1/2}(t_*/10^9 \text{ yr}) \text{ kpc}, \quad (26)$$

and a wide range of characteristic sizes is clearly possible. We have previously given an expression for the amount of disk material M_d that can become centrifugally supported by collapse through the halo. Viscosity acts to redistribute the material into an exponential disk of characteristic scale α^{-1} . Consequently, the central surface density of the disk can be estimated as

$$\Sigma_{\text{central}} = M_d \alpha^2 / (2\pi) \\ = 2000(\lambda/0.06)^2 (0.1/\eta)^{3/2} (\Sigma_h/100 M_\odot \text{ pc}^{-2}) \\ \times (R_h/10\bar{v}_0 t_*)^3 M_\odot \text{ pc}^{-2}. \quad (27)$$

This value is consistent with the central surface density of disk galaxies (Freeman 1970), and $M/L \sim 10$.

c) Bulge Characteristics

The spheroidal components of disk galaxies form analogously to elliptical galaxies, cloud collisions and

dissipation resulting in the formation of centrally condensed bulges. The surface brightness profiles and metallicity gradients should resemble those of ellipticals. For at least two reasons, rotational support will generally be significant. The bulges have formed by infall through a dark halo, thereby enhancing their specific angular momentum by a factor $\sim \lambda^{-1}$ (§ IVd). Moreover, continuing accretion of disk matter and star formation will result in the acquisition of angular momentum, as well as the erasure of any preexisting metallicity gradient. The resulting velocity dispersions may be reduced because of rotational support, but this effect should not be very pronounced because there is likely to be a broad range in V_p/σ values. Bulges are therefore expected to satisfy a similar L - σ relation to ellipticals, with a slightly lower σ at a given value of L . Observations are consistent with this expectation (Whitmore, Kirshner, and Schechter 1979).

It is possible that the recent discovery of a correlation between near-infrared luminosity and H I line width for disk systems (Aaronson, Huchra, and Mould 1979) can also be interpreted in a similar manner. The correlation can be expressed as $L_{\text{IR}} \propto V_p^4$, and is similar to that for the spheroidal component, a result that should not be unexpected if the disk is dominated by the halo and bulge potentials and L_{IR} measures the old stellar population of red giants. This old population is similar in color indices to elliptical galaxy colors. The near-infrared luminosity is likely to be bulge-dominated for early-type spirals, although for late-type systems (Sc's and irregulars) it is likely to be predominantly due to the old disk population. The fact that the old disk component exhibits a L - σ relation similar to that found for ellipticals suggests that a truncation similar to that hypothesized for the spheroidal components may have limited the concentration of luminous matter, at least in the direction normal to the disk.

Our estimates of bulge and disk radii can be combined to yield a crude bulge-to-disk ratio. Note that the disk radius $\propto^{-1}$ is independent of disk mass and depends only on the uncertain turbulence and star formation parameters, whereas the bulge radius is approximately proportional to $M_B^{1/2}$. Consequently, applying equation (20) to the bulge component, we obtain

$$R_B/R_d \sim M_B^{1/2} [\frac{1}{3} v_0 t_* (\mu\eta)^{1/2}]^{-1}. \quad (28)$$

Quantitative estimates of bulge to disk ratios are given by Dressler (1980), namely $R_d/R_B = 1 \pm 1, 4 \pm 2, 8 \pm 2$ for Sa, Sb, and Sc galaxies, respectively, where, over this range, total luminosities decrease by approximately 5 mag in the near-infrared where the bulge and old disk components dominate the light (Aaronson, Huchra, and Mould 1979), roughly consistent with the predicted trend.

d) Disk Formation Time and Metallicity Gradients

It is obvious that any attempt to quantitatively estimate the viscosity in the protogalactic disk is fraught with uncertainty. We have assumed that, within a time $t \lesssim t_*$, the viscous drag is sufficient to establish an exponential disk. To demonstrate that the requirements on viscosity are not unreasonable, we note that at present the viscosity associated with collisions of molecular clouds in the

interstellar medium of our Galaxy imply a turbulent viscosity of $\nu \sim 2 \text{ km s}^{-1} \text{ kpc}$ (Norman and Silk 1980b). If, during disk formation, the viscosity is larger than at the present epoch by between one and two orders of magnitude, the disk formation time $\sim r^2/\nu \sim 10^9$ – 10^{10} years. Because the protogalaxy is largely gaseous and is considerably more active than the present quiescent disk, this requirement on viscosity does not seem unreasonable. The fact that the cloud motions are highly supersonic does not enter directly: the parameter η takes account of the dissipation efficiency. The required viscosity for slow disk formation implies $\bar{v}_0 \gtrsim 30 \text{ km s}^{-1}$ and $l_0 \gtrsim 2 \text{ kpc}$. Since we envisage the halo to provide a reservoir of orbiting and infalling clouds, these extreme values may be attainable. The relevant comparison at the present epoch is with the *outermost* disk, which our model implies must have formed most recently.

Moreover, for viscosity to be effective over a time scale in excess of 10^9 yr, star formation must be relatively slow. Provided that the star formation time scale is of the same order as the disk formation time scale, a metallicity gradient will develop. A similar argument has previously been given for the origin of metallicity gradients in spherical systems (§ IVg). We note here that an important prediction of our dissipative model of slow disk formation is that there should be no strong early burst of star formation.

VI. GALAXY CLUSTERING AND S0 FORMATION

Dissipative disk formation will only be able to occur in relatively quiescent regions. Interactions with neighboring galaxies will clearly tend to result in the destruction of nascent gaseous disks. We now argue that lenticular galaxy formation is associated with the enhanced rate of merging during galaxy clustering. The onset of galaxy clustering is generally considered to begin at $z \sim 5$ – 10 . During the initial collapse of a rich cluster, many galaxy collisions will occur and result in mergers. Once the protogalaxies have fallen through the potential well of the cluster, further mergers will be mostly inhibited. We shall show that evidence for the occurrence of protogalaxy mergers during cluster collapse is contained in the recent study by Dressler (1980) of galaxy type as a function of local density in rich clusters.

a) Dependence of Lenticular Fraction on Density

Consider the following simple model for the evolution of a protocluster. Protospirals fall together, acquiring a virialized velocity that first exceeds their typical internal velocity, σ_g , after a time $t_f \sim \sigma_g (G\rho R)^{-1}$, where ρ is the local density and R is the distance from the cluster center. After this time, merging ceases (Roos and Norman 1979). Let us assume that mergers result in lenticular or elliptical galaxy formation. If γ is the mean collision rate, the rate of mergers in a shell of radius r with initial density n_0 of spirals is $dn_m/dt = \gamma n_{\text{sp}}^2$, and the rate of spiral galaxy destruction is

$$\frac{dn_{\text{sp}}}{dt} = -2\gamma n_{\text{sp}}^2,$$

where $n_{sp} + 2n_m = n_0$ and n_{sp} and n_m denote the local densities of spiral and merged galaxies. It follows that

$$\frac{n_m}{n_0} = \frac{\gamma n_0 t}{1 + 2\gamma n_0 t}. \quad (29)$$

We have neglected radial mixing between shells. Identifying t with t_f and assuming a cluster density profile $\rho \propto (r^2 + a^2)^{-3/2}$, we infer that the fraction of merged galaxies depends on local density according to

$$\frac{n_m}{n_0} = \frac{a\rho^{1/3}}{1 + 2a\rho^{1/3}}. \quad (30)$$

This weak dependence on density is consistent with the results of Dressler's (1979) study if we consider that the local galaxy density can be regarded as a monotonic function of a scale parameter that refers either to distance from the cluster center or to the distance from the center of a bound substructure in the early stages of collapse, more appropriate to an unrelaxed cluster. Comparison of equation (30) with Dressler's data is shown in Figure 1. The observed weak density dependence of the merged fraction is also consistent with the N -body simulation results discussed by Roos and Norman (1979).

b) Relative Abundance of Galaxy Types

In fact, elliptical and S0 galaxies display substantially different variations with density. The lenticular fraction rises rather slowly whereas the elliptical fraction is much more sharply peaked at high galaxy densities. To explore the implications of this, we must develop a more general model for the evolution of morphological type in mergers.

Let us suppose that the initial merger between two protospirals results in the formation of an S0. The idea is that sufficient gas remains after the merger to form a disk.

Since the merger effectively triggers star formation in the gas clouds that remain bound to the resulting galaxy, one is likely to end up with a less prominent and relatively gas poor disk than is associated with a spiral. In the case of the S0, the continuing supply of gas clouds to refuel star formation and disk accretion effectively ceases. A protospiral, however, is characterized by continuing gas supply. Our formation theory leads to an initial differentiation between S0's and spirals, and there may be no need to invoke such secondary effects as supernova-driven winds or ram-pressure stripping as primary S0 formation mechanisms. Subsequent mergers, between protospirals and proto-S0's or between pairs of proto-S0's, are expected to destroy any residual gas and erase any disk, forming elliptical galaxies.

There is one complicating factor that must be considered, however. This involves the relative velocities of the colliding protogalaxies. If the relative velocity is too great, a merger will not occur, although a protospiral may be depleted of gas to form an S0. In rich clusters, mergers are not presently occurring when galaxies collide. Mergers are very likely to occur, of course, during the initial collapse phase, prior to cluster virialization. Later mergers, between proto-S0's, and mergers involving ellipticals will occur with decreasing probability because of virialization.

A simple merger scheme that allows for many of these possibilities is indicated in Table 2. We assume that protospirals collide to form proto-S0's with unit probability, since these systems represent the basic units. Suppose that a protospiral and a proto-S0 collide, to merge into a protoelliptical with probability a (< 1) or else to result in two proto-S0's (the protospiral being effectively stripped). Two proto-S0's collide, to merge into a protoelliptical with probability b (< 1). Similarly, protoellipticals merge with protospirals, proto-S0's, or protoellipticals to form protoellipticals with respective probabilities c , d , f . In all cases, collision results in change

FIG. 1.—Merged galaxy fraction from equation (30) as function of local galaxy density (dashed line), compared with data from Dressler (1980) (solid line).

TABLE 2
MERGER SCHEME

1.	Sp + Sp → S0
2.	Sp + S0 $\xrightarrow{1-a}$ E ↓ S0 + S0
3.	SO + SO $\xrightarrow{1-b}$ E ↓ SO + SO
4.	E + Sp $\xrightarrow{1-c}$ E ↓ E + S0
5.	E + SO $\xrightarrow{1-d}$ E ↓ E + SO
6.	E + E $\xrightarrow{1-e}$ E ↓ E + E

in morphological type either when a protospiral collides or when a merger occurs.

Under these simple assumptions, we can now write down a set of kinetic equations for morphological type evolution as a consequence of galaxy mergers. The equations for morphological type evolution can now be expressed in the following form:

$$\frac{1}{\gamma} \frac{dn_{sp}}{dt} = -2n_{sp}^2 - n_{s0}n_{sp} - n_en_{sp}, \quad (31)$$

$$\frac{1}{\gamma} \frac{dn_{s0}}{dt} = n_{sp}^2 + (1-2a)n_{sp}n_{s0} - 2bn_{s0}^2 + (1-c)n_en_{sp} - dn_en_{s0}, \quad (32)$$

and

$$\frac{1}{\gamma} \frac{dn_e}{dt} = bn_{s0}^2 + an_{s0}n_{sp} - fn_e^2. \quad (33)$$

All of these collisional mergers are understood to occur during the initial phases of cluster collapse or in regions where the relative collision velocities are less than the internal velocity dispersion of the colliding systems.

By considering the ratios of equations (31), (32), and (33), first integrals of the system can be obtained. These directly yield the relative abundances of different types of galaxies. It is convenient to introduce the variable $x \equiv n_{sp}/n_{s0}$, which decreases monotonically with increasing galaxy density. The following implicit solutions to equations (31), (32) and (33) are obtained in the limit that collisions of protospirals and proto-S0's with proto-ellipticals are neglected:

$$\frac{n_{sp}}{n_0} = \left[\frac{x^2}{n^2 + (3-2a)x + 1-2b} \right]^{1/2(1-2b)} \times \exp \left[\frac{1+2a-4b}{(1-2b)\Delta} \left(\tan^{-1} \left(\frac{2x+3-2a}{\Delta} \right) - \frac{\pi}{2} \right) \right], \quad (34)$$

$$n_{s0} = n_{sp} x^{-1}, \quad (35)$$

and

$$n_e = \int_x^\infty \frac{(b+ax)n_{sp}}{x^2(x^2 + (3-2a)x + 1-2b)}, \quad (36)$$

where

$$0 < b < \frac{1}{2} - \frac{1}{2}(\frac{3}{2} - a)^2, \quad \frac{1}{2} < a < 1. \quad (37)$$

We have assumed that $n_e = 0$ as $x \rightarrow \infty$ (i.e., in the field where $n_{sp} \gg n_{s0}$). This is required by our simple model, which, however, could obviously be generalized if necessary to include an additional uniform component of ellipticals. In Figure 2, we display our solution for $a = 1$, $b = 0$, and for comparison show also data obtained from the study by Dressler (1980). One would obviously prefer a solution of equations (31)–(33) in which the mergers between protoellipticals and disk systems were not neglected. There are sufficiently many parameters in the full system of equations that it seems fruitless to pursue this

FIG. 2.—Fractional abundances of elliptical, spiral and irregular, and S0 galaxies, expressed as a function of the ratio of S0 to spiral (and irregular) galaxies and computed from a simple merger model (eqs. [31], [32], and [33]). Solid lines are calculated for $a = 1$, $b = 0$, and compared with data from Dressler (1980).

point, other than to note that the asymptotic behavior is as follows: define $y = n_e/n_{s0}$, then for large x and small y ,

$$xy = a, \quad n_{sp} \sim \exp(-2/x) \quad (38)$$

and for small x and large y ,

$$y \sim x^{-(d-f)/(1-d)}, \quad n_{sp} \sim x^{1/(1-d)}. \quad (39)$$

This solution has the same asymptotic behaviour as the solution to the restricted set of equations plotted in Figure 2, provided that $d > f$.

How reasonable are the restrictions imposed on the merger parameters in equations (31)–(33)? Clearly, the progression toward virialization makes it seem plausible that $a > b > c > d > f$. One cannot really say any more without recourse to the complications of a dynamical model. In general, theory and observation are in remarkably good agreement apart from one data point, namely that for field ellipticals. Allowance for an unclustered distribution of ellipticals, possible due to dynamical escape from denser regions, while detracting from the simplicity of the merger model, could completely reconcile model and data points if the uniform component amounts to a fractional population less than 0.1. The fact that the fractional populations of galaxy types are determined over a wide range of density (10^{-2} to 10^{+2} galaxies Mpc^{-3}) strongly supports the merging hypothesis. Moreover, it seems clear that *mergers of protospiral systems form S0 galaxies and subsequent mergers between S0 and protospiral galaxies of comparable mass result in elliptical galaxy formation*. This suggests that the morphology of galaxies spans a continuous range.

We can attempt to understand these results as follows. The disks of massive galaxies will be slow to form, and

their formation will be inhibited by galaxy collisions, particularly when clustering commences. An additional indication that galaxy formation is still occurring at the epoch of clustering is inferred from evidence (Strom and Strom 1979) for a correlation between galaxy orientation and cluster flattening. We may expect that in the initial collision of a protospiral galaxy with a similar system of comparable mass, the gas will only be partially removed, allowing disk formation to recommence. However, subsequent mergers will remove any remaining gas and also disrupt the stellar disk component, thereby forming elliptical systems. In this manner, the enriched gaseous material destined to form a disk in an isolated system is utilized in clusters to form the disks of potential S0 galaxies in the first collision, and is ejected subsequently into the intracluster medium (Binney and Silk 1978; Norman and Silk 1979).

Our results indicate that the initial number density of galaxies is reduced by $\sim 60\%$ as a consequence of mergers. This implies that the characteristic elliptical should be about 1 mag more luminous than the spheroidal component of a typical spiral with V_p/σ decreased by a factor of $\sim 2.5^{3/2} \sim 4$, and with S0 bulges forming an intermediate class. In general, a continuous sequence of properties is expected between ellipticals and the spheroidal components of S0's and of spirals. This should include rotation and metallicity, as well as luminosity.

VII. SUMMARY AND CONCLUSIONS

The hypothesis that protogalaxies consist of an aggregate of interacting gas clouds can account for many of the existing features of galaxies. This protogalaxy model is a natural consequence of the gaseous fragmentation theory in which galaxies form from massive primordial protogalaxy fluctuations. It can also be reconciled with the hierarchical clustering theory of galaxy formation if the clustering subunits are assumed to retain their gaseous character by means of internal energy input from nondissipative star formation.

A principal feature of the theory is that the system of bound gas clouds can behave in many ways like test particles in N -body simulations, while also undergoing dissipative processes, such as internal star formation and inelastic cloud-cloud collisions. Our galaxy formation model depicts a protogalaxy as a system of self-gravitating, self-sustaining gas clouds, long-lived by virtue of internal nondisruptive star formation. In isolation, such a system gradually dissipates its kinetic energy via cloud-cloud collisions. Angular momentum is conserved, and the system spins up, V_p/σ increasing by a factor $\sim \lambda^{-1}$ in the presence of a dark halo. We identify this system as a protospiral, the disk of which forms over 10^9 – 10^{10} years as the clouds collide, induce star formation, and disrupt. The bulge forms in the initial collapse, as a consequence of the violent relaxation of core-halo segregation of the interacting clouds. The bulge component is defined by the condition that $t_{\text{coll}} < t_{\text{dyn}}$: clouds satisfying this condition are expected to rapidly collide and form stars. Preexisting stars will acquire less central concentration than newly formed stars as a result of the

collisions; hence a metallicity gradient will develop in the bulge.

This formation mechanism induces a characteristic (L, R, σ) -relation. Protogalaxy mergers at the onset of clustering lead first to S0 formation, and, in subsequent mergers, to elliptical formation. Successive mergers destroy any preexisting disk or cloud reservoir that can potentially form a disk. We argue that the first merger between protospirals forms an S0, whose disk forms as a consequence of merger-induced star formation. A merger injects a supply of clouds into the core of the new system; these clouds collide and form stars, resulting in the formation of the condensed cores of ellipticals. The core parameters are defined by the condition $t_{\text{coll}} < t_{\text{dyn}}$; this yields characteristic (L, R, σ) -relations, as for bulge components. The low rotation of ellipticals can be understood in terms of mergers from marginally bound orbits: bulges of disk galaxies are expected to acquire a degree of rotational support that is intermediate between disks and ellipticals. We have inferred from the dependence of morphological type on local density that the typical elliptical in a rich cluster has undergone from two to three protospiral mergers; this suffices to reduce λ from ~ 0.4 (characteristic of disks) by a factor $\sim 2^{3/2}$ – $3^{3/2}$ or ~ 3 – 5 to the value ~ 0.07 characteristic of ellipticals (cf. Fall 1979). It would be extremely interesting to compare "field" and cluster ellipticals: our arguments indicate that the former may have undergone fewer mergers and therefore have higher values of λ . In general, S0's form a population whose properties are intermediate between spirals and ellipticals as a consequence of their formation via single mergers between protospirals.

We have found it possible to give a plausible explanation for the origin of both the velocity anisotropy and metallicity gradients in elliptical galaxies via protogalaxy mergers, where some memory of the collision dynamics is retained in the stars formed as a consequence of cloud collisions. This encouraged us to attempt to quantitatively explain many of the fundamental observed properties of galaxies listed in Table 1. Although the estimates given are undoubtedly simplistic, we can crudely account in a quantitative way for the most important properties of elliptical and disk galaxies.

Massive halos may play an important role in establishing the initial conditions of our model, helping to determine the angular momentum acquired by the luminous material in relatively isolated disk galaxies. In contrast, elliptical galaxies are formed by successive mergers of protogalaxies where it is likely that the extended dark halos have been tidally stripped as galaxy clustering develops. Presumably ellipticals in groups of galaxies could have retained any preexisting dark halos: their low rotation is attributed to the short-lived role of dissipation during their formation, in contrast to the continuous dissipative growth of disk galaxies: even the bulges of disk systems will accrete lower angular momentum material and consequently acquire more rotation. Protogalaxy mergers during the initial collapse of groups and clusters account in detail for the relative abundances of spiral, S0, and elliptical galaxies.

The slow gas dissipation rate in an isolated protogalaxy implies that it gradually develops a disk over 10^9 – 10^{10} yr. A simple viscous interaction model enables us to infer a quasi-exponential density distribution for the disk. Continuous infall of gas into the disk implies that the observed low frequency of occurrence of low-metallicity disk stars can be understood. As rich clusters of galaxies develop over a similar time scale, much of the enriched gas that otherwise would form a disk in an isolated system is collisionally stripped and fills the intracluster medium.

It is possible that relatively isolated clouds can remain gas-rich until late epochs, and could either effectively exhaust the bulk of their gas supply or be tidally shocked, the net effect being catastrophic star formation at a relatively recent epoch. The formation of young globular cluster-like systems and gas-rich dwarfs would be possible consequences. We finally note various consequences and predictions of our model that distinguish it from alternative schemes of galaxy formation. Galaxy formation is *not* associated with a strong burst of star formation. Rather, star formation occurs gradually in preexisting clouds, and galaxies form out of enriched material with $Z \lesssim 10^{-3} Z_{\odot}$. Protogalaxy mergers will trigger enhanced star formation through cloud-cloud collisions, and elliptical galaxies should therefore have a bright protogalactic phase at the epoch of galaxy clustering, $z \lesssim 5$ – 10 . The duration of this phase is determined by the cluster collapse time, typically $\sim 10^9$ yr, when collisional mergers occur. Protospiral galaxies will be much less conspicuous than ellipticals, because star formation proceeds continuously for $\lesssim 10^{10}$ yr. The disks of spiral and S0 galaxies should be significantly younger than the spheroidal components, especially in the outermost regions. A limiting binding energy is predicted for luminous matter, most notably for the cores of elliptical galaxies. The amount of *enriched* intergalactic matter per unit luminosity stripped from galaxies should be substantially more in rich clusters than in loose groups of galaxies. Essentially none should be present in the intergalactic medium, unless galaxy and cluster winds have played a major role in ejecting enriched material. Elliptical galaxies, particularly members of groups or in the outlying regions of clusters, could possess nascent stellar disks, since collisional mergers and stripping should not have been as effective for these systems as for those in higher

density regions. Many of the structural properties induced by mergers, such as velocity anisotropy, v_p/σ , flattening, and metallicity gradients, should reveal a dependence on local density. A substantial number of ellipticals do not possess metallicity gradients. Dissipation, which tends to increase v_p/σ and velocity anisotropy, should be correlated with the weakness of metallicity gradients.

Throughout the preceding discussion, we have stressed the role of protogalaxies in mergers. It is possible that many of our arguments are equally valid for mergers between conventional spiral galaxies. We cannot eliminate this possibility, other than to remark that an exceedingly high efficiency of star formation is required. Even if purely dynamical effects can produce the centrally condensed cores of ellipticals, the real hurdle to be surmounted is the generation of metallicity gradients. These do not survive the violent relaxation associated with mergers, and must be generated anew from the available gas supply. The entire gas supply of two late-type spirals would barely suffice to yield the required gradient in the merged system, if converted to stars with great efficiency. A similar constraint is required for spiral mergers to produce the characteristically large bulges of S0 galaxies. Protogalaxy mergers provide a natural means of circumventing this difficulty: the gas supply available for star formation is greatly increased.

We emphasize finally that our hybrid model simultaneously resolves two principal characteristics of galaxies that existing theories of galaxy formation fail to adequately address, namely the velocity anisotropies and metallicity gradients in spheroidal systems. At first sight, these characteristics seem mutually contradictory: we contend that they provide the best indication that both dissipation and violent relaxation play fundamental roles in galaxy formation. Such a hybrid approach seems to us to be required to quantitatively account for many of the detailed structural properties of galaxies.

This research has been supported in part by the NATO Scientific Affairs Division and by NASA grant NGRO5-003-578. We acknowledge helpful discussions with Drs. S. M. Fall, G. Illingworth, J. Kormendy, N. Roos, F. Schweizer, and S. D. M. White. We are especially indebted to Dr S. M. Faber for her detailed and constructive comments in the manuscript.

APPENDIX A

CLOUD HYDRODYNAMICS

The generalized kinetic equation for a system of clouds with distribution function $N = N(M, r, v, t)$ in a gravitational potential is

$$\begin{aligned} \frac{\partial N}{\partial t} + \mathbf{v} \cdot \nabla N + \nabla \phi \cdot \frac{\partial N}{\partial \mathbf{v}} &= \frac{1}{2} \int_0^M \int_{-\infty}^{\infty} N \left(M - M', r, \frac{M\mathbf{v} - M'\mathbf{v}'}{M - M'}, t \right) N(M', r, \mathbf{v}', t) \sigma(M - M', M') \left| \frac{M}{M - M'} (\mathbf{v} - \mathbf{v}') \right| \left(\frac{M}{M - M'} \right)^3 d^3 \mathbf{v}' dM' \\ &\quad - N(M, r, \mathbf{v}, t) \int_0^{\infty} \int_{-\infty}^{\infty} N(M', r, \mathbf{v}', t) \sigma(M - M', M') \left| \frac{M}{M - M'} (\mathbf{v} - \mathbf{v}') \right| \left(\frac{M}{M - M'} \right)^3 d^3 \mathbf{v}' dM' \\ &\quad + \frac{1}{2} q g(M, p) \int_0^{M_1} \int_{-\infty}^{\infty} N \left(M - M', r, \frac{M\mathbf{v} - M'\mathbf{v}'}{M - M'}, t \right) N(M', r, \mathbf{v}', t) \sigma(M - M', M') \\ &\quad \times \left| \frac{M}{M - M'} (\mathbf{v} - \mathbf{v}') \right| \left(\frac{M}{M - M'} \right)^3 d^3 \mathbf{v}' dM', \end{aligned} \quad (\text{A1})$$

where σ is the cloud-cloud collision cross section, and where we have assumed that a fraction q of massive clouds above mass M_1 are injected into the distribution with initial distribution $g(M, p)$. Here we assume that the cross section is independent of velocity.

This is the momentum-conserving Boltzmann-coagulation equation in a gravitational field. Usually, the coagulation equation is solved in a homogeneous infinite medium, but in this case it is necessary to include effects of the spatially inhomogeneous gravitational potential and cloud distribution. Let us first consider two basic limits.

$$a) \quad t_{\text{coll}} > t_{\text{dynamical}}$$

Here collisions and coagulation are unimportant, and the collisionless Vlasov equation describes the evolution of the system by

$$\frac{\partial N}{\partial t} + \mathbf{v} \cdot \nabla N + \nabla \phi \cdot \frac{\partial N}{\partial \mathbf{v}} = 0. \quad (\text{A2})$$

In this case of merging of protogalaxies, the N -body simulation results will be valid over time scales up to $\sim t_{\text{coll}}$. Merging, for example, takes several crossing times; and if $t_{\text{coll}} \gtrsim 3t_{\text{cross}}$, relaxation will have occurred before any collisional effects take place. Consequently, the resulting radial distribution should be approximately a Hubble law in mass density as found in the simulations (Roos 1979; White 1979a). We envisage that this violently relaxed state is reached outside the cores of merged systems.

$$b) \quad t_{\text{coll}} \leq t_{\text{dynamical}}$$

In this case, not only is there buildup of cloud masses by coalescence, but there is also loss of orbital kinetic energy in the momentum-conserving inelastic collisions, due to radiative cooling in collision-induced shocks. In this limit, we have to address the full kinetic coagulation equation. Although this is a complex problem, our task is greatly simplified if we seek a steady state distribution. This problem naturally appears as a two-time-scale problem. On short time scales $\sim t_{\text{coll}}$, it is only the coagulation terms on the right hand side of (A1) which drive the evolution. Denoting these terms by $C \equiv C(M, \mathbf{v}, r, t)$, the kinetic equation becomes

$$\frac{N}{t} = C(M, \mathbf{v}, r, t). \quad (\text{A3})$$

On a longer dynamical time scale, the system will respond to the Vlasov equation

$$\frac{\partial N}{\partial t} + \mathbf{v} \cdot \nabla N + \nabla \phi \cdot \frac{\partial N}{\partial \mathbf{v}} = 0, \quad (\text{A4})$$

but with $C = 0$.

Using this two-time-scale approximation, our work is now greatly simplified. If the injected mass is all injected at the same lowest mass M in the spectrum and if the injected momentum distribution is Gaussian,

$$g(M_0, v) = \frac{1}{2^{1/2}\pi v_0} \exp\left(-\frac{\frac{1}{2}Mv^2}{v_0^2}\right), \quad (\text{A5})$$

the solution of equation (A3) is a Maxwellian velocity distribution (Penston *et al.* 1969)

$$N(M, v, r) = R(r) \frac{(M/M_0 - \frac{3}{2})!}{(M/M_0)!} T \frac{M}{M_0} \frac{(M/M_0)^{1/2}}{2^{1/2}\pi m_0 v_0} \exp\left(-\frac{Mv^2}{2v_0^2}\right). \quad (\text{A6})$$

Here $R(r)$ is a space-dependent factor, T is the injection term in equation (A1) written as $T\delta(M - M_0)$, and we have specifically assumed that $\langle\sigma v\rangle$ is constant. The solution of equation (A4) for a stationary state then gives us the spatial dependence term in (A6), yielding the general solution as

$$N(M, v, r) = \frac{(M/M_0 - \frac{3}{2})!}{(M/M_0)!} T \frac{M}{M_0} \frac{M/M_0}{2^{1/2}\pi m_0 v_0} \exp\left(-\frac{\frac{1}{2}Mv^2 + \phi(r)}{v_0^2}\right). \quad (\text{A7})$$

This shows that there is equipartition of energy among the clouds in the core regions, assuming the potential is given. Hence, the more massive clouds will segregate toward the central regions. The self-consistent solution to this equation when the clouds themselves dominate the gravitational potential should lead to a core-halo structure.

Another possible approach is to look at moments of the kinetic equations. Here a useful simplifying approximation is available if the mass spectrum of clouds is such that most of the mass in clouds is at the low mass end of the spectrum. Let us assume that in this system there is no external source of clouds. The destruction term is

$$\frac{N(M, v, r, t)}{\tau(M, v, r, t)},$$

where

$$\tau^{-1}(M, v, r, t) = \int_{-\infty}^{\infty} N(v', M', r, t) \sigma(M - M', M') \left| \frac{M}{M - M'} |v - v'| \right| d^3v' dM',$$

and the kinetic equation becomes

$$\frac{\partial N}{\partial t} + v \cdot \nabla N + \nabla \phi \cdot \frac{\partial N}{\partial v} = -\frac{N}{\tau}. \quad (\text{A8})$$

For τ independent of N , this is the normal Vlasov equation with a collision term. However if τ is a function of N , the situation is more complex. In taking moments of the equation, we shall assume that we can model the right hand side of equation (A8) by

$$-\frac{N}{\tau} = -\alpha_1 N^2 \sigma v_{cl}(r), \quad (\text{A9})$$

where α_1 is a factor of order unity. More specifically, if all clouds have the same mass, cross section, and velocity, then this expression would be exact with $\alpha_1 = 1$.

We further assume that an individual cloud loses a constant fraction ϵ of its energy per collision, that clouds spend most of their time in a given spherical shell whose thickness is small compared to the overall system size, and that at a given radius all clouds have the same mean mass. The cloud hydrodynamic equations now become under these approximations:

$$\frac{d}{dt} N(r) = -N^2 \sigma v_{cl}(r) \quad (\text{A10})$$

and

$$\frac{d}{dt} E_{cl} = -N \sigma V_{cl} E_{cl}, \quad (\text{A11})$$

where

$$E_{cl} = \frac{1}{2} M v_{cl}^2$$

and

$$\frac{d}{dr}(NMv_{cl}^2) = -4\pi G \frac{NM}{r^2} \int_0^r r'^2 N(r') M(r') dr' . \quad (A12)$$

We look for a solution of steady mass inflow at a rate $\dot{M}$, due to orbital energy dissipation in cloud-cloud collisions, with

$$\dot{M} = 4\pi r^2 v_{cl,r}(r) M(r) N(r) \quad (A13)$$

and

$$\frac{d}{dt} \equiv \frac{\dot{M}}{4\pi r^2 M(r) N_{cl}(r)} \frac{d}{dr} . \quad (A14)$$

We have assumed a constant cloud cross section, and $v_{cl,r}$ is the mean cloud radial velocity.

Using the cloud hydrodynamics equation (A10)–(A14), we find from (A10) and (A11), a first integral

$$N = KMv_{cl}^2 . \quad (A15)$$

From (A15) and (A10) we find

$$v_{cl}^{-1} = \frac{\dot{M}K}{4\pi r^2 \sigma_{cl} N^{\epsilon+2}} \frac{dN}{dr} , \quad (A16)$$

which gives with (A15)

$$M = \frac{K\dot{M}^2}{(4\pi)^2 r^4 \sigma} N^{-(\epsilon+4)} \left(\frac{dN}{dr} \right)^2 , \quad (A17)$$

Then (A12), (A16), and (A17) produce

$$\frac{d}{dr} N(r) = - \frac{GK^3 \dot{M}^4 N^{-(2\epsilon+3)}}{(4\pi)^3 r^6 \sigma^4 (\epsilon+1)} \left(\frac{dN}{dr} \right)^2 \int_0^r \frac{N^{-(\epsilon+3)}}{r'^2} \left(\frac{dN}{dr'} \right)^2 dr' . \quad (A18)$$

With a given potential ϕ or given distribution with galaxy mass $M_g(r)$, we use the hydrostatic equation

$$\frac{d}{dr}(NMv_{cl}^2) = 4\pi GNM \frac{M_g(r)}{r^2} . \quad (A19)$$

Then

$$N(r) = \left[\frac{8\pi\sigma(\epsilon+1)}{GK^2 \dot{M}^2} \int_0^r \frac{r'^6}{M_g(r')} dr' + N(0)^{2(\epsilon+1)} \right]^{-1/[2(\epsilon+1)]} . \quad (A20)$$

For a fully self-consistent problem, the cloud hydrodynamics give an integro-differential equation (A18) for $N_{cl}(r)$. If we seek a power-law solution, we obtain the result

$$N(r) \propto r^{-2/(1+3\epsilon/5)} \propto r^{-2} . \quad (A21)$$

In the non-self-consistent case, where cloud self-gravity is neglected and the dominant mass inside galactic radius r is $M_g(r)$, we find from equations (A19) and (A20) that

$$N(r) \propto [1 + (r/r_1)^7]^{(2(\epsilon+1))^{-1}} \propto r^{-3.5} \quad \text{for } M_g(r) = \text{constant} \quad (A22)$$

and

$$N(r) \propto [1 + (r/r_2)^6]^{(2(\epsilon+1))^{-1}} \propto r^{-3} \quad \text{for } M_g(r) \propto r , \quad (A23)$$

where r_1 and r_2 are constants and we take ϵ to be small compared to unity. If $N(r) \propto r^{-\beta}$, then from (A17) the mass distribution $\rho(r) = N(r)M(r) \propto r^{\beta(1+\epsilon)-6}$ and for $\beta = 3$, say, $\rho(r) \propto r^{-3(1+\epsilon)}$.

This simple model of the behavior of the coagulating self-gravitating system surprisingly yields power law solutions for the mass density, similar to a Hubble law. Thus both the numerical N -body simulations, which are applicable outside the core, and the analytical model for the interior region yield a similar power-law density profile.

REFERENCES

Aaronson, M., Huchra, J., and Mould, J. 1979, <i>Ap. J.</i> , 229 , 1.	Binney, J., and Silk, J. 1978, <i>Comments Astr. Ap.</i> , 7 , 139.
Aarseth, S., and Fall, S. M. 1980, <i>Ap. J.</i> , 236 , 43.	Bosma, A. 1977, Ph.D. thesis, University of Groningen.
Aarseth, S. J., Gott, J. R., and Turner, E. L. 1979, <i>Ap. J.</i> , 228 , 664.	Bosma, A., and Freeman, K. C. 1980 (in preparation).
Binney, J. 1978, <i>M.N.R.A.S.</i> , 183 , 501.	Brosche, P. 1970, <i>Astr. Ap.</i> , 6 , 240.
Binney, J., and de Vaucouleurs, G. 1981, <i>M.N.R.A.S.</i> (in press).	Burstein, D. 1979, <i>Ap. J.</i> , 234 , 435.

- Chevalier, R. 1974, *Ap. J.*, **188**, 501.
 Davis, M., and Peebles, P. J. E. 1977, *Ap. J. Suppl.*, **34**, 425.
 de Vaucouleurs, G. 1974, in *IAU Symposium 58, The Formation and Dynamics of Galaxies*, ed. J. R. Shakeshaft (Dordrecht: Reidel), p. 335.
 Doroshkevich, A. G., Shandarin, S. F., and Saar, E. 1978, *M.N.R.A.S.*, **184**, 643.
 Dressler, A., 1980, *Ap. J.*, **236**, 351.
 Efstathiou, G. 1979, *M.N.R.A.S.*, **187**, 117.
 Efstathiou, G., and Ellis, G. 1978, *M.N.R.A.S.*, **185**, 55.
 Efstathiou, G., and Jones, B. J. T. 1979, *M.N.R.A.S.*, **186**, 133.
 ———. 1980, *Comments Ap.*, **8**, 169.
 Einasto, J., Kaasik, A., and Saar, E. 1974, *Nature*, **250**, 309.
 Faber, S. 1977, in *The Evolution of Galaxies and Stellar Populations*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 157.
 Fall, S. M. 1979, *Nature*, **281**, 200.
 Fall, S. M., and Efstathiou, G. 1980, *M.N.R.A.S.*, **193**, 189.
 Fall, S. M., and Rees, M. J. 1977, *M.N.R.A.S.*, **181**, 37.
 Fish, R. 1964, *Ap. J.*, **139**, 284.
 Freeman, K. C. 1970, *Ap. J.*, **160**, 811.
 ———. 1975, in *Galaxies and the Universe*, ed. A. Sandage, M. Sandage, and J. Kristian (Chicago: University of Chicago Press), p. 409.
 Gott, J. R. 1975, *Ap. J.*, **201**, 296.
 ———. 1977, *Ann. Rev. Astr. Ap.*, **15**, 235.
 Gott, J. R., and Thuan, T. X. 1976, *Ap. J.*, **204**, 609.
 Gott, J. R., Turner, E. L., and Aarseth, S. 1979, *Ap. J.*, **234**, 13.
 Illingworth, G. 1977, *Ap. J. (Letters)*, **218**, L43.
 Kormendy, J. 1977, *Ap. J.*, **218**, 333.
 ———. 1980, in *Proc. ESO Workshop on Two-Dimensional Photometry*, ed. P. Crane and K. Kjar (Geneva: European Southern Observatory), p. 191.
 Kormendy, J., and Illingworth, G. 1981, in preparation.
 Larson, R. B. 1974, *M.N.R.A.S.*, **166**, 585.
 ———. 1976, *M.N.R.A.S.*, **176**, 31.
 Lynden-Bell, D., and Pringle, J. E. 1974, *M.N.R.A.S.*, **168**, 603.
 Marchant, A., and Shapiro, S. 1977, *Ap. J.*, **215**, 1.
 Melnick, J., and Sargent, W. L. W. 1977, *Ap. J.*, **215**, 401.
 Norman, C., and Silk, J. 1979, *Ap. J. (Letters)*, **233**, L1.
 ———. 1980a, *Ap. J.*, **238**, 158.
 ———. 1980b, in *IAU Symposium 87, Interstellar Molecules*, ed. B. Andrew (Dordrecht: Reidel), in press.
 Ostriker, J. P. 1980, *Comments Ap.*, **8**, 177.
 Ostriker, J. P., Peebles, P. J. E., and Yahil, A. 1974, *Ap. J. (Letters)*, **193**, L1.
 Pagel, B. J. 1979, in *Stars and Star Systems*, ed. B. Westerlund (Dordrecht: Reidel), p. 17.
 Peebles, P. J. E. 1974, *Ap. J. (Letters)*, **184**, L51.
 Peebles, P. J. E., and Dicke, R. H. 1968, *Ap. J.*, **154**, 891.
 Penston, M., Munday, V., Strickland, D., and Penston, M. 1969, *M.N.R.A.S.*, **142**, 355.
 Press, W., and Schechter, P. 1974, *Ap. J.*, **187**, 425.
 Rees, M. J., and Ostriker, J. P. 1977, *M.N.R.A.S.*, **179**, 541.
 Roos, N. 1981, *Astr. Ap.*, in press.
 Roos, N., and Norman, C. 1979, *Astr. Ap.*, **76**, 75.
 Sandage, A., and Visvanathan, N. 1978, *Ap. J.*, **225**, 742.
 Sanders, R., and van Albada, T. 1979, preprint.
 Schechter, P. 1976, *Ap. J.*, **203**, 297.
 Schechter, P. L., and Gunn, J. E. 1979, *Ap. J.*, **229**, 472.
 Schweizer, F. 1979, *Ap. J.*, **233**, 23.
 Shields, G., and Searle, L. 1978, *Ap. J.*, **222**, 821.
 Silk, J. 1977, *Ap. J.*, **211**, 638.
 ———. 1978, *Ap. J.*, **220**, 390.
 Silk, J., and Norman, C. 1979, *Ap. J.*, **234**, 86.
 Silk, J., and White, S. D. M. 1978, *Ap. J. (Letters)*, **226**, L103.
 Strom, S., and Strom, K. 1978, *A.J.*, **83**, 1293.
 ———. 1979, *A.J.*, **84**, 1091.
 Sunyaev, R. A., and Zel'dovich, Ya. B. 1972, *Astr. Ap.*, **20**, 189.
 Tinsley, B. M., and Larson, R. B. 1978, *Ap. J.*, **221**, 554.
 ———. 1979, *M.N.R.A.S.*, **186**, 503.
 Toomre, A. 1977, in *The Evolution of Galaxies and their Population*, ed. B. M. Tinsley and R. B. Larson (New Haven: Yale University Observatory), p. 401.
 van den Bergh, S. 1979, *Ap. J. (Letters)*, **230**, L161.
 White, S. D. M. 1976, *M.N.R.A.S.*, **177**, 717.
 ———. 1979a, *M.N.R.A.S.*, **189**, 831.
 ———. 1979b, *Ap. J. (Letters)*, **229**, L9.
 White, S. D. M., and Rees, M. J. 1978, *M.N.R.A.S.*, **183**, 341.
 Whitmore, B. C., Kirshner, R. P., and Schechter, P. L. 1979, *Ap. J.*, **234**, 68.
 Zel'dovich, Ya. B. 1978, in *IAU Symposium 79, The Large Scale Structure of the Universe*, ed. M. S. Longair and J. Einasto (Dordrecht: Reidel), p. 409.

COLIN NORMAN: Huygens Laboratory, University of Leiden, Wassenaarseweg 78, Leiden 2300 RA, Netherlands

JOSEPH SILK: Department of Astronomy, University of California, Berkeley, CA 94720