THE ASTROPHYSICAL JOURNAL, 392:41–53, 1992 June 10

MERGING IN COSMIC STRUCTURES

A. CAVALIERE,[1] S. COLAFRANCESCO,[2] AND N. MENCI[1]

Received 1991 August 7; accepted 1991 December 4

ABSTRACT

We discuss the evolution of gravitational systems under merging interactions between their components. Specifically, we consider galaxies in groups or in the field, and subclusters in forming clusters. The mass distribution evolving under aggregations is described with a kinetic equation, for which numerical and analytical solutions are presented and compared. The results agree quantitatively and show, after a short transient, two regimes: self-similar evolution; or a fast, critical phenomenon occurring over a few crossing times in finite systems with relatively low velocity dispersions. The latter bears the marks of a gravitational phase transition. We compare these findings with observations of groups and clusters of galaxies, and conclude that such gravitational phase transitions are indeed effective or even dominant in two typical environments: in groups, leading to the formation of a giant elliptical or a cD-like galaxy, and in forming clusters, causing fast erasure of substructures. In the "open" field we find that aggregations play only a complementary role relative to direct collapses from initial density perturbations. An intermediate situation may prevail in large-scale structures with weak contrast and slow expansion, like the sheets and filaments which actually modulate the field.

Subject headings: galaxies: clustering — galaxies: interactions

1. THE PROBLEM

The formation of such cosmic structures as groups and clusters of galaxies has been discussed mainly in terms of direct collapses from initially small overdensities that are weakly gravitationally unstable even in a critical FRW universe. Such perturbations are often taken to constitute a random-phase Gaussian field, with power spectrum $\langle|\delta_k|^2\rangle \propto k^v t^{4/3}$, where $v \lesssim -1$ at cluster and smaller scales (see Peebles 1980). On crossing the threshold of nonlinearity the perturbations collapse and virialize in a few dynamical times. For a homogeneous sphere formed from an overdensity of rms amplitude in a critical universe, this implies $\delta \approx \sigma \propto M^{-a} t^{2/3} \sim 1$ with $a = (v + 3)/6$, and leads to a characteristic virialized mass $M_c \propto (1 + z)^{-1/a}$.

In the direct hierarchical clustering (DHC) scenario every step of the hierarchy is directly related to the distribution of the initial overdensities, and the internal densities of the condensations at virialization scale in step with the external background $\rho_a \propto (1 + z)^3$. The nonlinear stages of the collapses have been followed in detail with N-body simulations, which for the high-contrast condensations yield time-dependent distributions of mass (MDs) within a limited dynamic range (see, e.g., Efstathiou et al. 1988; Carlberg & Couchman 1989).

Analytic forms of the MDs have been derived in a quasi-static approximation, either by assuming at each given epoch collapse to take place in all overdense volumes (Press & Schechter 1974); or by assuming collapses to begin only at high peaks of the overdensity field, and estimating a statistically relevant measure of the infallen mass (Doroshkevich 1970; Bardeen et al. 1986; Colafrancesco, Lucchin, & Matarrese 1989). Alternatively, Cavaliere, Colafrancesco, & Scaramella (1991, hereafter CCS 1991) propose a rate equation of purely *differential* structure resolved over times of the order of the dynamical time t_d and with a right-hand side *linear* in the MD, to account for primary collapses at high peaks followed by secondary infall of halos.

All such computations assume substructure to be erased on time scales close to the minimum resolution. The effective resolution is of several t_d in the quasi-static approach, which generally causes problems with overcounting substructures (see discussion by Bond et al. 1991). The resolution is still of order t_d in the approach of CCS (1991). But erasure develops over a variety of scales; generally it takes a few crossing times of the forming structure as a whole, but it may take much longer times for galaxies in large groups or in clusters.

To describe such phenomena in full we must resolve time scales $\lesssim t_d$. In fact, we see both in the sky and in N-body simulations *interactions* in the making among substructures, that involve transfers of energy to internal degrees of freedom. These play an important role during the formation and the subsequent evolution of groups and clusters, not only for permanence of substructure, but also because their back-reactions affect the duration of the overall collapse as discussed by Cavaliere & Colafrancesco (1990, hereafter CC 1990).

Aggregations of subclusters in cluster simulations have been discussed in the papers by Cavaliere et al. (1986); Efstathiou et al. (1988); West, Oemler, & Dekel (1988); and CC 1990. Observational evidence of substructures present, or on their way to merge, stems mainly from X-ray observations (see Jones & Forman 1984; Forman & Jones 1990; CC 1990; Briel et al. 1991). Additional evidence comes from optical counts associated with spectroscopic analyses (see Geller & Beers 1982; Binggeli, Tammann, & Sandage 1987; Dressler & Shectman 1988).

The aggregations of galaxies have been widely investigated, in two limits. The first considers direct merging which follows encounters between *comparable* galaxies, with stickiness increasing with decreasing relative speeds (Toomre 1977; see also Saslaw 1985; Binney & Tremaine 1987). References relevant to such interactions of galaxies in groups include Roos & Norman (1979), Carnevali, Cavaliere, & Santangelo (1981, hereafter CCS 1981), Ishizawa et al. (1983), Merrit (1983), Rich-

[1] Astrofisica, Dipartimento di Fisica II Università di Roma, via E. Carnevale, I-00173 Roma, Italy.
[2] Osservatorio Astronomico di Roma via dell'Osservatorio, I-00040 Monteporzio, Italy.

stone & Malumuth (1983), Barnes (1989), Mamon (1990), and Quinn, Salomon, & Zurek (1991). The second limit envisages interactions between *unequal* galaxies with high relative velocities typical of rich environments; such encounters are best described in terms of dynamical friction (see Alladin, Narasimham, & Ballabh 1988; Richstone 1990).

In the following we will investigate the effect on the MD of interactions leading to aggregations between high-contrast units, both inside bound systems, and out in the "open" field.

The plan of the paper is as follows. In § 2 we set up our analytical tool to describe gravitational "collisions," in the form of an aggregation kinetic equation with a *nonlinear integro-differential* structure. In § 3 we define and discuss specific models for interactions in the field and in bound systems, and give a heuristic preview of the behaviors to be expected for the solutions. Our numerical solutions of the kinetic equation, that substantially improve and extend previous work by Nakano (1966), are presented in § 4. In §§ 5 and 6 we derive analytic asymptotic solutions of the kinetic equation. Comparison of the numerical with the analytical results is made in § 7. In § 8 we summarize and interpret our findings and our technique, and discuss interesting extensions.

2. THE FRAMEWORK

To attack the complexity of merging phenomena we adopt the following framework: $\mathcal{N}$ units of individual mass M, size r, internal density ρ, compose a system with total mass $\mathcal{M}$, overall size R, and average density $\rho_a < \rho$. The system evolves mainly through random binary interactions.

Our formalism to treat the ensuing evolution of the MD will start from the classic aggregation equation of von Smoluchowski (1916). Possible cosmogonic applications have been pointed out by among others Silk & White (1978) and Lucchin (1988). However, the rich mathematical content of this integro-differential equation has been discovered and discussed only recently; see Ernst (1986) for a review oriented toward the physical chemistry of suspensions and Cavaliere, Colafrancesco, & Menci (1991) for phenomena relevant to cosmogony.

We start with the continuous form

$$\frac{\partial N}{\partial t} = \frac{1}{2}\int_0^M dM' K(M', M-M', t)\, N(M', t)\, N(M-M', t) - N(M, t)\int_0^\infty dM' K(M, M', t)\, N(M', t)\,, \tag{2.1}$$

where $N(M, t)$ is the MD in "comoving" form, that is, divided by the ambient mass density ρ_a and normalized to system mass $\mathcal{M}$. The lower integration limits actually mean masses $M_\ell \ll \mathcal{M}$; if necessary, the equation may be formulated in an alternative form free of canceling infinities. In the intrinsically *finite* systems we shall consider, the upper infinite limit is to be replaced by $M_u < \mathcal{M}$.

The formal structure of the right-hand side includes a construction and a destruction term, such as to yield for the moments of the MD

$$\langle M^p\rangle = \int_0^\infty dM N(M, t) M^p \tag{2.2}$$

the relationships

$$\langle \dot{M}^0\rangle < 0\,, \quad \langle \dot{M}\rangle = 0\,, \quad \langle \dot{M}^2\rangle > 0\,, \tag{2.3}$$

when the MD has an upper cutoff.[3] That the second relationship $\langle \dot{M}\rangle = 0$ does not always hold will be one major issue of this paper. Correspondingly, the total mass $\mathcal{M} = \langle M\rangle$ in the system described by equation (2.1) is not conserved.

The physics of the interactions is contained in the kernel $K = \rho_a \overline{\Sigma V}$, where V is the relative velocity and Σ is the gravitational cross section. Their product, averaged over the distribution of relative velocities, is amenable to simple expressions (see Saslaw 1985) in three different regimes depending on the ratio v^2/V^2, where v is the parabolic velocity of escape at the closest approach p. We shall often approximate v with the internal velocity dispersion of the component bodies, that has the same scaling and slightly larger values.

The first regime applies to the case of high-velocity encounters with $V \gg v$, when technically Σ can be calculated in the impulsive approximation (neglecting the change in the potential energy during the encounter, that is, along straight orbits). Then the effective cross section scales as the square of the objects radii, multiplied by a small efficiency for energy transfers $\eta \lesssim v^2/V^2 \ll 1$. Small efficiency implies long time scales for merging, during which other evolutionary phenomena are likely to take over. Such encounters are unlikely to dominate the evolution of cosmic structures that we consider here, and so we shall not pursue them in detail.

A second regime applies when $v^2 \sim V^2$and $p \sim r$ hold. In these conditions, the cross section for merging of two members of masses M and M' not widely different reads $\Sigma \approx \pi(r + r')^2(1 + 2G[M + M']/[r + r']V^2)$. In the numerical computations to follow we shall consider the full expression for the cross section, but useful approximations obtain from considering the two addenda separately, which constitute partial cross sections scaling like M^λ with a different homogeneity degree λ.

The first component describes purely geometrical collisions (GCs) where $\Sigma = \pi(r + r')^2$. The second describes the focused, resonant interactions (FIs) with cross section $\Sigma \approx 2\pi(r + r')^2 G(M + M')/(r + r')V^2$, that prevail in the range $V^2/v^2 \lesssim$ a few, and are most effective for merging.

The member radii scale as $r \propto (M/\rho)^{1/3}$ in terms of the internal density ρ. Alternatively, for galaxies the Faber-Jackson relationship $L \propto v^4$ (Faber 1982) holds at least before they are involved in fast merging activity, and this constraint implies the scaling $r \propto M^{1/2}$ to hold at $M/L \sim$ const, independently of ρ (M. Vietri, 1991 private communication). During the phenomena we are to investigate these two extremes bracket the galaxy scalings. For the sake of simplicity, our exposition will be mainly in terms of the former scaling, and will be complemented as necessary with the (small) changes in the results ensuing from the latter.

The scalings of $\overline{\Sigma V}$ with *mass* and *time* will constitute the key features in equation (2.1). Considering first the characteristic mass $M_*(t) = \langle M^2\rangle/\langle M\rangle$ and the normalized mass $m \equiv M/M_*$, the following scalings obtain: $\overline{\Sigma V} \propto M_*^{2/3} V \rho^{-2/3}\psi(m, m')$ with $\psi \propto (m^{1/3} + m'^{1/3})^2$ for the GC component; or $\overline{\Sigma V} \propto M_*^{4/3}\, V^{-1}\rho^{-1/3}\psi(m, m')$ with $\psi(m, m') \propto (m^{1/3} + m'^{1/3})(m + m')$ for the FI component. To wit, either for

[3] This is a necessary condition for the time derivative of the moments to be written in the symmetrized form

$$\langle \dot{M}^p\rangle = \frac{1}{2}\int_0^\infty \int_0^\infty dM\, dM' K(M, M', t)\, N(M, t)\, N(M', t) \times [(M + M')^p - M^p - M'^p]\,.$$

GCs or for FIs the cross section $\overline{\Sigma V}$ is a symmetric homogeneous function of M and M', with degree $\lambda = 2/3$ or $\lambda = 4/3$, respectively; the latter becomes $\lambda = 3/2$ when the Faber-Jackson relationship applies. The scalings with t are given in the next section.

3. TIME SCALES

We shall see that the solutions of the integro-differential equation (2.1) begin with a transient stage that still remembers the initial conditions, but rapidly goes into a self-similar stage where such memory is lost. Remarkably, in cases where $\lambda > 1$ the memory loss may take the extreme form of a runaway phenomenon that remolds the whole distribution over a few to several crossing times $t_d \sim 2R/V$.

The average merging time is given by $\tau \sim 1/n\overline{\Sigma V} \sim t_d/\mathcal{N}(r/R)^2$. As long as the mass in normal members $\mathcal{M}$ is conserved in the system, their number density will scale like $n \propto \rho_a M_*^{-1}$. Then the average merging rates effective in the kernel of equation (2.1) read

$$\tau^{-1} \propto \frac{\rho_a V}{\rho^{2/3}} M_*^{-1/3} \text{ or } \tau^{-1} \propto \frac{\rho_a}{\rho^{1/3} V} M_*^{1/3} , \tag{3.1}$$

for the GC component or for the FI component, respectively. In a first heuristic look at equation (2.1), a rate $\tau^{-1}(t)$ accelerating with time will point to a runaway process.

When GCs prevail, the explicitly t-dependent coefficient in equation (3.1a) reads

$$\mathcal{F}_{\rm GC}(t) = \rho_a^{1/3}(\rho_a/\rho)^{2/3} V . \tag{3.2}$$

Equation (3.1a) implies that any attempt of GCs to accelerate the rate $\tau^{-1} \propto \mathcal{F}_{\rm GC}(t) M_*^{-1/3}$ will be counteracted by the negative feedback provided by the very growth of $M_*(t)$.

For FIs, on the other hand, the time-dependent coefficient of τ^{-1} in equation (3.1b) reads

$$\mathcal{F}_{\rm FI}(t) = G\rho_a^{2/3}(\rho_a/\rho)^{1/3} V^{-1} . \tag{3.3}$$

As $M_*(t)$ grows by merging, the condition for τ^{-1} to run away is that $\mathcal{F}_{\rm FI}(t)$ does not decrease rapidly. Then a *positive* feedback loop for $M_*(t)$ may set in, leading to divergence of the rate or equivalently to very short effective times, and conceivably inducing peculiar behaviors of the solutions of equation (2.1).

To proceed to a quantitative analysis of such behaviors, we parameterize $\rho_a(t) \propto t^d$, $\rho(t) \propto t^s$, and $V(t) \propto t^u$. Hence the behaviors of $\mathcal{F}_{\rm FI}(t)$ and $\mathcal{F}_{\rm GC}(t)$ may be described in terms of a common form $\mathcal{F} \equiv \mathcal{F}_0(t/t_0)^f$, but with different expressions and values for the exponent f and for the dimensional factor $\mathcal{F}_0$. Specifically, for GCs the expression $f = d - 2s/3 + u$ obtains from equation (3.2), while for FIs we have $f = d - s/3 - u$ from equation (3.3). The quantities d, s, and u take on different values in the field and inside the structures like groups or clusters.

In the *field* of a FRW universe the values $d = -2$ or -3 apply for values of the density parameter $\Omega_0 = 1$ or 0. As for the velocity field, the linear analysis of large-scale motions yields $u \approx \frac{1}{3}$ (see Vittorio & Turner 1987).

Within *bound* structures $\rho_a(t)$ is expected to rise sharply during the virialization stage, and to grow slowly thereafter. Meanwhile $V(t)$ undergoes only a modest increase like $t^{1/3}$ or stays nearly constant. We parameterize these behaviors with $d \approx \frac{2}{3}$ and $u \approx \frac{1}{3}$.

The internal density $\rho(t)$ is poorly known, and, if anything, it tends to decrease after a merging. We consider two limiting cases: $s \approx 0$, that is, nearly constant internal density for the interacting units, which will provide a relevant bound within structures; and $s \approx d$, that is, nearly constant contrast with the ambient, which will provide relevant bounds for the field.

In summary, our pivotal rates both scale as

$$\tau^{-1} \propto t^f M^{\lambda - 1} . \tag{3.4}$$

Physically relevant values of f and λ are as follows.

For GCs we have $\lambda = \frac{2}{3}$. In addition, in the field with three-dimensional critical expansion $f \approx -5/3$ holds when constant internal density applies, and $f \approx -\frac{1}{3}$ when constant contrast applies. Inside structures instead, with a velocity field $V \propto t^{1/3}$ we have $f \approx 5/9$ for constant density, and $f \approx 1$ for constant contrast.

With FIs, $\lambda = 4/3$ (or 3/2) applies. As to f, under the above conditions we have $f \approx -7/3$ to $-5/3$ in the field, and $f \approx$ 0–1/3 inside structures.

For a constant velocity field the value of f will decrease by one-third for GCs and increase by one-third for FIs. We collect in Table 1 pairs of λ and f which are either of direct physical relevance, or useful for comparing behaviors in different ambients.

We are now ready to compute numerically the full range of evolutions of the MD under GC and FI interactions, either in the field or inside bound systems.

TABLE 1

Cross Section	λ	f = −7/3	f = −5/3	f = −1/3	f = 0	f = 1/3	f = 5/9	f = 1
GC	2/3		Field $\rho_a \propto t^{-2}$ $\rho \sim$ const	Field $\rho_a \propto t^{-2}$ $\rho \propto \rho_a$			Bound $\rho_a \propto t^{2/3}$ $\rho \propto \rho_a$	Bound $\rho_a \propto t^{2/3}$ $\rho \sim$ const
FI ..... $r \propto M^{1/3}$	4/3	Field $\rho_a \propto t^{-2}$ $\rho \sim$ const	Field $\rho_a \propto t^{-2}$ $\rho \propto \rho_a$		Bound $\rho_a \propto t^{2/3}$ $\rho \propto \rho_a$	Bound $\rho_a \propto t^{2/3}$ $\rho \sim$ const		
FI ..... $r \propto M^{1/2}$	3/2	Field $\rho_a \propto t^{-2}$ $\rho \sim$ const				Bound $\rho_a \propto t^{2/3}$ $\rho \sim$ const		

NOTES.—Pivotal models corresponding to the values of f and λ discussed in § 3. The asymptotic behavior of $M_*(t)$ grows steeper from left to right, and from top to bottom. Phase transitions will occur when $M_*(t)$ has a vertical asymptote at a finite time, i.e., for $\lambda > 1$ and $f > -1$. A velocity field $V \propto t^{1/3}$ is implied. V = constant would yield f values decreased or increased by one-third for GCs or FIs, respectively.

4. NUMERICAL RESULTS

We solve numerically the integro-differential equation (2.1) for the relevant parameters indicated in § 3 using time-iterative integrations with a time step $\Delta t = 10^{-3}\tau$, and a mass step $\Delta M \leq M_0$, where M_0 is our mass unit. When required by numerical stability the step ΔM may go down to $\sim 10^{-2} M_0$; this implies a rescaling of the physical time intervals corresponding to the computational time step.

For galaxies in a group we use $M_0 = 10^{11}\ M_\odot$, and for subclusters in a cluster $M_0 = 5 \times 10^{12}\ M_\odot$. The numerical values for the interaction kernels are set on the basis of equations (3.1) using $t_0 = t_d$ as the physical unit of time. The initial number density of units that appears in the physical time scale will be $n_0 \sim 10$ Mpc^{-3} galaxies in a group or $n_0 \sim$ a few Mpc^{-3} subclusters in a cluster.

We consider, to begin with, initial conditions within a narrow mass range. Specifically, we shall mainly use a narrow top hat $N(M, t_i) = N_0 H(M_0, 2M_0)$ as an approximation to a monodisperse distribution, and compare results with those from the Gaussian distribution $N(M, t_i) = N_0 e^{-(M/M_0)^2/2}/(2\pi)^{1/2} M_0$. We discuss in § 8 different initial distributions.

In general, for time intervals $\sim t_0$ (decreasing for increasing values of f) the figures to follow show transients that remember the initial conditions, both as for the shape and for the evolution of the MDs. But then the MDs go into asymptotic regimes independent of initial details, with self-similar evolution.

Figures 1–3 illustrate the evolution of the MDs for the GC case from the nearly monodisperse initial condition and for $f = -\frac{1}{3}$, 0, 1, to illustrate how the behavior changes for increasing values of f. The fully developed self-similar shape of the MDs at their low-mass end is a power law with logarithmic slope $\xi \approx 1.3$, independent even of the values of f.

With time increasing, the MDs extend toward large masses, while the small-mass end lowers, thus conserving the mass involved (we check that overall numerical uncertainties concerning the total mass remain within 10^{-2} during our integrations). Such evolution is strongly dependent on f. In particular for the value $f = -\frac{1}{3}$ illustrated in Figure 1 the evolution is slowed down, but is still appreciable; we have checked with similar computations that for values $f < -\frac{2}{3}$ the evolution comes to an effective standstill after a time $\lesssim 1 t_0$, and specifically after a time $0.6 t_0$ for $f = -5/3$.

We illustrate in Figures 4, 5, and 6 the evolution under FIs with $\lambda = 4/3$, starting from the nearly monodisperse condition at different times (see captions), again for increasing values of $f = -5/3$, 0, 1/3. While the first case shows a standstill after a very short transient, the other cases (see Fig. 7) show an interesting phenomenon, namely *nonconservation* of the total mass in the system, as directly computed from $\mathcal{M} = \int dM M N(M, t)$. This occurs near and after a *finite* time t_∞ which takes on values progressively shorter for increasing f, namely, $t_\infty = 5t_0$ for $f = 0$ and $t_\infty = 2t_0$ for $f = \frac{1}{3}$. Correspondingly but independently, the shape of the MD at $t \gtrsim t_\infty$ becomes a pure power

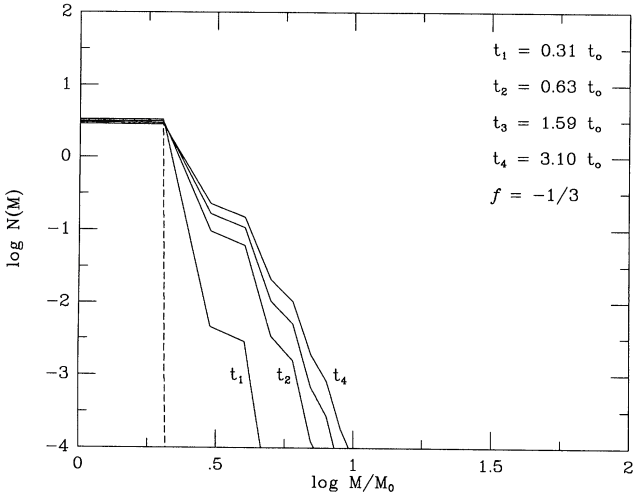

FIG. 1.—The evolution of the mass distribution under GCs ($\lambda = \frac{2}{3}$) with $f = -\frac{1}{3}$ is computed numerically from the monodisperse initial condition shown at the extreme left (*dashed line*). Our unit of mass M_0 corresponds to $10^{11}\ M_\odot$ for galaxies in a group, and to $5 \times 10^{12}\ M_\odot$ for subclusters in a cluster. The solid lines refer to the MDs evaluated at the times (*from left to right*) $t_1/t_0 = 0.31$, $t_2/t_0 = 0.63$, $t_3/t_0 = 1.59$, and $t_4/t_0 = 3.1$ (we recall that our initial time t_0 is taken equal to the dynamical time t_d).

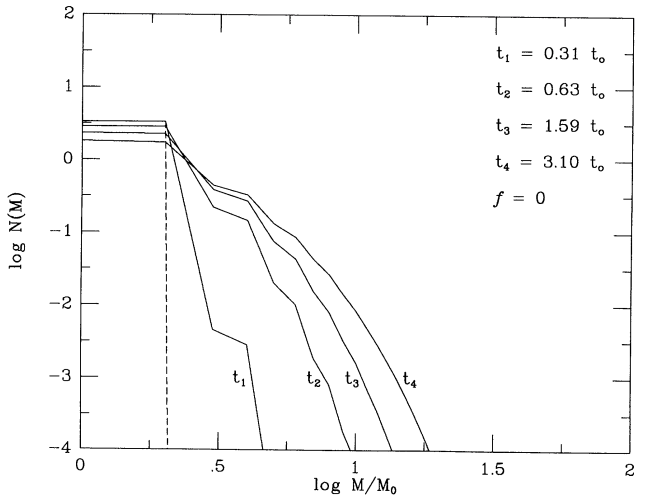

FIG. 2.—Same as Fig. 1 but with $f = 0$. Although this is not a physically relevant value, it is shown to illustrate the effect of f on the evolution.

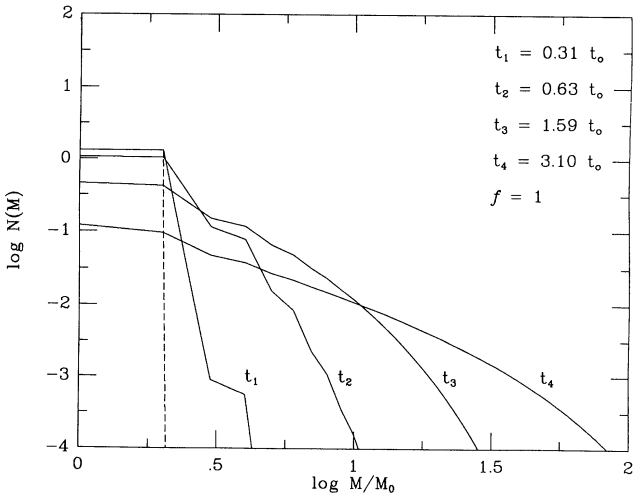

FIG. 3.—Same as Fig. 1 but with $f = 1$. Note the time scale shortened by ~ one-third.

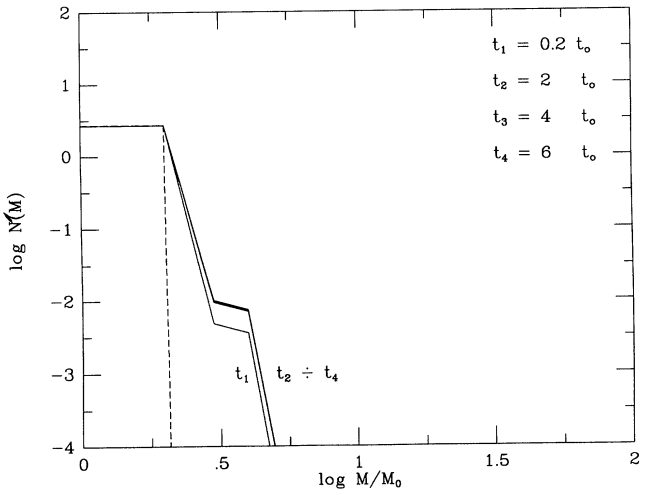

FIG. 4.—The mass distribution under FIs with $\lambda = 4/3$ and $f = -5/3$ and $f = -5/3$ is evaluated at $t_1/t_0 = 0.2$, $t_2/t_0 = 2$, $t_3 - t_0 = 4$, $t_4/t_0 = 6$. The last three curves overlap, indicating a standstill after a short transient.

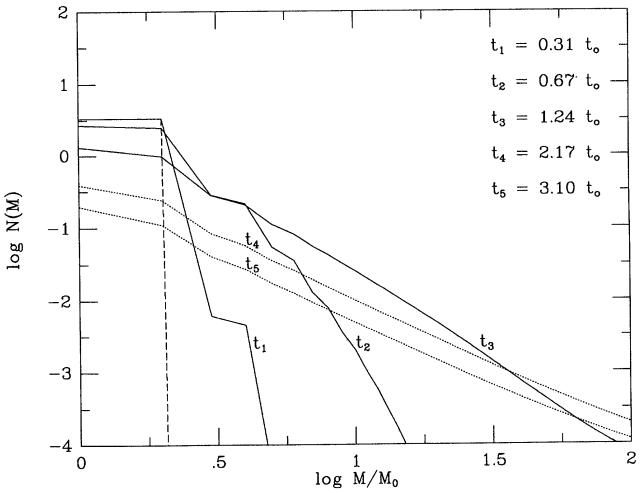

FIG. 6.—As above for $f = \frac{1}{3}$. The times now are $t_1/t_0 = 0.31$, $t_2/t_0 = 0.67$, $t_3/t_0 = 1.24$, $t_4/t_0 = 2.17$, $t_5/t_0 = 3.1$.

law with slope constant at the value ≈ 2.15; the normalization lowers, which indeed implies a decrease of the total mass in the system described. Note that the average object mass also decreases, thus inverting the canonical trend of the MDs before t_∞. As we shall see in the following, these features signal the occurrence of a phenomenon with the clear mark of a phase transition, corresponding to the formation of a merger.

The evolution of $N(M, t)$ for FIs with $\lambda = 3/2$ is shown in Figure 8. The behavior is analogous to the previous case, only the evolution is faster at given f; for example, for $f = 0$ we now find $t_\infty \approx 3.3t_0$, to be compared with $t_\infty = 5t_0$ holding for $\lambda = 4/3$.

As discussed in § 2 the actual cross section is the sum of the GC and the FI components. We give in Figure 9 the results for this realistic case. Here the GC component—with its larger value of f for given exponents d, s, u—initially drives and leads the MD evolution toward large masses, but then the FI component with its larger value of λ takes over and causes the phase transition to take place. The overall result is to shorten t_∞ relative to the case where the FI component alone was considered.

Similar results obtain from different initial conditions, in particular from a Gaussian distribution with the same initial total mass as shown in Figure 10 for $f = 0$. Starting instead from initial MDs already extending out to large masses, the conditions of the preceding paragraph apply even earlier, and again a faster transition obtains. So the evolutionary behavior is robust relative to variations of the initial conditions.

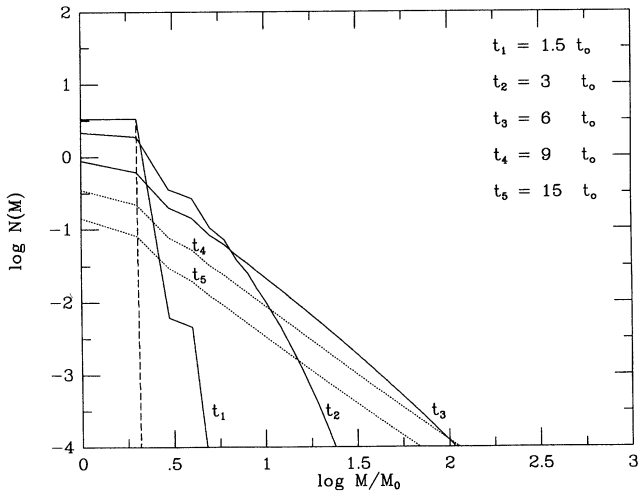

FIG. 5.—The mass distribution under FIs with $\lambda = 4/3$ and $f = 0$ is evaluated at the times $t_1/t_0 = 1.5$, $t_2/t_0 = 3$, $t_3/t_0 = 6$, $t_4/t_0 = 9$, $t_5/t_0 = 15$. The labels refer to the line at their left. The initial condition is shown at the extreme left (*dashed line*). The solid lines refer to the MDs before the transition, characterized by an exponential cutoff; the dotted lines refer to the MDs after the trasition, with the typical power-law shape. It is not represented by the merger discussed in § 6: a spike at $\mathcal{M} = \mathcal{M}_0 - M(t)$. This also applies to Figs. 6, 8, 9, and 10.

5. TIME BEHAVIORS OF THE CHARACTERISTIC MASS

In this and in the following section we look for *analytic* solutions of equation (2.1) from generic initial distributions.

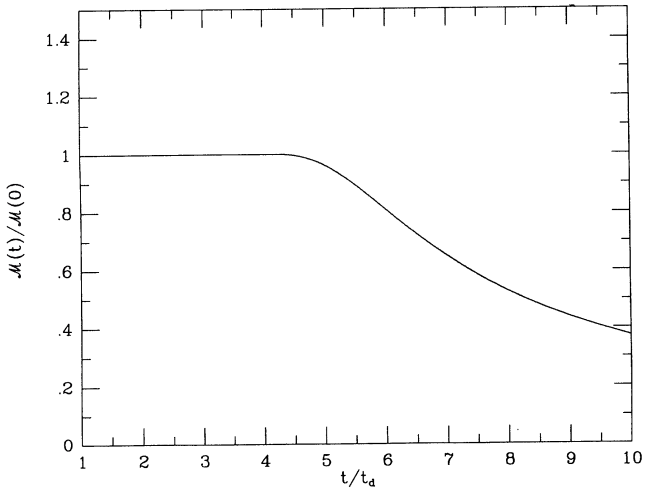

FIG. 7.—The mass of the system $\mathcal{M}$ for the FI case with $\lambda = 4/3$, and $f = 0$ is plotted as a function of time. The phase transition occurs at $t_\infty \approx 5t_0$.

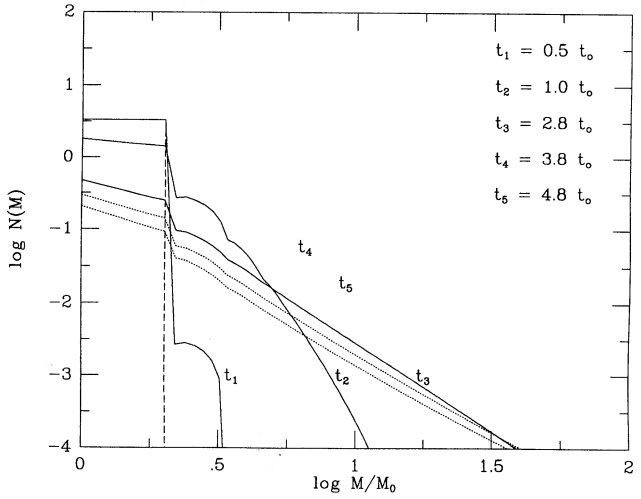

FIG. 8.—The mass distribution under FIs with $\lambda = 3/2$, computed with $f = 1/3$ for easy comparison with Fig. 5. The times are $t_1/t_0 = 0.5$, $t_2/t_0 = 1$, $t_3/t_0 = 2.8$, $t_4/t_0 = 3.8$, $t_5/t_0 = 4.8$. Again solid and dotted lines refer to the MDs before and after the transition, respectively.

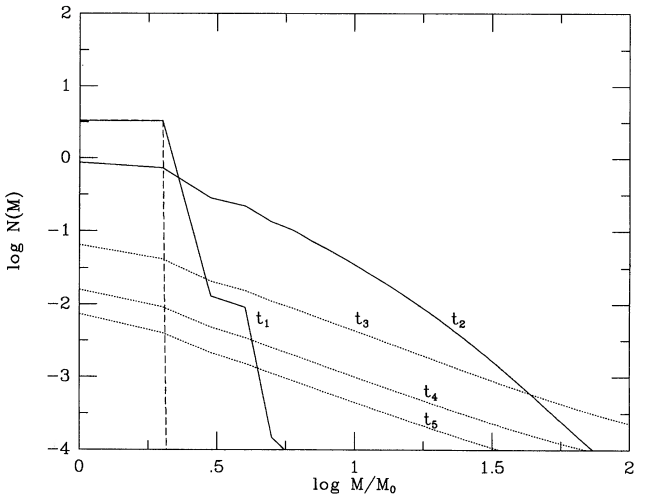

FIG. 9.—The mass distribution for the complete cross section (see § 2), with the GC and the FC components proportional to t and to $t^{1/3}$, respectively, corresponding to the same $\rho_a \propto t^{2/3}$, $V \propto t^{1/3}$, and $\rho \sim$ constant. Times are as in Fig. 8.

Then we shall compare these with the numerical results presented in the preceding section.

We look for solutions of the canonical form $N(M, t) \rightarrow M_*(t)^{-\alpha}\phi(m)$ in terms of the characteristic mass $M_*(t)$ and of the normalized mass $m = M/M_*$.[4] Such form is conveniently normalized to the total mass of the system $\mathcal{M}$, and so the previous *Ansatz* writes in full

$$N(M, t) \propto (\mathcal{M}/M_*)^{\alpha-2}\phi(m)/M_*^2 . \quad (5.1)$$

Note that when $\alpha = 2$ applies mass conservation $\int dMMN(M) = \mathcal{M}$ holds.

On substituting the *Ansatz* (5.1) in equation (2.1), the latter is separated into two *parallel* equations: an m-dependent equation for $\phi(m)$ that reads

$$\left[m\frac{d\phi(m)}{dm} + \alpha\phi(m)\right] = \phi(m)\int_0^\infty dm'\psi(m, m')\phi(m') - \frac{1}{2}\int_0^m dm'\psi(m', m - m')\phi(m')\phi(m - m') \quad (5.2)$$

to within an arbitrary separation constant that is set by our definition $t_0 = t_d$; and a t-dependent equation that reads

$$\dot{M}_* = \mathcal{M}^{\alpha-2}\mathcal{F}(t)M_*^{2+\lambda-\alpha} , \quad (5.3)$$

where $\mathcal{F}(t) = \mathcal{F}_0(t/t_0)^f$ contains the time dependence of the interaction kernel.

These two parallel equations are *connected* by the value of α, which is set by a consistency argument between the t-behavior and the m-behavior. In our serial exposition we shall consider such behaviors in turn.

The solutions of equation (5.3) for $M_*(t)$ clearly depend on the value of α used in the *Ansatz* (5.1). Let us adopt first the canonical value $\alpha = 2$ that implies canonical mass conservation. Then the solutions of equation (5.3) read

$$M_*(t) = M_{*0}\left\{1 - k\left[\left(\frac{t}{t_0}\right)^{1+f} - 1\right]\right\}^{1/(1-\lambda)} , \quad (5.4)$$

where

$$k = \frac{(\lambda - 1)}{(f + 1)}\, n_0\, \Sigma R\left(\frac{\mathcal{M}}{M_{*0}}\right)^{(1+\lambda)/2} . \quad (5.5)$$

For $f < -1$, with either $\lambda < 1$ or $\lambda > 1$, such solutions saturate after a short rise (see Fig. 11).

For $f > -1$ with $\lambda < 1$ the behavior after a slow start goes into a slanting asymptote $M_* \propto t^{(f+1)/(1-\lambda)}$ (Fig. 12). For $\lambda > 1$ instead, even an inspection of equation (5.4) obtained upon insisting on the canonical value $\alpha = 2$ will indicate a *divergence* of $M_*(t)$ at a *finite* time. We are about to show in the next

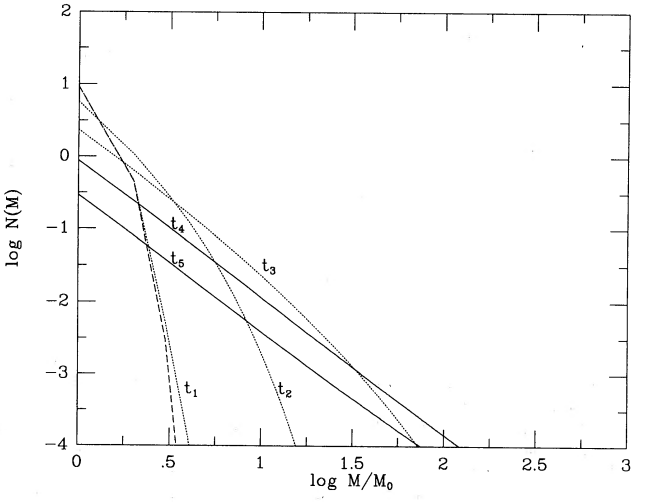

FIG. 10.—Same as Fig. 5 but with a Gaussian initial condition (*dashed line at the extreme left*).

[4] Uniqueness of the solution of eq. (2.1) has been proved for the kernels where $\psi(m, m')$ is bounded from above by $\psi_B \propto mm'$ (see McLeod 1962).

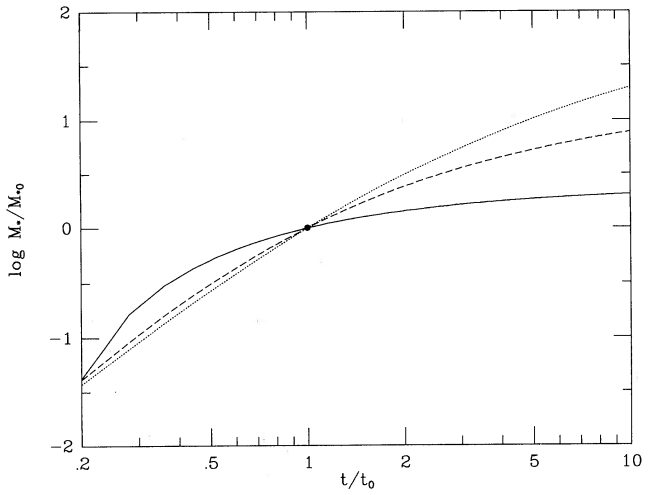

FIG. 11.—The characteristic mass M_* as a function of time with $f = -5/3$, for GCs (*solid line*), FIs with $\lambda = 4/3$ (*dashed line*), and FIs with $\lambda = 3/2$ (*dotted line*).

section that any such divergence of M_* will cause mass non-conservation, and then the appropriate value of α will be provided by a detailed analysis of the parallel equation for the shape of the MD, to read $\alpha = (\lambda + 3)/2$. Assuming this value for the time being, the strength of the actual divergence is doubled relative to a naive extrapolation of equation (5.4) to cases with $\lambda > 1$, and the consistent solution of equation (5.3) writes

$$M_*(t) = M_{*0}\left\{1 - \frac{k}{2}\left[\left(\frac{t}{t_0}\right)^{f+1} - 1\right]\right\}^{2/(1-\lambda)}, \quad (5.6)$$

with $k > 0$. Figure 13 illustrates the behavior of $M_*(t)$ for the values $\lambda = 4/3$ and $\lambda = 3/2$, and exhibits a strong formal divergence at the finite time

$$t_\infty = t_0\left(1 + \frac{2}{k}\right)^{1/(f+1)}. \quad (5.7)$$

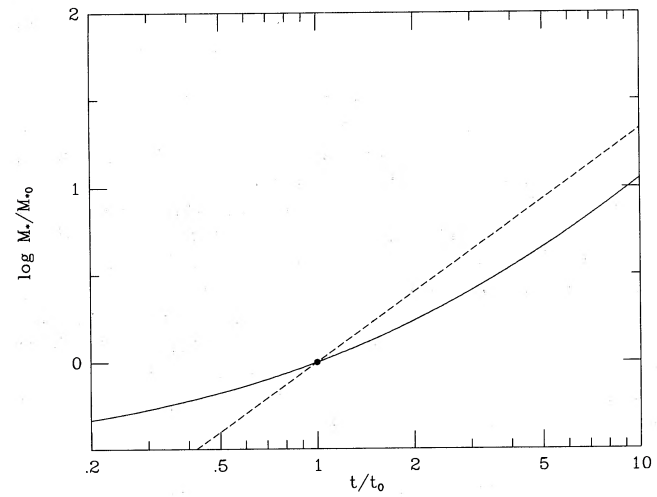

FIG. 12.—The characteristic mass M_* as a function of time for GCs in the field with $f = -\frac{1}{3}$ (*solid line*), is compared with $M_c(t)$ according to direct hierarchical clustering with a perturbation spectral index $\nu = -1.2$ (*dashed line*).

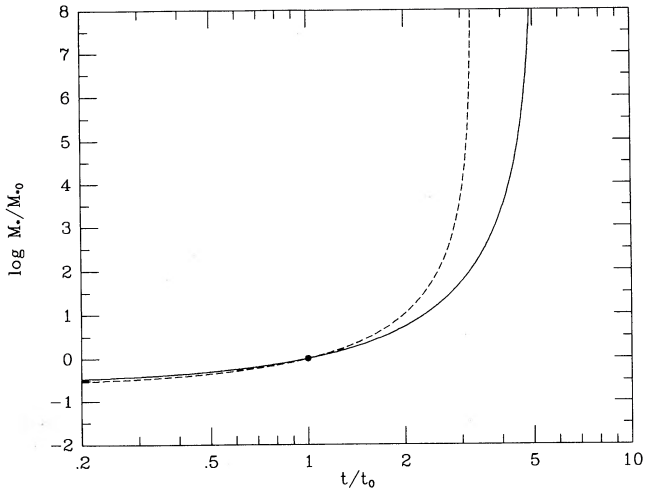

FIG. 13.—The characteristic mass M_* as a function of time for FIs with $f = 0$. The solid line refers to $\lambda = 4/3$, and the dashed one to $\lambda = 3/2$.

Specifically, when $\lambda = 4/3$ we obtain $t_\infty \approx 5t_0$ for $f = 0$, and $t_\infty \approx 2t_0$ for $f = \frac{1}{3}$.

We stress that the characteristic mass $M_*(t) = \langle M^2\rangle/\langle M\rangle$ may also be interpreted as the mass "correlation length," and in general is not associated with any specific object.

We now proceed to prove that a divergence in $M_*(t)$ causes in fact $\mathcal{M} \neq$ const to hold, and ultimately implies the value of α that was just assumed in this section.

6. ASYMPTOTIC SHAPE OF THE MD

To proceed, we need to derive m-behaviors of $\phi(m)$ from equation (5.2). These will differ radically for $\lambda < 1$ and $\lambda > 1$.

6.1. *When Mass Is Conserved*

We consider first the case $\alpha = 2$, corresponding to MDs of the form $\phi(m)/M_*^2$ that conserve the total mass of the system $\mathcal{M}$.

The asymptotic shape in the limit $m \gg 1$ obtains from the dominant terms in equation (5.2), including the construction one on the right-hand side,

$$\frac{d\phi(m)}{dm} \sim -\frac{1}{2m}\int_0^{m\to\infty} dm'\psi(m', m - m')\phi(m')\phi(m - m'). \quad (6.1)$$

To ensure convergence of the integral on the right-hand side in spite of the diverging behavior of $\psi \propto m^\lambda$, it is natural to try $\phi(m) \propto m^{-\lambda}\theta(m)$. The function $\theta(m)$ is found by direct substitution into equation (6.1). In fact, in the limit $m \gg 1$ one obtains $\theta(m) \propto e^{-m}$ so that the full asymptotic behavior reads

$$\phi(m) \sim m^{-\lambda}e^{-m}. \quad (6.2)$$

To find the behavior of $\phi(m)$ for $m \ll 1$ we try a solution

$$\phi(m) = Bm^{-\xi}, \quad (6.3)$$

and look for the values B and ξ by direct substitution in equation (5.2) to find

$$\frac{2 - \xi}{B} = \int_0^\infty dm'\psi(m, m')\phi(m') - \frac{m^\xi}{2} \times \int_0^m dm'\psi(m, m - m')\phi(m)\phi(m'). \quad (6.4)$$

The canonical procedure (Ernst 1986) would be as follows: first, infer $\xi = (\lambda + 1)$ by power counting in the above equation; then use the value so obtained in the equation to get $B = (1 - \lambda)/L(\xi)$, where

$$L(\xi) = \frac{1}{2}\int_0^1 dx\psi(x, 1-x)[x(1-x)]^{-\xi} \times [x^{2\xi-\lambda-2} + (1-x)^{2\xi-\lambda-2} - 1] . \quad (6.5)$$

However, with our specific cross sections dominated for small m by mixed terms in m and m', the integrand in $L(\xi)$ diverges for $x \to 0$ and $x \to 1$, and we have to resort to equation (6.4) in full to derive ξ. The result is

$$\xi = 2 - p_\lambda , \quad (6.6)$$

where p_λ is the moment of order λ of the function $\phi(m)$. Since ϕ is not yet entirely determined, this exponent cannot be determined analytically beyond setting the bound $\xi < 2$ from equation (6.6). For the GCs, our numerical computations give $\xi \approx 1.3$ in the relevant self-similar regime. For FIs, we shall see next that actually mass conservation does not hold, and we must start anew our quest for the asymptotic shape of the MD, which is crucial for predicting the evolution of the system.

The key tool is provided by the mass flux crossing a given, large value of M (see van Dongen & Ernst 1985) that we denote by $\hat{M}$:

$$\dot{\mathscr{M}}(\hat{M}, t) = -\int_0^{\hat{M}} dM N(M, t) M \int_{\hat{M}-M}^{\infty} dM' N(M', t) K(M, M', t) . \quad (6.7)$$

This expression is derived by taking the first moment of equation (2.1) with suitable interchanges of variables that do not require a cutoff in the MD. Substituting the *Ansatz* (5.1) one obtains a *bistable* behavior of $\dot{\mathscr{M}}(\hat{M}, t)$, depending on the shape $\phi(m)$ of the MD.

When $\phi(m)$ is of the canonical type in equation (6.2) with an upper cutoff, the mass flux satisfies $\lim_{\hat{M}\to\infty} \dot{\mathscr{M}}(\hat{M}, t) \to 0$, implying conservation of the total mass $\mathscr{M}$. So mass conservation, a cutoff for $N(M)$, and $M_*(t)$ finite at any finite time constitute a set consistent with $\lambda < 1$.

6.2. *When Mass Is Not Conserved*

On the other hand, a remarkable phenomenon arises when $m \ll 1$ applies even as M is large, which formally requires a divergence of M_* at a finite time. Indeed, we already know from § 5 that such a divergence is to occur when $\lambda > 1$ (and $f > -1$). Then the asymptotic shape of the MD takes on the form stretched out as given by equation (6.3) also for *large* absolute values of M, and equation (6.7) yields the limiting mass flux

$$\lim_{\hat{M}\to\infty} \dot{\mathscr{M}}(\hat{M}, t) \propto -\hat{M}^{3+\lambda-2\xi} \neq 0 . \quad (6.8)$$

But a finite value for $\dot{\mathscr{M}}$ is to be required for a successful *time-resolved* description in terms of a *single* time scale (infinite values, equivalent to discontinuities of physical conditions on the basic scale, would require a second and finer time resolution). In fact, the condition for $\dot{\mathscr{M}}(\hat{M}, t)$ to remain *finite* reads simply

$$\xi = (\lambda + 3)/2 . \quad (6.9)$$

Furthermore, just because $\mathscr{M}$ is not conserved, now we cannot retain the canonical value $\alpha = 2$ corresponding to mass conservation. The right choice of α is obtained by direct substitution of the *Ansatz* (5.1) into equation (5.2) to yield

$$B = (\xi - \alpha)/L(\xi) . \quad (6.10)$$

If $\xi = (\lambda + 3)/2$ is to hold, then $L(\xi) = 0$ follows (see last factor in the integrand [6.5]). But then a finite B requires $\alpha = \xi$.

Beyond the critical point, the condition (6.9) must still hold if the mass flux is to remain finite.

Note, as anticipated at the end of § 5, that the equalities

$$\alpha = \xi = (\lambda + 3)/2 \quad (6.11)$$

also fix the strength of the divergence of $M_*(t)$ that occurs for $\lambda > 1$. So the formal argument closes up, having made use of these consistency relationships between the parallel t-behavior and m-behavior.

The meaning of $\dot{\mathscr{M}} < 0$ is as follows. A divergence of $M_*(t)$ implies the tail of $N(M)$ to be stretched out to large values of M. So the few objects in the tail will have effective merging times τ quite below the typical value for most members, and will coalesce soon into a single large *merger* with a time scale shorter yet. As the Smoluchowski equation in the form (2.1) cannot describe kinetics with widely diverse time scales, the subsystem constituted by the merger is no longer described by the same equation as applies to the *normal* objects on the scale given by equation (3.1b) with the typical mass suited to the body of the distribution. The detailed kinetics of the merger interactions as seen by the normal members is replaced by a boundary condition on the M axis, that is, by the integrated outcome of the interactions constituted by the mass loss $\dot{\mathscr{M}} < 0$. Such a loss is consistent with the destruction term being strongly dominant for the normal members (cf. equation [5.2] with $m \ll 1$), which in turn produces the power-law shape (6.3) of $N(M)$. All this constitutes a valid description on the time scale of the normal galaxies provided that $\dot{\mathscr{M}}$ be finite, which can be met as shown by equation (6.9). Overall consistency with global mass conservation obtains as the mass leaving the subsystem still described by equation (2.1) flows into the merger.

Thus the system is to break into a bimodal distribution: (1) a first phase constituted by the normal members still following equation (2.1), therefore distributed after equation (6.4) and loosing mass at a rate $\dot{\mathscr{M}} < 0$; and (2) a second phase constituted by the forming merger accreting mass at a rate $|\dot{\mathscr{M}}|$, that is, with a δ-like mass distribution centered on the value $\mathscr{M}_0 - \mathscr{M}(t)$.

In short, the set of propositions consistent with $\lambda > 1$ (and $f > -1$) comprises $M_*(t)$ diverging at a finite t; $\dot{\mathscr{M}} < 0$; and $N(M)$ splitting into a power-law excess of small masses and a δ-like spike at a high and increasing mass.

Such splitting and remolding of the MD provides, for example, a description for the formation of a large, dominant cD-like galaxy in groups (Cavaliere et al. 1991). In a broader perspective, it constitutes a *gravitational phase transition.* The divergence of M_* at a finite time t_∞, necessary to have $\dot{\mathscr{M}}(\hat{M}, t) \neq 0$ and finite, plays a role analogous to the divergence of the correlation length in theories of critical phenomena (see Domb & Lebowitz 1988). The appropriate order parameter is given by $1 - \mathscr{M}/\mathscr{M}_0$.

7. COMPARISON OF ANALYTIC AND NUMERICAL RESULTS

Here we compare the numerical results given in § 4 with the asymptotic behaviors of the MDs given in §§ 5 and 6.

The mass nonconservation and the lowering power-law MDs noted for $t > t_\infty$ in the numerical results for FIs ($\lambda = 4/3$) with $f = 0, \frac{1}{3}$ (Figs. 4, 5) correspond to the divergence of $M_*(t)$ appearing in equation (5.6) for $\lambda > 1$ and $f > -1$, and to the mass flux $\dot{\mathcal{M}} \neq 0$ in equation (6.8).

The slope of the postmerger distribution (i.e., at $t \geq t_\infty$) as computed numerically from Figure 5 is $\xi \approx 2.15$, and agrees to within 1% with the value $(\lambda + 3)/2 = 13/6$ evaluated analytically in § 5 at and beyond the critical time.

As for GCs with $-\frac{2}{3} \leq f \leq \frac{2}{3}$, the numerical results (Figs. 1–3) show a time change in the MDs toward a self-similar shape consistent with equations (6.2) and (6.3). In terms of M and t, the self-similar distributions $N(N, t) \propto M_*^{-2}\phi(M/M_*)$ evolve at the rate given by $M_*(t)$ in equation (5.4) with $\lambda < 1$. The slope of the power-law section is ≈ 1.3 (see Fig. 3), the one result that cannot be predicted analytically for reasons discussed in § 6. By way of contrast, the MDs in Figure 1 are not yet out of the transient regime, due to the low value of f used there.

For both GCs and FIs a faster evolution numerically obtains with increased values of f (cf. Figs. 1–3, 5–6), a fact explained analytically in terms of a larger exponent in the expressions for $M_*(t)$ given by equations 5.4–5.6. Furthermore, the standstill of the MD observed—after a short transient—in the numerical solutions for $f < -1$ (Fig. 4) reflects the saturation of the $M_*(t)$ shown by equations (5.4) and (5.6).

We also computed numerical solutions for the value $\lambda = 1$, a critical value where several of the asymptotic expressions (see, e.g., § 5) lose meaning. We find behaviors qualitatively similar to those for $\lambda > 1$ at equal values of f.

A final remark is that the numerical computations solve, in fact, the discrete counterpart of equation (2.1) for the concentration N_k of the mass $M_k = kM_0$,

$$\dot{N}_k = \tfrac{1}{2} \sum_{i+j=k} K(i, j) N_i N_j - N_k \sum_{j=1}^{j_{\max}} K(k, j) N_j , \qquad (7.1)$$

where summations over a finite number of members replace the integrals. The agreement of the numerical solutions with the analytical results from equation (2.1) proves the equivalence of the two formulations.

8. SUMMARY AND DISCUSSION

8.1. *Summary*

We propose a pattern that captures essential aspects of the complex phenomena of merging between high-contrast members of cosmic structures. Specifically, we consider galaxies in groups or in the field, and subclusters in forming clusters.

Starting from the aggregation kinetic equation of Smoluchowski, we derive *numerical* and *analytical* solutions, that agree quantitatively; compare for example, the key results in Figures 5 or 6, 7, and 13 with those in equations (6.3), (6.8), and (5.6). Both kinds concur with a long sequence of simulations of galaxy merging started by Toomre & Toomre (1972), and taken up for galaxies in groups by, for example Aarseth & Fall (1980), CCS (1981), Barnes (1989). They concur also with simulations of cluster formation by Cavaliere et al. (1986), West et al. (1988), and Carlberg & Couchman (1989).

The general message is that in finite, self-gravitating systems the component condensations are sensitive to the action of gravitational aggregations. Shape and evolution of the member MD tend to become *independent* of initial conditions. The evolution may undergo a merging *runaway* capable of remolding the MD over a few dynamical times in systems with longer lifetime and limited membership. In galaxy groups such processes correspond to aggregation of member galaxies into a large cD-like merger; in clusters they correspond to aggregation of subclumps into a relaxed cluster configuration. These processes bear the marks of a critical phenomenon, and specifically of a gravitational *phase transition.* Aggregation, as a two-body process, is sensitive to the cosmic density decrease and hence would not sustain a lasting growth of structures in the "open" field.

Quantitatively, a positive feedback loop is started yielding shorter and shorter time scales $\tau \propto t^{-f}M^{1-\lambda}$, when aggregations take place by focused interactions (FIs) of condensations with a nonlinear cross section ($\lambda > 1$), and with a number density not strongly decreasing ($f > -1$). As a result, the a merging *runaway* develops, and on the scale of a few dynamical times t_d leads to rapid growth of a large merger at the expenses of the normal system members, whose original total mass $\mathcal{M}_0$ is *not* conserved. Specifically, the following phenomena occur: the characteristic mass $M_*(t) \propto (t_\infty^{f+1} - t^{f+1})^{-2/(\lambda-1)}$ formally *diverges* in a finite time; the shape of the mass distribution $N(M, t)$ corresponding to the normal galaxies tend to be stretched out into a scale-free *power law* $B(t)$ $M^{-(\lambda+3)/2}$, while a *second* δ-like component develops, corresponding to the formation and growth of a single merger; a finite mass flux $\dot{\mathcal{M}} < 0$ transfers mass from the first to the second component of this *bimodal* distribution; and the flux is related to the lowering amplitude of the power-law component by $B(t) \propto |\dot{\mathcal{M}}|^{1/2}$.

8.2. *Analytic Technique*

Since they constitute a pattern of specific and general relevance, we summarize our main technical points in the form of three consistency arguments, complemented by three stability checks.

1. A consistent description in terms of a single time scale.

Our main thrust is toward a description of $\partial N/\partial t$ from aggregations, resolved on a *single* time scale. But when $M_*(t)$ diverges at a finite time, $N(M)$ tends to be stretched out into a power law toward large values of M, so that the system as a whole would develop a range of time scales (in a naive approximation $\tau \sim 1/n\overline{\Sigma V} \propto t^{-f}M^{1-\lambda}$) extending toward zero. This is too wide a dynamic range to be described by the Smoluchowski equation on the time scale characteristic for the body of the MD.

Instabilities of numerical computations at high masses stress the point dramatically and to be controlled require finer meshes in M and t, leading to very long computing times.

On the analytic side, the largest masses corresponding to the shortest merging times are conveniently separated out, and replaced by a boundary condition describing their integrated action onto the residual "normal" galaxies. Such action corresponds to the mass flux $\dot{\mathcal{M}} < 0$ toward the merger, related to dominance of the destruction term in the equation for the normal subsystem.

A *finite* value for $\dot{\mathcal{M}}$ is to be required for a *time-resolved* description in terms of a *single* time scale.

The whole scheme meets success since the requirement of such a finite $\dot{\mathcal{M}} < 0$ can be indeed enforced by the composite, but ultimately simple, consistency argument summarized below.

2. Consistent t- and M-behaviors.

The original Smoluchowski equation with the *Ansatz* $N(M, t) \propto M_*^{-\alpha}(t)\phi(M/M_*)$ separates into a t-dependent and

an M-dependent equation, related by the value of the parameter α that both contain.

The M-equation is solved asymptotically for large and for small values of M/M_*. In the latter regime power-law behaviors are found, corresponding to dominance of the destruction term.

On the other hand, when $\lambda > 1$ holds (with $f > -1$), the t-equation yields a diverging $M_*(t) \propto (t_\infty^{f+1} - t^{f+1})^{1/(1+\lambda-\alpha)}$. But then asymptotic form for $M/M_* \ll 1$ applies even in the body of the M-distribution $\phi(M/M_*)$. The ensuing solution $\phi \to (M/M_*)^{-\alpha}$ corresponds to unbalanced destruction. This implies mass loss from the body of the MD, at a rate $\dot{\mathcal{M}} \propto -M^{3+\lambda-2\alpha}$.

A finite rate requires $\alpha = (\lambda + 3)/2$. At a single stroke, this determines both the true strength of the divergence in $M_*(t)$ and the slope of $N(M, t)$ at the critical point, and actually beyond.

3. Consistency with global mass conservation.

This obtains as the mass leaving the subsystem still described by equation (2.1) flows into the merger. Thus the system is to break into a bimodal distribution.

4. Stability.

Our description is *stable* relative to different values of the finite system mass $\mathcal{M}$. This is shown—once the condition $\alpha = (\lambda + 3)/2$ is satisfied—by the full, finite-size version of equation (6.8) which reads $\dot{\mathcal{M}} \propto -M^{3+\lambda-2\alpha}[1 + O(M_\ell/\mathcal{M}_0)^{(\lambda-1)/2}]$.

The evolutionary behavior is also robust relative to variations of the initial conditions. Indeed, widely dispersed initial conditions accelerate the onset of the critical phenomenon.

Finally, we discuss below in § 8.5 how the critical phenomenon is actually accelerated when a realistic cross section is considered, including a geometric and a focusing term.

8.3. *Gravitational Phase Transitions*

The merging runaway gone to near completion constitutes a neat instance of a cosmic *phase transition* of a gravitational nature in the relatively nearby universe. In the perspective provided by theories of critical phenomena, the characteristic value of the mass $M_* = \langle M^2 \rangle / \langle M \rangle$ plays on the mass axis a role analogous to the correlation length in ordinary space, whose divergence is indicative of organized behavior embracing the system. At the divergence $N(M, t)$ becomes scale-free, and instead $\mathcal{M}(t)$ takes over in the role of the physically relevant quantity; in fact, the order parameter appropriate to the present transition is constituted by $1 - \mathcal{M}(t)/\mathcal{M}_0$. The gravitational free energy $\propto N^2$, so tat the critical phenomenon constitutes a *second-order* transition.

While the qualitative outline of the phase transition is provided by the mean field equation (2.1) in dilute approximation, it will constitute matter of further study—in analogy with the development of other theories of critical phenomena (cf. Domb & Lebowitz 1988)—to find how the quantitative values of the critical indexes may be modified by a proper account of realistic and growing correlations. Here we note that a condition for equation (2.1) to apply, namely no pairwise correlations of initial velocities, implies $(\mathcal{M}/M)^{2/3} \gg 1$. When aggregation operates among a few units, the initial velocities may instead be correlated, and equation (2.1) may not apply literally. But then (at given angular momentum) the equation provides an upper bound to the time for aggregation to occur.

The gravitational transitions are closely analogous to the phase transition sol to gel ("gelation") occurring in suspensions of aggregating particles, which is represented by a closely similar, but in fact simpler, aggregation equation with the integro-differential structure (2.1) as reviewed by Ernst (1986). We have, in fact, two additional structural features: the first is constituted by the finiteness of our gravitational systems; the second feature is constituted by the explicit time dependence of the kernel, see equation (3.4), mainly related to the drifting density of colliding condensations.

8.4. *Environmental Effects*

The two-body aggregations are clearly sensitive to the t-dependence of the ambient density, an effect directly embodied into the time scale $\tau \propto t^{-f} M^{1-\lambda}$. Bound systems provide quite a different density run from the field. In addition, in finite systems the velocity dispersions entering the cross section depend on relative sizes.

8.5. *Bound Systems*

In self-gravitating, finite systems the relevant time scale is provided by the internal time t_d set at system collapse, multiplied by a factor governed by the mass ratio $\mathcal{M}/M_*$, as specified by the expression for t_∞ in equations (5.7) and (5.5). Continued evolution obtains as the ambient density is increasing or constant, such as to yield $f > -1$. With GCs alone, the negative feedback by the growing aggregates (expressed by eq. [3.4] with $\lambda < 1$) holds back the evolution and causes the MDs to settle to a mild, self-similar regime. Instead, with FIs important the condition $\lambda > 1$ holds, the feedback is positive and drives a runaway ending up in the phase transition.

The FI model discussed in the text uses the simple scaling $r \propto (M/\rho)^{1/3}$, and yields a cross section scaling as $M^{4/3}$. Alternatively, the empirical Faber-Jackson relation $L \propto v^4$ (Faber 1982) presumably describes the specific galactic structure, and with the constraint $M/L \sim$ const yields the scaling $M^{3/2}$, independently of ρ. As long as this holds, not only the runaway takes place because the key requirement $\lambda > 1$ is satisfied, but in fact it is even faster, as shown by equations (5.4) and (5.6). The quantitative effect is to decrease t_∞ by a factor ~ 1.5, from $\approx 5t_d$ to $\approx 3.3t_d$ for $f = 0$.

The transition from GCs to FIs is marked by values $v^2/V^2 \gtrsim \frac{1}{3}$. In bound systems $v^2/V^2 \sim R/r\mathcal{N}$ holds, and the condition $v^2/V^2 \gtrsim \frac{1}{3}$ for FIs to prevail implies $\mathcal{N} \lesssim 30$. Even when this condition is not strictly satisfied, and aggregations are governed by GCs, the stronger M-dependence of the FI component will cause it to prevail eventually, as shown in Figure 9. A similar behavior also occurs with the time dependence of the cross sections. Given behaviors of $\rho(t)$, $\rho_a(t)$, and $V(t)$ combine to yield for FIs values of $f \sim \frac{1}{3}$, smaller than the corresponding values for GCs. In such conditions, the stronger t-dependence of GCs drives the initial evolution of $M_*(t)$ so that $N(M)$ extends toward large values of M. Then the stronger M-dependence of FIs takes over and is responsible for the final transition. Widely dispersed initial conditions will only speed up the prevalence of FIs.

Considering a wider range of V/v values, *two* extreme regimes may be envisaged: for $V/v \sim 1$, resonant interactions end up to dominate and start the runaway, as said above; when instead $V^2/v^2 \gg 1$ holds initially, the evolution by the GC process is self-similar and slow, liable to an early termination by inclusion of the system into a still larger cluster.

One condition that tends to enhance V^2 at given luminous mass is the presence of diffuse dark matter (as distinct from individual dark galaxian halos); then the merging times tend to be stretched out. Actually, the situation with dark mass is complex, because the former effect competes with increased dynamical friction. On the other hand, individual halos may be stripped and shed by encounters, and by tidal forces from the overall potential. Such interplay of three effects warrants simulations aimed to cover relevant ranges of galaxian and of diffuse M/L.

In rich clusters the velocity dispersion is large enough to suppress the direct mergings, making it difficult to build up a cD body inside rich clusters by this process (CCS 1981; Merrit 1983; Richstone & Malumuth 1983; Bothun & Schombert 1988). In such environments, a slower form of merging may prevail (see Richstone 1990); first, dynamical friction segregates the galaxies to the center, then merging or cannibalism take place.

We note other physical parameters opposing direct galaxy merging: (1) Crowded systems, implying interference of interlopers with the two-body collisions and slingshot effects (Saslaw 1985; Mamon 1990): but high average density and large membership do not occur together in cosmic structures. (2) A large systemic angular momentum: this is not statistically expected in many-body systems from the action of external tides and of aggregations (see Saslaw 1985), while selected initial conditions (see Governato, Bathia, & Chincarini 1991) can be embodied in our dscription under the reduced form of a small η multiplying $\overline{\Sigma V}$. (3) Strong radial dependence of V in the potential well: but our model theory still describes the evolution of the central region, while the halo contributes a few late comer components with high velocities that may survive several collisions.

Even under favorable conditions the runaway is eventually stabilized as resonant interactions are quenched by the decrease in number and size of the surviving galaxies. These become small also because the least bound external regions are peeled off by grazing interactions. The residual interactions with the merger are best described in terms of dynamical friction, with asymmetrical cross sections (see Alladin et al. 1988), of satellites with large angular momenta.

With all the caveats said, the *reality* of runaways gone to near completion is supported not only by the results from the sequence of simulations referred to in the Introduction, but also by real observations of such groups dominated by a cD-like galaxy (e.g., MKW11, AWM4, AWM7) as cataloged by Morgan, Kayser, & White (1975) and by Albert, White, & Morgan (1977). Indirect evidence is provided by X-ray emission from groups (cf. Schwartz, Schwarz, & Tucker 1980; Biermann, Kronberg, & Madore 1982; Kriss, Cioffi, & Canizares 1983), since extensive merging action induces shrinking of the overall configuration so increasing the density of the intergalactic gas. Thus the bremsstrahlung emissivity is boosted, and sustained over several dynamical times.

The observed low frequency of such cD-dominated groups is the result of competing time scales. The pure merging runaway takes a time around $3t_d$. This is longer than the average time needed for a reshuffling of the group into a larger unit, an event that is likely to brake or stop the process. The percentage of resulting systems containing cD or cD-like galaxies is around 7%, as for the process here described.

In rich clusters, instead, merging occurs between subclusters and groups falling together at formation. The phase transition appears as rapid erasure of substructure during the collapse and virialization of the overall structure (see also White & Rees 1978). Such processes at large scales may restart merging of the cD or the cD-like galaxies already formed in subclusters and carried with them into larger condensations (Ostriker & Tremaine 1975; Edge 1991). If so, the luminosity of the brightest cluster members must correlate in detail with Bautz-Morgan morphologies of the host clusters; on average, it may *increase* with cosmic epoch as long as the merging rate of the host systems overcomes decay of the stellar populations.

Technically, the present formalism holds as it stands when many subclusters aggregate to form a large cluster, a condition prevailing when the spectral index of the initial perturbations $v \gtrsim -1$ (see CC 1990). Because r/R is relatively larger in this case, the time scale $\tau \sim t_d/n_0 \Sigma R$ is relatively shorter compared with galaxies in groups, $t_\infty = 2\text{–}2.5t_d$ for $f = 0\text{–}\frac{1}{3}$ and with $n_0 = 5$ Mpc^{-3}. In addition, when only a few subclusters aggregate, the correlations may shorten the scale further. Because rich clusters are young systems forming around the present epoch, the process is likely to be observed in action in a fair fraction of such structures, as in fact is shown by the optical and especially by the X-ray evidence mentioned in the Introduction. Once again, N-body simulations, as discussed by CC (1990), show in detail the course and the statistics of the process.

8.6 *The Field*

In the field the role of merging interactions is very different from that in self-gravitating systems. In the canonical, "open" field of a FRW universe the strong decrease of the ambient density, $\rho_a \propto t^{-2}$ or faster, implies $f \lesssim -1$ which drastically reduces the efficiency of merging interactions like all two-body interactions. Then the evolution driven by either FIs or GCs is braked to a standstill, as indicated by the numerical results of § 4 and by the analytical results of §§ 5 and 6.

With FIs, although the mass dependence is stronger, given behaviors of $\rho(t)$, $\rho_a(t)$, $V(t)$ combine to yield $f < -1$, capable of freezing the evolution anyway. In sum, in a three-dimensional field of a FRW universe, the decreasing ambient density inhibits merging runaways. The GC component alone would produce appreciable self-similar evolution in the field under the unlikely condition $f \gtrsim -\frac{1}{3}$ (see Table 1).

Even with pure GCs, our solutions of equation (2.1) *differ* sharply from the fully self-similar solutions (see Silk & White 1978) forced to evolve on a cosmic time scale by the assumed condition $\tau \propto t$. In a heuristic look, such a condition might be induced in the field of a critical universe by $\rho_a \propto t^{-2}$, since the condition $\tau \sim 1/(G\rho_a)^{1/2} \to t$ may in principle apply. However, for this to hold quantitatively quite some fine tuning must be enforced for two quantities of different stand and origin: $\lambda = 1$, associated with the nature of the two-body interactions; *and* $f = -1$, reflecting the ambient conditions. These two equalities are to hold together to yield $t^{(f+1)/(\lambda-1)} \propto t$. In an open universe the even stronger expansion eventually freezes the evolution anyway, because $\tau/t \to \infty$.

The competition between merging interactions and direct collapses may be discussed in terms of a comparison of equation (2.1) with the other equation

$$\frac{\partial N}{\partial t} = \frac{N}{\tau_+} - \frac{N}{\tau_-}, \tag{8.1}$$

that describes DHC theories; $\tau_- = 3t/2$ and $\tau_+ = \tau_- m^{-2a}$ apply in the simplest instance (CCS 1991).

The two equations have analogies, including the structure of the destruction term and the existence of self-similar regimes. However, they differ radically in the memory of the initial conditions they retain: limited to a short transient for the aggregation process, and persistent at all epochs in the DHC process. In fact, in the latter case the solutions of equation (8.1) are always self-similar, with the characteristic mass $M_c(t) \propto t^{4/(\nu+3)}$. For $\nu < 0$ this is faster than its counterpart $M_*(t) \propto t^{(f+1)/(\lambda-1)}$ from self-similar aggregations (see Fig. 12) indicating better efficiency of DHCs in building up new structures in the field. In addition, the slope of DHC solutions are steeper than the aggregation solutions in their self-similar regime, indicating dominance of the former also at low masses.

Actually, the two processes play different and *complementary* roles in structure formation. Initially the DHCs set up forerunner modulations of the density field on the overall scale of the collapse; then the aggregations become effective within such forming structures. The share of DHCs grows for $\nu \lesssim -1$ (the perturbation field is dominated by relatively few, accreting peaks), that of aggregations for $\nu \gtrsim 0$ (interactions between equals become more likely).

We mention (from Cavaliere, Colafrancesco, & Menci, in preparation) that pure merging processes *in the field* have a hard time explaining the rapid anti-evolution of clusters luminosities in X-rays at $z \gtrsim 0.2$ observed by Gioia et al. (1990), Edge et al. (1990), and Henry et al. (1991). As said, field aggregations cannot sustain a fast and lasting growth of structures. In their self-similar regime they are too slow to match the DHCs with $\nu < -1$, even considering that for the X-ray luminosities $L \propto M^{2.5}$ may hold in this context (Edge et al. 1990), ao that $L_c(z)$ is accelerated relative to $M_c(z)$. The initial, short transient regime may be construed to be fast by ignoring the kernel dependence on time and adjusting the efficiency. But then it would produce an almost equally strong anti-evolution in the optical band at $z \sim 0.5$, for which to now there is little, or rather contrary evidence (Dressler & Gunn 1988; Gunn 1990; Ellis 1991).

8.7. *Low-Dimensionality Structures*

On the other hand, the role of merging is enhanced not only in truly bound structures, but also in all environments where the density does not decrease strongly. As an extreme case, merging phenomena would be widespread and overwhelming, at an ever increasing rate, in the recollapse stage of a supercritical universe.

Even when the average $\Omega \leq 1$ holds, a homogeneous isotropic field expanding like $\rho_a \propto t^{-d}$ with $d = 2$–3 may be a myth after all. Deep redshift surveys (e.g., Sutherland 1988; Ramella, Geller, & Huchra 1989), large-scale simulations (e.g., Efstathiou et al. 1988; Villumsen 1989; Carlberg & Couchman 1989), and quasi-linear analyses (see Shandrin & Zel'dovich 1989; V. Lukash 1991; private communication) all concur in outlining a universe tessellated by voids surrounded by precursor ridges and filaments. Such low-dimensionality structures looming out at larger z in scaled form may provide especially at their intersections "protected" sites with enhanced contrast, reduced expansion, and nonradial velocities, where merging activity is likely to be triggered and sustained both for galaxies and for groups and clusters.

Such milder or incomplete, but widespread variants of the aggregation activity may be relevant (Cavaliere, Colafrancesco & Menci, in preparation) both for X-ray cluster anti-evolution with increasing z, and also for positive number evolution that may produce the excess of faint galaxy counts discussed by Tyson (1988), Cowie et al. (1990), Koo (1990), and Guiderdoni & Rocca-Volmerange (1990).

We are grateful for helpful discussions to F. Lucchin, S. Matarrese, M. Vietri, and W. C. Saslaw. Thanks are due to the referee for his helpful and stimulating comments that induced considerable improvements in the presentation. We acknowledge grants from ASI and MURST.

REFERENCES

Aarseth, S. J., & Fall, S. M. 1980, ApJ, 236, 43
Albert, C. E., White, R. A., & Morgan, W. W. 1977, ApJ, 211, 309
Alladin, S. M., Narasimham, K. S. V. S., & Ballabh, G. M. 1988, in The Few Body Problem, ed. M. J. Valtonen (Dordrecht: Kluwer), 327
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Barnes, J. E. 1989, Nature, 338, 123
Biermann, P., Kronberg, P. P., & Madore, B. F. 1982, ApJ, 256, L37
Binggeli, B., Tammann, G. A., & Sandage, A. 1987, AJ, 94, 251
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, in preparation
Bothun, G. D., & Schombert, J. M. 1988, ApJ, 335, 617
Briel, U. G., et al. 1991, A&A, 246, L10
Carlberg, R. G., & Couchman, H. M. P. 1989, ApJ, 340, 47
Carnevali, P., Cavaliere, A., & Santangelo, P. 1981, ApJ, 249, 449 (CCS 1981)
Cavaliere, A., & Colafrancesco, S. 1990, in Clusters of Galaxies, ed. W. R. Oegerle, M. J. Fitchett, & L. Danly (STScI Symposium Series) (Cambridge: Cambridge Univ. Press), 43 (CC 1990)
Cavaliere, A., Colafrancesco, S., & Menci, N. 1991, ApJ, 376, L37
Cavaliere, A., Colafrancesco, S., & Scaramella, R. 1991, ApJ, 380, 15 (CCS 1991)
Cavaliere, A., Santangelo, P., Tarquini, G., & Vittorio, N. 1986, ApJ, 305, 651
Colafrancesco, S., Lucchin, F., & Matarrese, S. 1989, ApJ, 341, 3
Cowie, L. L., Gardner, J. P., Lilly, S. J., & McLean, I. 1990, ApJ, 360, L1
Domb, C., & Lebowitz, J. L., ed. 1988, Phase Transition and Critical Phenomena (London: Academic)
Doroshkevich, A. G. 1970, Astrofizika, 6, 320
Dressler, A., & Gunn, J. I. 1988, in IAU Symp. 130, The Large Scale Structure of the Universe, ed. J. Audouze (Dordrecht: Reidel), 311
Dressler, A., & Shectman, S. A. 1988, AJ, 95, 985
Edge, A. C. 1991, MNRAS, 250, 103
Edge, A. C., Stewart, G. C., Fabian, A. C., & Arnaud, K. A. 1990, MNRAS, 245, 559
Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, MNRAS, 235, 715
Ellis, R. G. 1991, in Proc. of the NATO ASI Clusters and Superclusters of Galaxies, in press
Ernst, M. H. 1986, in Fractals in Physics, ed. L. Pietronero & E. Tosatti (New York: Elsevier), 289
Faber, S. M. 1982, in Astrophysical Cosmology, ed. H. A. Brück, G. V. Coyne, & M. S. Longair (Città del Vaticano: Pontificia Academia Scientiarum), 191
Forman, W., & Jones, J. C. 1990, in Clusters of Galaxies, ed. W. R. Oegerle, M. J. Fitchett, & L. Danly (STScI Symposium Series) (Cambridge: Cambridge Univ. Press), 257
Geller, M. J., & Beers, T. C. 1982, PASP, 94, 421
Gioia, I. M., Henry, J. P., Maccacaro, T., Morris, S. L., Stocke, J. T., & Wolter, A. 1990, ApJ, 356, L35
Governato, F., Bathia, R., & Chincarini, G. 1991, AJ, 371, L15
Guiderdoni, B., & Rocca-Volmerange, B. 1990, A&A, 227, 362
Henry, J. P., Gioia, I. M., Maccacaro, T., Morris, S. L., Stocke, J. T., & Wolter, A. 1991, preprint
Ishizawa, T., Matsumoto, R., Tajima, T., Kageyama, H., & Sakai, H. 1983, PASJ, 35, 61
Jones, J. C., & Forman, W. 1984, ApJ, 276, 38
Koo, D. C. 1990, in The Evolution of Universe of Galaxies: The E. Hubble Centennial Symposium, ed. R. C. Kron, in press
Kriss, G. A., Cioffi, D. F., & Canizares, C. R. 1983, ApJ, 272, 439
Lucchin, F. 1989, in Lecture Notes in Physics, Vol. 332, Morphological Cosmology, ed. P. Flin (Berlin: Springer), 284
Mamon, G. A. 1990, in IAU Colloq. 124, Paired and Interacting Galaxies, ed. J. W. Sulentic, W. C. Keel, & C. M. Telesco (NASA CP 3098), 609
McLeod, J. V. 1962, Quart. J. Math. Oxford, 13, 119
Morgan, W. W., Kayser, S., & White, R. A. 1975, ApJ, 199, 545
Merrit, D. 1983, ApJ, 264, 24
Nakano, T. 1966, Prog. Theor. Phys., 36, 515
Ostriker, J. P., & Tremaine, S. D. 1975, ApJ, 202, L113

Peebles, P. J. 1980, The Large Scale Structure of the Universe (Princeton: Princeton Univ. Press)
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Quinn, P. J., Salomon, J. K., & Zurek, W. H. 1991, preprint
Ramella, M., Geller, M. J., & Huchra, J. P. 1989, ApJ, 344, 57
Richstone, D. 1990, in Clusters of Galaxies, ed. W. R. Oegerle, M. J. Fitchett, & L. Danly (STScI Symposium Series) (Cambridge: Cambridge Univ. Press), 231
Richstone, D. O., & Malumuth, E. M. 1983, ApJ, 268, 30
Roos, N., & Norman, C. A. 1979, A&A, 76, 75
Saslaw, W. C. 1985, Gravitational Physics of Stellar and Galactic Systems (Cambridge: Cambridge Univ. Press)
Schwartz, D. A., Schwarz, Y., & Tucker, W. 1980, ApJ, 238, L59
Shandarin, S. F., & Zel'dovich, Ya. B. 1989, Rev. Mod. Phys., 61, 185
Silk, J., & White, S. D. M. 1978, ApJ, 223, L59
Sutherland, W. 1988, MNRAS, 243, 159
Toomre, A. 1977, in The Evolution of Galaxies and Stellar Population, ed. B. M. Tinsley & R. B. Larson (Cambridge: Cambridge Univ. Press), 111
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Tyson, A. 1988, AJ, 96, 1
van Dongen, P. G. J., & Ernst, M. H. 1985, Phys. Rev. Lett., 54, 1396
Villumsen, J. 1989, ApJS, 71, 407
Vittorio, N., & Turner, M. S. 1987, ApJ, 316, 475
von Smoluchowski, M. 1916, Phys. Z., 17, 557
West, M. J., Oemler, A., & Dekel, A. 1988, ApJ, 327, 1
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341