INTERNAL DYNAMICS OF GALAXY CLUSTERS

DAVID MERRITT
Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 1A1, Canada

ABSTRACT Recent results on the dynamics and post–collapse evolution of rich clusters are reviewed, with an emphasis on the Coma cluster. Current constraints on the distribution of dark matter in clusters are still rather weak; contrary to general belief, the limits derivable from kinematical data are about as good as those derivable from the existing X–ray data. In the case of the Coma cluster, the allowed values of M/L span nearly an order of magnitude; high–mass models require galaxy orbits that are very radial in the outer parts. If the dark matter is distributed like the galaxies, then tidal forces are sufficiently strong to limit the masses of dark halos to $\sim 5 \times 10^{11} M_\odot$ in the centers of clusters like Coma. The orbits of such massive galaxies could decay appreciably over a cluster lifetime due to dynamical friction; this decay has probably been observed as an enhancement in the density of galaxies very near the centers of many rich clusters. Observations of the morphology and kinematics of multiple–nucleus cD galaxies suggest that "cannibalism" is a real effect, although it does not appear to be occurring at a fast enough rate to produce a cD galaxy in a cluster that did not contain one initially. Thus, if cD galaxies are the end–products of repeated mergers, it is likely that most of these mergers took place before or during cluster formation. N–body simulations of cluster formation suggest that this is a viable hypothesis.

INTRODUCTION

Rich galaxy clusters are the largest dynamically relaxed systems in the universe, and as such provide some of the best available sites for studying the distribution of dark matter. Because of their high central densities, clusters are also useful as "laboratories" for studying the physics of galaxy–galaxy and galaxy–dark matter interactions. Over the last few years, as redshift surveys have begun to shed light on the distribution of galaxies on larger and larger scales, interest in the internal dynamics of galaxy clusters has declined. This is too bad, because new observational techniques have greatly increased the speed with which photometric and kinematical data on cluster galaxies can be obtained. Furthermore, there are a number of fundamental questions concerning the dynamics and evolution of galaxy clusters that remain to be convincingly answered. These questions include: What is the detailed distribution of dark matter in clusters? Is it more or less centrally concentrated than the galaxies? What fraction of the dark matter is bound to galaxies, i.e., do galaxies in clusters have massive dark halos? Do encounters between galaxies in clusters significantly affect their internal dynamics or their observable properties? Do galaxy orbits decay appreciably after cluster formation, and if so, is it possible to form a very bright "cD" galaxy through the repeated accretion of galaxies whose orbits have decayed?

This article will review our current understanding of the dynamics and post–collapse evolution of rich clusters. It will be assumed throughout that clusters may be usefully approximated as spherical, relaxed systems that are dynamically uncoupled from the surrounding universe. Such an assumption is dangerous, since much recent work suggests that subclustering and infall are common phenomena, and hence that many clusters are still in the throes of formation. Nevertheless a simple model may still be relevant for the inner regions of clusters where dynamical times are much shorter than the age of the universe.

Because of the prominent role which it has played in observational and theoretical studies, the Coma cluster will be singled out for attention here, even though Coma may not be typical of the majority of rich clusters in terms of its richness or density (both of which are very high).

THE DISTRIBUTION OF DARK MATTER IN CLUSTERS

The starting point for estimating the mass of a stellar or galactic system in dynamical equilibrium is the time–independent Jeans equation:

$$n\nabla\Phi = -n\overline{\mathbf{v}}\cdot\nabla\overline{\mathbf{v}} - \nabla\cdot n\sigma^2. \tag{1}$$

Here $n(\mathbf{r})$ is the number density of some "tracer" population (stars, galaxies, etc.), $\overline{\mathbf{v}}(\mathbf{r})$ is their mean velocity, $\Phi(\mathbf{r})$ is the total potential, and σ^2 is the velocity dispersion tensor:

$$\sigma_{ij}^2 = \overline{(v_i - \overline{v_i})\,(v_j - \overline{v_j})}. \tag{2}$$

The Jeans equation is a moment (over velocity space) of the collisionless Boltzmann equation, which describes the "flow" of a set of particles through phase-space in response to the acceleration induced by some gravitational potential $\Phi(\mathbf{r})$. Equation (1) is more useful than the full Boltzmann equation because the right hand side depends only on the lowest moments of the phase-space distribution function $f(\mathbf{r},\mathbf{v})$, which are generally far more accessible to observation than the distribution function itself. The Jeans equation is however incomplete in the sense that the distribution function of a collisionless system is not uniquely specified by its lowest moments. Thus it is possible to find solutions to equation (1) which appear reasonable, but are not consistent with any completely nonnegative phase-space density. (This fact can sometimes be used to rule out certain potentials $\Phi(\mathbf{r})$; see below.) Furthermore, reduction of the kinematical data to its lowest moments (e.g. $n(r), \sigma^2(r)$) can easily conceal the fact that a particular system is far from equilibrium, and therefore not describable by a time-independent equation. This point is particularly important given the recent evidence that many galaxy clusters contain statistically significant substructure (see review by M. Fitchett in this volume).

Galaxy clusters are generically fairly round and slowly-rotating. Assuming spherical symmetry and setting the mean-motion terms to zero, the Jeans equation becomes

$$n\frac{d\Phi}{dr} = n\frac{GM(r)}{r^2} = -\frac{d\left(n\sigma_r^2\right)}{dr} - \frac{2n}{r}\left(\sigma_r^2 - \sigma_t^2\right). \tag{3}$$

Here $M(r)$ is the total mass contained within r, and σ_r and σ_t are the galaxy velocity dispersions along and tangential to any radius vector. Equation (3) allows us to estimate the mass distribution in a cluster given knowledge of the three functions $\{n(r), \sigma_r(r), \sigma_t(r)\}$. Until recently, equation (3) was rarely used for this purpose, primarily because of the difficulty of obtaining a usefully-large sample of galaxy radial velocities. This is no longer the case; new techniques, such as multi-object fiber spectrometry, allow one to measure dozens of radial velocities in a single observing run. There are now several clusters for which more than 100 radial velocities have been measured, and for the Coma and Virgo clusters, this number exceeds 300. A more fundamental problem with equation (3) arises from the nature of the information required to evaluate its right hand side. By measuring the radial velocity of a large sample of galaxies, we can in principle determine $\sigma_{los}(R)$, the dependence of the line-of-sight velocity dispersion on (projected) radius from the cluster center. But there is no way to deconvolve a *single* function of radius $\sigma_{los}(R)$ to obtain the *two* desired functions $\{\sigma_r(r), \sigma_t(r)\}$. Physically, this indeterminacy reflects the fact that spatial variations in either velocity anisotropy or cluster mass-to-light ratio may be responsible for the observed variation of σ_{los} with R.

Given this indeterminacy, there are several possible ways of proceeding, none completely satisfactory. By far the most common is to make an *a priori* assumption about the form of the mass distribution—e.g., that the matter is distributed like the observed galaxies—and then to derive a total mass using the virial theorem. If we multiply equation (3) by $4\pi r^3$ and integrate from zero to infinity, the result is

$$\left\langle r\frac{d\Phi}{dr}\right\rangle = \left\langle v^2\right\rangle. \tag{4}$$

The brackets indicate spatial averages over the observed sample of galaxies. Writing $d\Phi/dr = GM(r)/r^2 = GM_\infty F(r)/r^2$, where $F(r) \leq 1$ is the mass fraction within r, equation (4) becomes

$$GM_\infty = \frac{\langle v^2\rangle}{\langle r^{-1}F\rangle}. \tag{5}$$

Equation (5), which is a form of the virial theorem, relates the total cluster mass M_∞ to the velocity dispersion of the observed sample, and a quantity $\langle r^{-1}F\rangle$, which depends on the (generally unknown) form of the matter distribution. Note that, by taking the proper moment of the Jeans equation, we obtained an expression for the total mass that depends only on the total mean square velocity $\langle v^2\rangle = 3\langle v_{los}^2\rangle$. It is this lack of dependence on velocity anisotropy (at least in the context of spherical systems) that makes the virial theorem so useful. However it is clear from equation (5) that the inferred total mass will depend strongly on its assumed distribution. If the mass is distributed like the observed galaxies—in the form of "heavy halos", for instance—then the appropriate form of equation (5) is easily shown to be

$$GM_\infty = \frac{3\pi}{2}\frac{\langle v^2\rangle}{\langle r_{ij}^{-1}\rangle} \tag{6}$$

(Limber and Mathews 1960), where r_{ij} is the projected distance between any pair of galaxies. An equation similar to (6) was used by Zwicky (1933) to infer the mass of the Coma cluster from a sample of seven galaxies. He found a mass–to–light ratio consistent with the best modern estimates, i.e.

$$\left(\frac{M}{L}\right)_{Coma} \approx 350h\left(\frac{M}{L}\right)_{sun} \tag{7}$$

(e.g. Kent and Gunn 1982), where $h = H_0$ in units of $100\,\mathrm{km\,s^{-1}}$. Applying this mass–to–light ratio to the local universe gives a cosmological density parameter $\Omega_0 \approx 0.15$.

If the dark matter has a different distribution than the galaxies—as it must if, for instance, $\Omega_0 = 1$—then the virial theorem is not very useful. One can easily derive a *lower* limit on the mass required to bind a cluster by setting $F(r) = M(r)/M_\infty = 1$, i.e. by putting all of the dark matter at the cluster center. For a cluster like Coma, this assumption reduces the required mass below Zwicky's value by a factor of about five. A more reasonable model (at least in the eyes of most cosmologists) would have a dark matter distribution that is *more* extended than the galaxies, in which case $F(r)$ is small and the total mass can be arbitrarily large. Thus, although the virial theorem does not constrain the total mass very well, it does imply a *relation* between the total mass and its distribution: the more centrally concentrated the matter, the less is required to bind the cluster, and vice versa.

One can reduce (though not eliminate) this indeterminacy by making use of the additional information contained within the line–of–sight velocity dispersion profile $\sigma_{los}(R)$. For any assumed dark matter potential $\Phi(r)$, there is a unique set of functions $\{\sigma_r^2(r), \sigma_t^2(r)\}$ which satisfy the Jeans equation (3) and give the correct projected profile $\sigma_{los}(R)$. But for certain $\Phi(r)$, the derived σ_r^2 or σ_t^2 may be negative at some radii, corresponding to models with negative numbers of galaxies on some orbits. Such models can be ruled out as unphysical. Figure 1 illustrates this technique with the Kent and Gunn (1982) Coma data, and an assumed potential

$$\Phi(r) \propto \log(r^2 + r_o^2), \tag{8}$$

corresponding to a mass density that falls off as r^{-2} at large radii (more slowly than the galaxies). Large values of r_o—i.e., nearly uniform dark matter distributions—require very radial orbits, and for $r_o \gtrsim 5h^{-1}$ Mpc, no solution is possible.

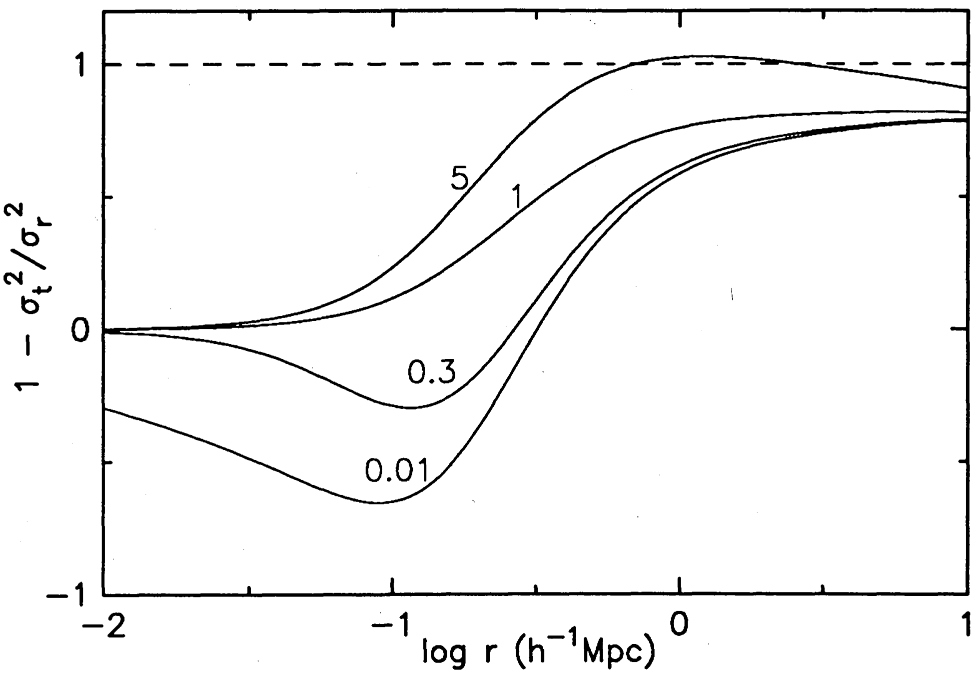

Fig. 1 Dependence of Coma galaxy velocity anisotropy on radius, assuming the dark–matter potential of eqn. (8). Curves are labelled by the dark–matter core radius r_0 in Mpc. Anisotropies greater than unity are unphysical.

Testing a large number of assumed potentials $\Phi(r)$ in this way gives approximate upper and lower bounds on the Coma mass–to–light ratio:

$$150h \lesssim \left(\frac{M}{L}\right)_{Coma} / \left(\frac{M}{L}\right)_{sun} \lesssim 1200h \qquad (9)$$

(The and White 1986; Merritt 1987). The low-mass models are perhaps doubly unphysical, in that they require both a high central concentration of dark matter, as well as galaxy orbits that are close to circular at large radii ($\gtrsim 1h^{-1}$ Mpc). Most theories of the early universe predict that the galaxies would be, if anything, more clustered than the dark matter. Furthermore, to the extent that clusters form via gravitational collapse, the galaxy orbits should be predominantly radial, not circular. However, even excluding the tangentially anisotropic models, the mass of the Coma cluster is still uncertain by a factor of about three.

If the number of galaxy radial velocities available in the Coma cluster were much larger—of order 10^3 or 10^4—one could make further progress at constraining its mass by investigating the form of the line–of–sight velocity

distribution as a function of projected radius. In effect, this technique would require any model of Coma to be consistent not only with the Jeans equation, but also with the more detailed Boltzmann equation from which the Jeans equation is derived. No one has yet described the best way to carry out this task (nor is there any cluster that is both sufficiently well observed, and convincingly close to equilibrium, to justify such an analysis). In the case of Coma, one way to make use of the extra information contained within the full velocity distribution function is illustrated in Figure 2. The overall velocity histogram appears marginally most consistent with a high–mass, radial–orbit model; a model with low mass and circular orbits appears strongly inconsistent. There are, however, a number of reasons to be cautious about this comparison. The shape of the velocity histogram can be strongly affected by processes such

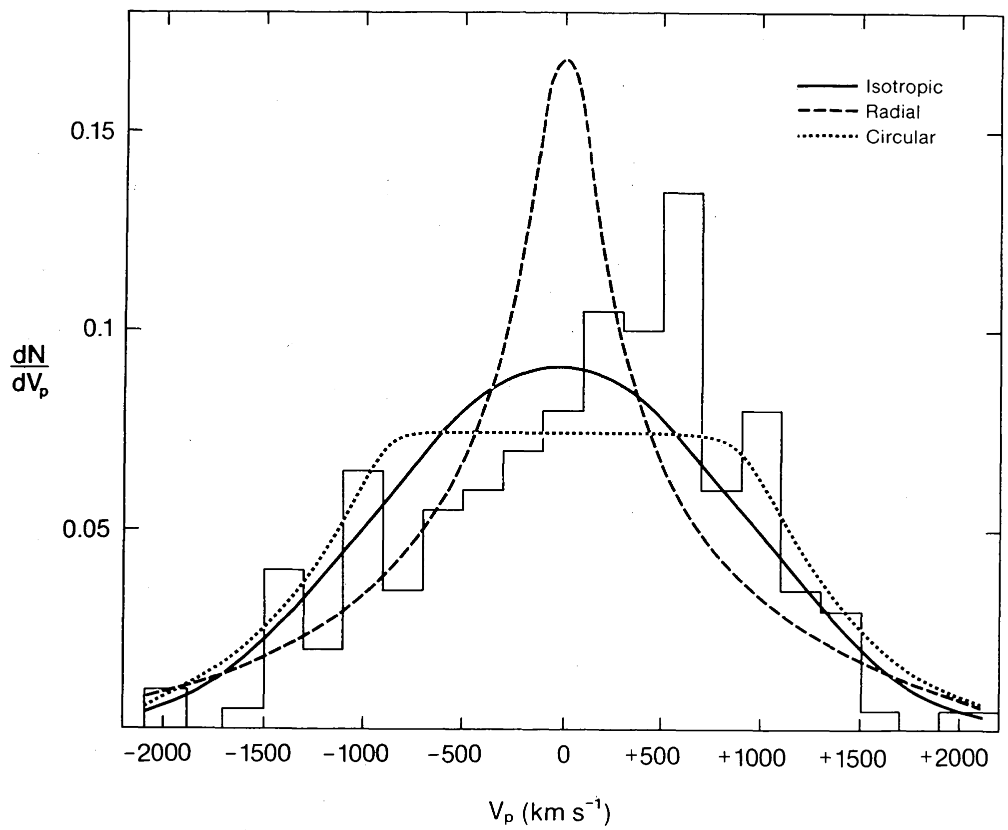

Fig. 2 Velocity histogram for galaxies in Coma. The three curves are derived from models in which the galaxy orbits are isotropic, radial, and circular; the dark matter distributions have been adjusted to give the same line–of–sight velocity dispersion profile in each case. (From Merritt, *Ap. J.*, **313**, 121.)

as rotation and infall which we have so far neglected. In fact, inspection of Figure 2 reveals a possibly significant (~ 97% confidence) degree of skewness in the observed distribution. Cluster rotation by itself would tend to broaden the velocity distribution rather than make it skew; furthermore, Coma does not exhibit significant rotation (Rood et al. 1972). However there is good reason to believe that contamination by foreground galaxies might explain the low–velocity "tail". De Lapparent, Geller and Huchra (1986) show that the Coma cluster appears to sit at the intersection of a number of large–scale galaxy "shells", one of which lies nearly along the line of sight to Coma. These foreground galaxies could significantly affect the form of the overall velocity histogram, even if they have little effect on the inferred dynamics of the central regions.

A number of other techniques have been discussed for constraining the orbital kinematics of galaxies in clusters. Pryor and Geller (1984) attempted to use the observed tidal radii and gas content of galaxies in Coma to put limits on their orbital pericenters, and hence on the degree of velocity anisotropy. Their result (that the Coma cluster is close to isotropic within $1h^{-1}$ Mpc) is strongly dependent on the uncertain physics of tidal truncation and gas dynamical ablation; furthermore those authors only considered models in which the dark matter is distributed like the galaxies, while in fact the available velocity data imply a fairly tight relation between the mass distribution and the galaxy orbits, as discussed above. O'Dea, Sarazin and Owen (1987) used the orientation of "narrow angle tail" radio sources in clusters to constrain the distribution of galaxy orbits, under the assumption that the radio–luminous plasma ejected by a moving galaxy is bent into a tail which marks the path taken by the galaxy through the cluster. Since most clusters contain only a few such radio sources (Coma, for instance, contains only one), those authors were forced to superpose data from many clusters. They obtained the surprising result that galaxy orbits in the inner $\sim 0.5h^{-1}$ Mpc of their clusters are strongly *radial*; at large radii the distribution of tail orientations appears to be random. The correct interpretation of this result will probably have to await a better understanding of the gas ablation process. It may be, for instance, that the probability of observing a galaxy as a narrow angle tail depends strongly on its velocity with respect to the intracluster gas, in which case the observed sample could be kinematically biased.

For a long time it was hoped that X–ray observations of hot intracluster gas would resolve the indeterminacy of cluster masses. The equation of hydrostatic equilibrium, in spherical symmetry, states

$$\frac{d\Phi}{dr} = \frac{GM(r)}{r^2} = -\frac{1}{\rho_g}\frac{dP_g}{dr} = -\frac{1}{\rho_g}\frac{d}{dr}\left(\rho_g k T_g\right), \tag{10}$$

where ρ_g and T_g are the gas density and temperature. Equation (10) is simpler than the Jeans equation (3) since gas is a collisional fluid with an isotropic pressure; thus the two functions $\{\sigma_r(r), \sigma_t(r)\}$ are replaced by one, $T_g(r)$. Furthermore, the statistical accuracy of a mass determination based on the X–ray emitting gas can always be increased by lengthening the integration time, whereas the number of bright galaxies in a cluster is limited. Unfortunately, the spatial resolution of the spectral instruments on past X–ray satellites has not been very good, and at present there is no cluster (with the possible exception of Virgo) for which we have an accurate determination of $T_g(r)$. This problem is capable of solution; future satellites, such as AXAF, should yield accurate temperature profiles and hence accurate masses for nearby clusters. However we have so far learned little about cluster masses from X–ray studies that we could not have learned from the kinematics of relatively modest samples of cluster galaxies.

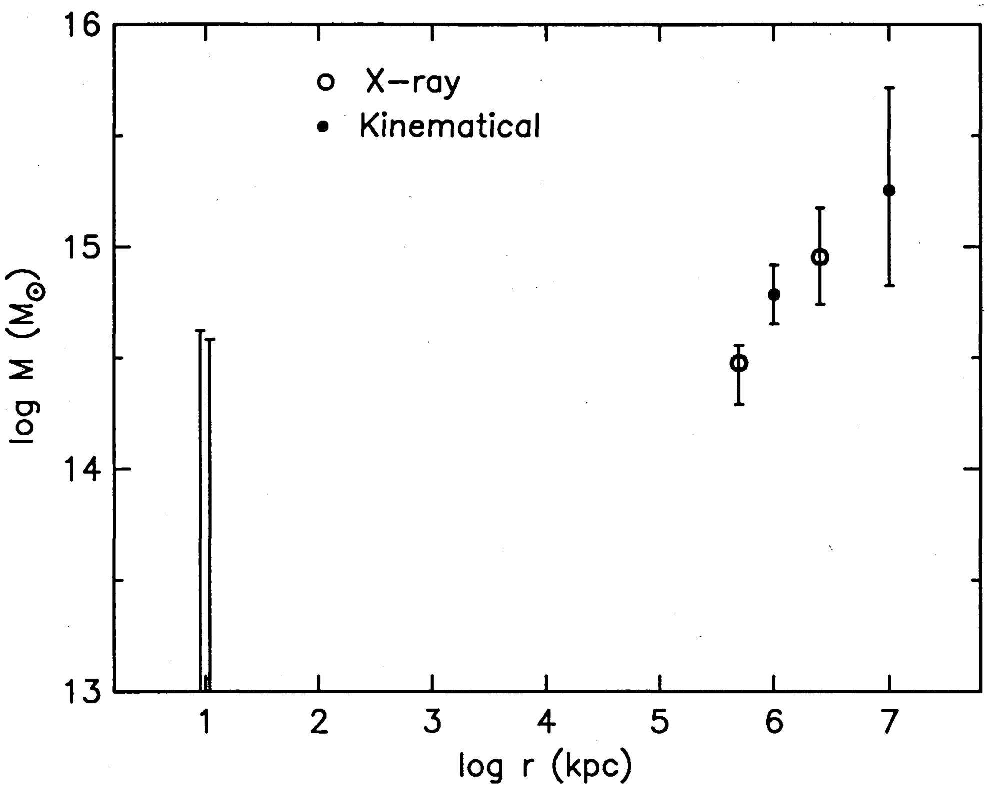

Fig. 3 Constraints on the mass distribution in Coma derived from kinematical and X–ray techniques ($h = 1$).

This point is made quantitatively in Figure 3, which presents limits on the mass distribution of the Coma cluster obtained from the kinematical technique described above, as well as the limits from the most recent analysis of the Coma X–ray data (Hughes 1988). In the vicinity of $\sim 1h^{-1}$ Mpc, both techniques give similar results for the enclosed mass, with error bars that span about a factor of two in both cases. Neither technique places interesting constraints on the central mass density; at large radii, the kinematical mass determination is superior, because currently available X-ray emissivity data for Coma do not extend beyond $\sim$ 50 arcmin $\approx 1h^{-1}$ Mpc.

DARK MATTER: BOUND TO GALAXIES OR SMOOTHLY DISTRIBUTED?

Given that the total masses of clusters like Coma greatly exceed their luminous masses—although by amounts that are still fairly uncertain—it is natural to wonder whether the dark matter is bound to the bright galaxies (in the form of massive halos, say) or whether it is more smoothly distributed. Here it is important to remember that there is currently no strong evidence for massive halos around "hot" stellar systems, and thus no reason to assume that the cluster dark matter is bound to galaxies, most of which are elliptical. The fraction of cluster mass that is associated with galaxies has important consequences for the rate of galaxy–galaxy interactions and orbital decay (see below). One of the first attempts to deal with this question was made by Noonan (1970), who noted that the tidal field produced by the potential well of a cluster might be sufficiently strong to pull *luminous* matter (along with, presumably, any dark–matter halos) from the brightest galaxies. Noonan's hypothesis has been unaccountably neglected; most subsequent work has focused instead on collisions as the physical process that determines galaxy masses in clusters.

There are two major sources of uncertainty that make it difficult to assess the importance of Noonan's tidal truncation hypothesis. 1. As discussed above, neither the (macroscopic) distribution of the cluster dark matter, nor the distribution of galaxy orbits, is well constrained, even in the best–observed clusters like Coma. 2. Although simple tidal theory makes a definite prediction about the limiting radius of a galaxy, with *specified* mass, as it orbits about a potential center, it says essentially nothing about the tidally–limited mass itself. This is because the total mass depends strongly on the distribution of matter near the galaxy's Lagrangian radius, and this distribution is essentially unknown. A straightforward analysis gives a feeling for the uncertainties involved. Consider a galaxy orbiting at a fixed radius r_0 from the center of a cluster. The equations of motion, in the rotating frame, of a star that remains close to the galaxy in the plane of its orbit are

$$\ddot{x} = -2\Omega\dot{y} - \frac{d\Phi_g}{d\delta} + x\left[\Omega^2 - \frac{d^2\Phi_{cl}}{dr^2}\,|_{r_0}\right], \qquad (11a)$$

$$\ddot{y} = +2\Omega\dot{x} - \frac{d\Phi_g}{d\delta}. \qquad (11b)$$

Here $\vec{\delta} = (x, y)$ is the position of the star relative to the galaxy center, with $\mathbf{x}$ parallel to the orbital radius vector $\mathbf{r_0}$; $\Phi_g(\delta)$ and $\Phi_{cl}(r)$ are the galaxy and cluster potentials, respectively; and Ω is the angular frequency of the galaxy orbit. Neglecting Coriolis terms, equation (11a) predicts that a star would feel a zero net force at the Lagrangian radius δ_L, where $\delta_L\left[\Omega^2 - \left(d^2\Phi_{cl}/dr^2\right)_{r_0}\right] = \left(d\Phi_g/d\delta\right)_{\delta_L}$. In order to derive a tidally-truncated mass from this equation, we need to assume some relation between the (truncated) mass m_g and the tidal radius r_t of the galaxy. Suppose that $Gm_g = \alpha^2\sigma_g^2 r_t/2$, where σ_g is the galaxy central velocity dispersion, and α is an unknown parameter that specifies the shape of the halo density profile. If the dark matter producing the tidal field is distributed roughly "isothermally" near the cluster center, with central density ρ_0, then $\Omega^2 \approx d^2\Phi_{cl}/dr^2 \approx G\rho_0$, and

$$Gm_g \approx \alpha^3\sigma_g^3(G\rho_0)^{-1/2}. \qquad (12)$$

A more careful calculation would give a value (dependent on orbital parameters) for the undetermined coefficient in equation (12). Note, however, that the predicted mass depends sensitively on the unknown structural parameter α. If the halo of the tidally truncated galaxy has the same mass distribution as an isotropic Michie–King model of high central concentration—a natural assumption, given that these models look so much like globular clusters, which are thought to be tidally truncated—then $\alpha \approx 1$. But even along the Michie–King sequence of models, α oscillates between ~0.85 and ~1.3, giving an uncertainty of nearly a factor of *four* to the predicted mass. Other reasonable families of models (eg. Wilson [1975] spheres) have halos that look very different. Thus, simple tidal theory can give only an order–of–magnitude indication of the mass that could remain bound to a galaxy in a cluster, even a cluster with a known potential.

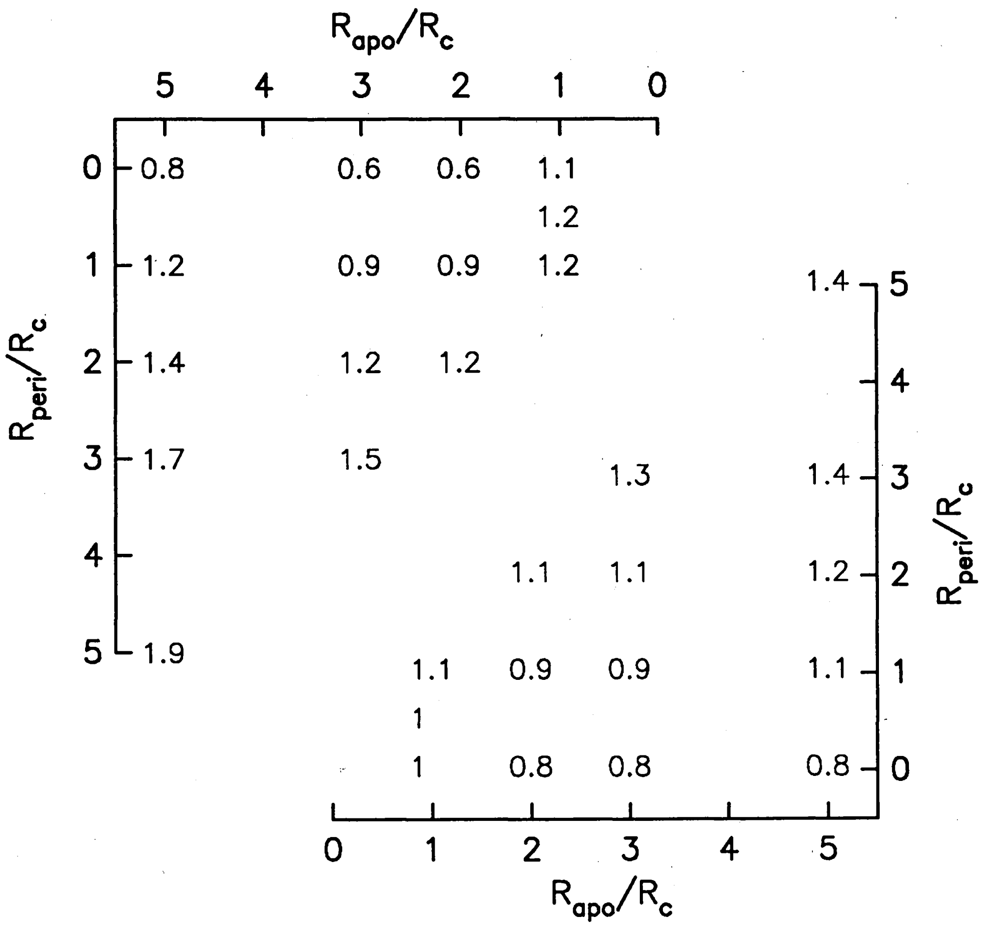

Fig. 4 Tidally–truncated masses, in units of $G^{-1}\sigma_g^3(G\rho_0)^{-1/2}$, of galaxies orbiting in a cluster with potential given by eqn. (13). Lower panel: $m_{initial} = 2$; upper panel: $m_{initial} = 3$. R_{apo}: orbital apocenter; R_{peri}:orbital pericenter.

The only way to reduce this uncertainty is to simulate the tidal truncation process via an N–body code. Figure 4 shows the final masses of model galaxies, containing 2000 particles initially, after orbiting for a time $\sim 40(G\rho_0)^{-1/2}$ about the center of a cluster with mass density

$$\rho_{cl}(r) = \rho_0 \left(1 + r^2/r_c^2\right)^{-3/2} \tag{13}$$

(see Merritt and White 1987 for a description of the technique). For Zwicky's Coma model, $\rho_0 \approx 1 \times 10^{-2} h^2 \mathrm{M}_\odot \mathrm{pc}^{-3}$, making the simulated evolution time $\sim 7 \times 10^9 h^{-1}$ years. According to Figure 4, galaxies on elongated, high–energy orbits are the most strongly truncated, although the dependence of final mass on orbital parameters is not great. Final masses of galaxies with pericenters near r_c (the dark matter "core radius") are given approximately by

$$\begin{aligned} m_f &\approx 1 \times G^{-1} \sigma_g^3 (G\rho_0)^{-1/2} \\ &\approx 6 \times 10^{11} M_\odot \left(\frac{\sigma_g}{250\,\mathrm{km\,s^{-1}}}\right)^3 \left(\frac{\rho_0}{0.01 \mathrm{M}_\odot \mathrm{pc}^{-3}}\right)^{-1/2} . \end{aligned} \tag{14}$$

Detailed examination of the final states of these "truncated galaxies" reveals that they contain about four times as much mass within their tidal radii as do $\alpha = 1$ Michie–King models. The final tidal and half–mass radii are

$$r_t \approx 30 \text{ kpc} \left(\frac{\sigma_g}{250\,\mathrm{km\,s^{-1}}}\right) \left(\frac{\rho_0}{0.01 \mathrm{M}_\odot \mathrm{pc}^{-3}}\right)^{-1/2} , \tag{15a}$$

$$r_{1/2} \approx 8 \text{ kpc} \left(\frac{\sigma_g}{250\,\mathrm{km\,s^{-1}}}\right) \left(\frac{\rho_0}{0.01 \mathrm{M}_\odot \mathrm{pc}^{-3}}\right)^{-1/2} . \tag{15b}$$

Equation (14) predicts that a bright galaxy orbiting near the center of the Coma cluster could indeed retain a fairly massive halo, similar in mass to the halos of bright spiral galaxies, as long as the central density of dark matter in Coma is not much greater than in Zwicky's model. Furthermore, equations (15) imply that little if any *luminous* matter would be tidally removed, since characteristic luminous radii of bright elliptical galaxies are $\lesssim 5h^{-1}$ kpc. Thus tidal truncation is probably not a viable mechanism for producing the "diffuse light" that is thought to be present in some clusters, including Coma (e.g. Thuan and Kormendy 1977). Both of these conclusions would have to be modified if cluster central densities could be shown to be much higher than in Zwicky's model. High central densities are likely in clusters containing "cD" galaxies, especially if the cD's are themselves surrounded by supermassive dark halos. This point is discussed further in the next section.

Finally, we can estimate the total fraction of the dark matter in Coma that could remain bound to galaxies in a model like Zwicky's. If we assume a Schechter (1976) galaxy luminosity function $N(L)dL \propto (L/L^*)^{-1} \exp(-L/L^*) d(L/L^*)$, and furthermore that the velocity dispersion of the halo material scales with the luminosity of its parent galaxy as $\sigma_g \approx 225\,\mathrm{km\,s^{-1}} (L/L^*)^{1/4}$, then equation (14) predicts that a fraction

$$\frac{M_g}{M_{cl}} \approx 0.15 \tag{16}$$

of the Coma cluster's mass should reside in galaxies. This estimate is an upper limit in the sense that it neglects the (probably greater, but transient) tidal stresses during cluster formation, galaxy–galaxy collisions, etc. It suggests that almost all of the dark matter in rich clusters is smoothly distributed.

GALAXY COLLISIONS

If the dark halos of galaxies in rich clusters are as large as the preceding discussion suggests, then a typical bright galaxy will undergo a number of interpenetrating collisions with other galaxies over a Hubble time. Because galaxy orbital velocities in clusters are usually much greater than their internal velocities, few if any of these encounters will result in mergers. (This may not be true for galaxies whose orbits confine them to the cluster core, especially if a supermassive "cD" galaxy is present; see the next section.) However close collisions might be effective at removing dark matter from galactic halos. One of the first discussions of collisional stripping in the context of galaxy clusters was that of Richstone (1976), who estimated that, if *all* of a cluster's dark matter was originally bound to the galaxies in the form of massive halos, all but $\sim$ 10% would be liberated in a Hubble time. Richstone's calculation ignored tidal forces due to the cluster mean field; as we have seen, tidal truncation by itself guarantees that roughly the same fraction of the dark matter will be smoothly distributed once a cluster has reach equilibrium, even in the absence of collisions. Nevertheless we cannot quite rule out collisions as unimportant in clusters, because Richstone's (1976) study (as well as most later studies) was based on rather crude estimates of the collisional mass loss rate. Furthermore, an occasional close collision between two massive galaxies could substantially affect their observable properties, even if *typical* collisions are relatively "mild".

Consider a gravitational encounter between two galaxies. During the encounter, each star in the "test" galaxy feels a tidal acceleration toward the "perturber" of order

$$\frac{dv}{dt} \approx \frac{Gm_p\delta}{r^3} \tag{17}$$

where δ and r are the separations of the star from the center of the test galaxy and the perturber, respectively, and m_p is the mass of the perturber. Since encounters in clusters typically occur at high speeds, we can ignore to first order the orbital motion of the star within the test galaxy when computing the total momentum given to it by the perturber (the "impulse approximation"). Thus

$$\Delta v \approx \frac{dv}{dt} \times \frac{p}{V} \approx \frac{Gm_p\delta}{p^2V}, \tag{18}$$

where p is the collision impact parameter and V is the encounter velocity. Equation (18) implies that close, slow encounters are the most effective at removing mass. Notice however that this equation can only be valid over a limited range of impact parameters. For $p \gg \delta$, the collision is slow, and the impulse approximation does not apply. For $p \lesssim \delta$, the galaxies are close or interpenetrating, and the tidal acceleration will not be given by the simple formula (17). Furthermore, the fraction of the test galaxy's mass that acquires positive binding energy and escapes must depend strongly on its internal velocity distribution. Thus an accurate estimate of the mass loss resulting from an encounter requires a full N–body simulation.

The most detailed study to date of this problem is that of Aguilar and White (1985), who simulated a large number of collisions between spherical galaxies with either of two density profiles. They found that close collisions, which are the only important ones for mass loss, are not well described by the tidal approximation; in particular, the fractional mass loss does not depend on simple powers of the encounter parameters as equation (18) suggests and as most previous workers (e.g. Richstone 1976, Merritt 1984a) have assumed. In order to estimate how greatly those earlier studies were in error, Aguilar and White (1985) calculated the mean rate at which a galaxy would lose mass due to encounters with a uniform background of equal–mass galaxies with a Maxwellian distribution of velocities. They found:

$$\left(-\frac{1}{m_g}\frac{dm_g}{dt}\right)^{-1} \equiv T_c \approx \left(40 n_g v_{cl} r_e^2\right)^{-1}\left(\frac{v_{cl}}{v_{rms}}\right)^3$$
$$\approx 1.6\times10^{10}\mathrm{yr}\left(\frac{n_g}{10^3\mathrm{Mpc}^{-3}}\right)^{-1}\left(\frac{v_{cl}}{10^3\,\mathrm{km\,s^{-1}}}\right)^{-1}\left(\frac{r_e}{10\mathrm{kpc}}\right)^{-2}\left(\frac{v_{cl}}{4v_{rms}}\right)^3 \quad (19)$$

where n_g is the number density of galaxies, r_e is the galaxy "effective" (projected half–mass) radius, and v_{rms} and v_{cl} are the galaxy internal and orbital velocity dispersions, respectively. Perhaps fortuitously, equation (19) agrees rather well with Richstone's (1976) estimate when $v_{cl}/v_{rms} \approx 4$, appropriate for the rich–cluster environment which Richstone considered. In these clusters, collisions would not be expected to liberate much mass in a Hubble time unless galaxies managed to retain halos much more extensive than those permitted by the mean tidal field. However the strong dependence of T_c on v_{cl} suggests that collisions might be very effective in poorer clusters, for which $v_{cl} \lesssim 2v_{rms}$.

It has occasionally been suggested (e.g. Gallagher and Ostriker 1972) that close collisions between galaxies in dense cluster environments might liberate a substantial amount of *luminous* material over a Hubble time, thus contributing to the luminosity of a central cD galaxy. The simulations of Aguilar and White (1985) suggest that this is very unlikely to be an important process unless a substantial fraction of a typical galaxy's light lies more than $\sim$ 10 kpc from its center. On the other hand, if a supermassive galaxy should happen to be present *ab initio* at the center of a rich cluster, then a modification of the arguments in the preceding section suggest that it might be responsible for tidally removing at least a moderate amount of light from other galaxies. If the mass of the central galaxy is distributed like an isothermal sphere, then its mass within a radius r is $m_{cD} \approx 2\sigma_{cD}^2 r/G$, and its average density is $\langle\rho\rangle \approx \sigma_{cD}^2/2Gr^2$. The tidal truncation calculations described above then give for the limiting mass of a galaxy that passes within this radius:

$$m_g \approx 1\times G^{-1}\sigma_g^3(G\langle\rho\rangle)^{-1/2}$$
$$\approx 4\times10^{11}M_\odot\left(\frac{\sigma_g}{250\,\mathrm{km\,s^{-1}}}\right)^3\left(\frac{\sigma_{cD}}{350\,\mathrm{km\,s^{-1}}}\right)^{-1}\left(\frac{r}{50\mathrm{kpc}}\right), \quad (20)$$

and for its limiting radius,

$$r_t \approx 20\ \mathrm{kpc} \left(\frac{\sigma_g}{250\,\mathrm{km\,s^{-1}}}\right)^3 \left(\frac{\sigma_{cD}}{350\,\mathrm{km\,s^{-1}}}\right)^{-1} \left(\frac{r}{50\mathrm{kpc}}\right). \tag{21}$$

Thus a bright galaxy whose pericenter lay within $\sim$ 30 kpc of the central cD would probably lose a significant fraction of its luminous matter after a few orbits; the mass of a cD galaxy that extends to this radius would be $\sim 2 \times 10^{12} M_\odot$, which is not unreasonable. Whether this mechanism is capable of liberating enough luminous matter in a Hubble time to produce a "cD envelope", with a total luminosity $\sim 5L^*$, is questionable, however.

ORBITAL DECAY AND "CANNIBALISM"

Given that the tidal field of a rich cluster is so efficient at limiting the sizes of galactic halos, it makes sense, at least to first order, to ignore encounters between galaxies when considering how clusters might evolve dynamically after their formation. In this simple model, the most important physical process that acts to modify the orbital motion of galaxies is dynamical friction (Chandrasekhar 1943), the force that results from polarization of the dark matter as a galaxy moves through it. Roughly speaking, every galaxy is followed in its orbit by a "wake" of slight overdensity that produces a constant deceleration; as the galaxy loses kinetic energy to the dark matter, its orbit decays. The equation of motion of a galaxy in a spherical cluster that is dominated by smoothly–distributed dark matter is, in Chandrasekhar's (1943) local approximation,

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{d\Phi}{dr}\mathbf{e}_r - 4\pi G^2 m_g \rho_b(v) \left(\frac{\mathbf{v}}{v^3}\right) \ln\Lambda. \tag{22}$$

The first term on the right–hand side of equation (22) defines the unperturbed motion of the galaxy in the cluster potential $\Phi(r)$. The second term is the drag force: m_g and $\mathbf{v}$ are the galaxy mass and velocity; $\rho_b(v)$ is the density of "background" particles with velocities, relative to the cluster center of mass, less than the orbital velocity $\mathbf{v}$; and $\ln\Lambda$ is the "Coulomb logarithm", about equal to 3 in the present context.

Equation (22) (which should be viewed as a rough approximation; see e.g. White 1983) predicts that galaxies which orbit in regions of the highest dark–matter density will undergo the most rapid orbital decay. Consider therefore a galaxy whose orbit is confined to the cluster core. If the dark matter in the core is uniformly distributed with density ρ_0, then the galaxy sees a potential which is nearly harmonic,

$$\Phi(r) \approx \frac{2}{3}\pi G \rho_0 r^2 + C. \tag{23}$$

The drag force, assuming a Maxwellian distribution of background particle velocities, is

$$\frac{d^2\mathbf{r}}{dt^2} = -4\pi G^2 m_g \left(\frac{\mathbf{v}}{v^3}\right) \ln\Lambda \int_0^v dv(4\pi v^2)\frac{\rho_0}{(2\pi v_{cl}^2)^{3/2}} e^{-v^2/2v_{cl}^2}$$
$$\approx -\frac{4}{3}\frac{\sqrt{2\pi}G^2 m_g \rho_0 \ln\Lambda}{v_{cl}^3}\frac{d\mathbf{r}}{dt}, \quad (24)$$

where v_{cl} is the one–dimensional velocity dispersion of the dark–matter "particles" (probably roughly equal to the velocity dispersion of the bright galaxies). The equation describing the galaxy's orbit is then

$$\frac{d^2\mathbf{r}}{dt^2} + \frac{1}{\tau}\frac{d\mathbf{r}}{dt} + \omega^2\mathbf{r} = 0, \quad (25)$$

where

$$\omega^2 \equiv \frac{4\pi}{3} G\rho_0 \quad (26)$$

and

$$\tau \equiv \frac{3}{4\sqrt{2\pi}}\frac{v_{cl}^3}{G^2 m_g \rho_0 \ln\Lambda}$$
$$\approx 5\times 10^8 \mathrm{yr} \left(\frac{v_{cl}}{10^3\,\mathrm{km\,s^{-1}}}\right)^3 \left(\frac{m_g}{10^{12}M_\odot}\right)^{-1} \left(\frac{\rho_0}{0.01 M_\odot \mathrm{pc}^{-3}}\right)^{-1}. \quad (27)$$

Equation (25), which is the equation of a damped harmonic oscillator, states that the orbits of galaxies in cluster cores will decay with a time constant $\sim 2\tau$; furthermore, the orbits retain their shape as they shrink. (The orbits of galaxies that spend most of their time *outside* of the core tend to *circularize* as they shrink; however the orbital decay time for these galaxies greatly exceeds a Hubble time.) If we assume that galaxies retain as much matter in their halos as the tidal field permits, then the decay time becomes (equations (14), (27)):

$$2\tau \approx 2\times 10^9 \mathrm{yr} \left(\frac{v_{cl}}{4\sigma_g}\right)^3 \left(\frac{\rho_o}{0.01 M_\odot \mathrm{pc}^{-3}}\right)^{-1/2}. \quad (28)$$

Thus the orbits of bright galaxies in and near the core may have decayed appreciably since the epoch of cluster formation.

What are the observable consequences of this decay? Since the decay time is inversely proportional to galaxy mass, we would expect the most massive galaxies to fall to the cluster center most quickly. Insofar as a galaxy's mass is correlated with its luminosity—a big "if", given that most of the matter is dark—the bright galaxies should therefore be more centrally concentrated and have lower orbital velocities than the faint galaxies. This "luminosity segregation" should not be terribly striking, however, because it is a *differential* effect, and because decay times are likely to be comparable to cluster lifetimes for all but the brightest galaxies. Figure 5 shows the dependence of orbital velocity dispersion on galaxy luminosity for galaxies near the center of the Coma cluster. Also shown are the predicted dependences, at times $\{0.6, 2, 6, 20\} \times \tau^*$, for an initially isothermal cluster with a Schechter (1976) luminosity function; here τ^* is the decay time of an L^* galaxy. The predicted amount of segregation, even in the extreme models, is small compared with the statistical uncertainties. Other studies (e.g. Rood 1965; White 1977) have similarly concluded that luminosity segregation in the Coma cluster is too small to be detected.

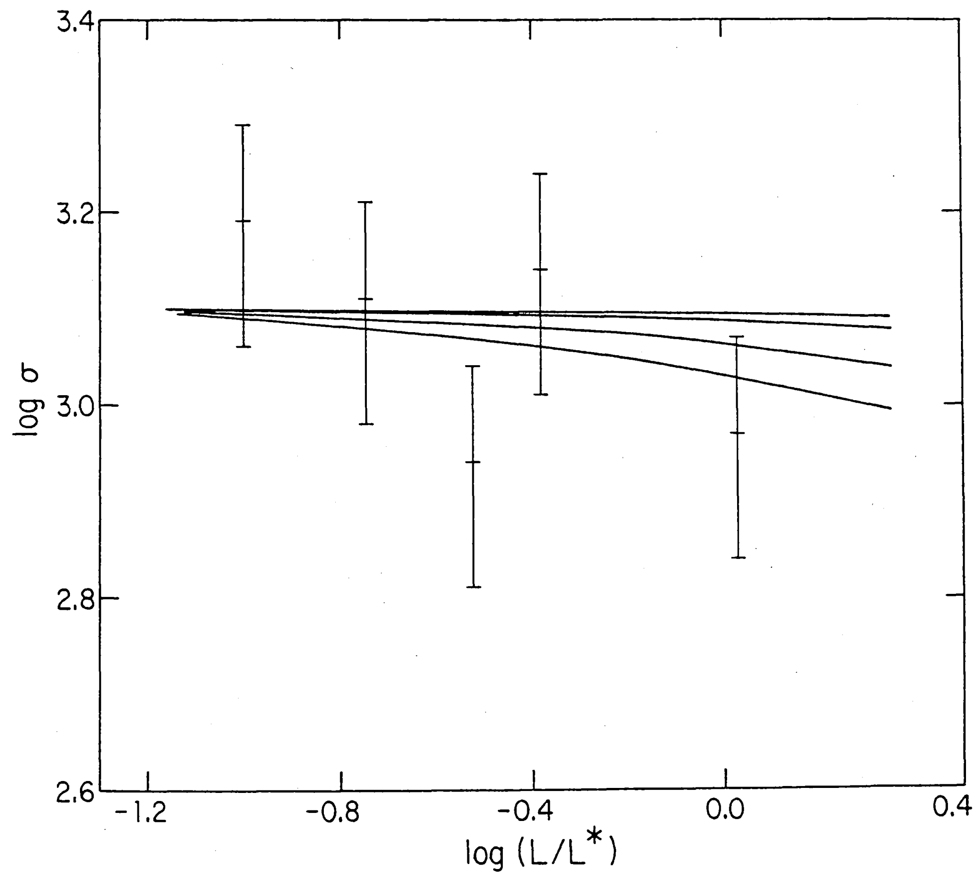

Fig. 5 Orbital velocity dispersion as a function of luminosity for galaxies within $0.35h^{-1}$ Mpc of the center of the Coma cluster. Solid line: predicted dependence, at four times, for an initially isothermal cluster with a Schechter luminosity function. (From Merritt, *Ap. J.*, **276**, 26)

A second, potentially observable consequence of orbital decay is that bright galaxies as a group should accumulate at the cluster center. It may be shown

(Merritt 1984b) that in an initially isothermal cluster, the central density of a group of galaxies of mass m_1 will increase with time constant

$$\left(\frac{1}{n_1}\frac{dn_1}{dt}\right)^{-1} = \frac{4\sqrt{2}}{3}\tau_1 \approx 2\tau_1 \tag{29}$$

for $t \lesssim \tau_1$; at later times the accumulation slows. Thus we might expect that the central density of the brighter galaxies will increase by a factor of ~ a few in a Hubble time. Such an enhancement has probably been observed. Beers and Tonry (1986) find that the combined density profile of a set of ~ 35 rich clusters is well described by a power law, $\rho \propto r^{-2}$, even at relatively small radii $\lesssim 0.25h^{-1}$ Mpc, as long as the cluster centers are assumed to coincide with either the location of a cD galaxy, or the peak of the X-ray surface brightness. Figure 6 shows the Beers and Tonry (1986) data, as well as the projected number density profiles of three initially isothermal clusters, containing equal-mass galaxies, which have been evolved for times $\{0,2,3\} \times \tau$. It is clear that moderate amounts of orbital decay (corresponding to times $2-3\tau$) are sufficient to "erase" the core and produce a density profile similar to that found by Beers and Tonry.

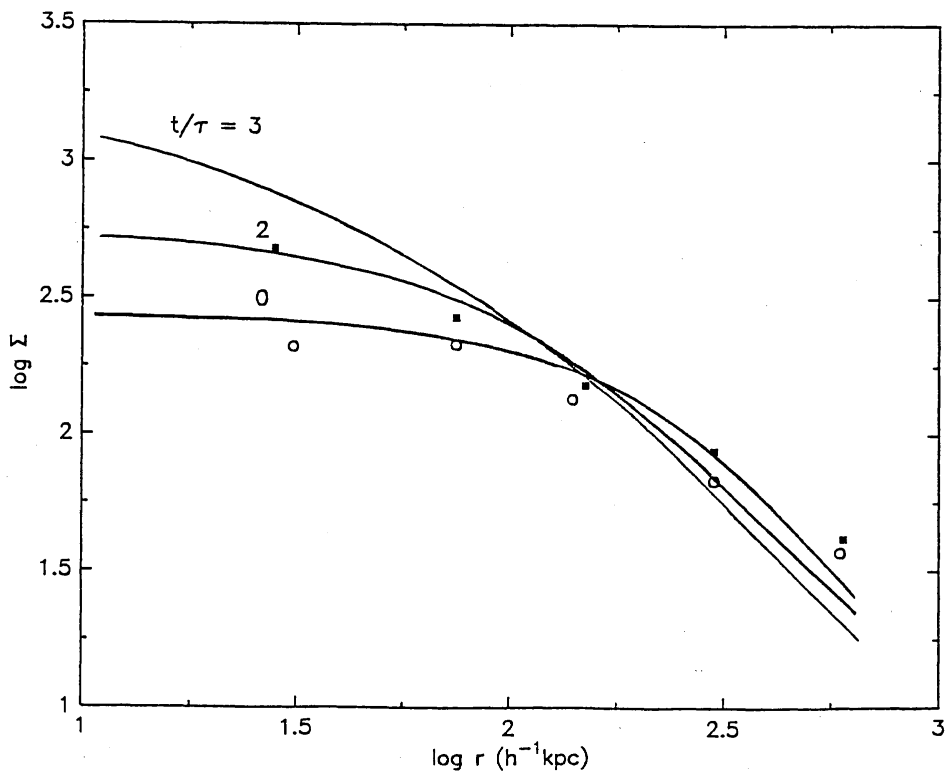

Fig. 6 Cluster surface density profiles. Squares: Beers and Tonry's (1986) combined profile, computed by superposing 36 clusters centered on the peak of their X-ray surface brightness profiles; circles: density profile of 49 clusters superposed on their median centers. Solid lines: density profiles of an initially isothermal cluster in which the galaxy orbits decay due to dynamical friction.

A third possible consequence of orbital decay, and by far the most interesting, is the formation of a very bright galaxy at the center of a cluster through the repeated accretion of other galaxies which spiral in. This is the so–called "cannibalism" model; its primary motivation is the existence, at the centers of roughly 10% of rich clusters, of very bright ($L \gtrsim 5L^*$) or "cD" galaxies. In almost every respect, the properties of cD galaxies appear to lie along a smooth continuation of the trends defined by less luminous galaxies (Tonry 1987). However, the fact that they are always located precisely at the centers of clusters, both spatially (Beers and Geller 1983) and kinematically (Quintana and Lawrie 1982), hints at a special formation process. Broadly speaking, theories which attempt to explain the central locations and high luminosities of cD galaxies by invoking mergers in cluster cores can be divided into two groups, which collectively define what might be called the "weak" and "strong" theories of cannibalism. According to the "weak" cannibalism hypothesis (Ostriker and Tremaine 1975), a massive galaxy which happens to lie near the center of a cluster will undergo a significant, although modest, increase in luminosity over a Hubble time as it accretes less massive neighbors and bound satellites. According to the "strong" cannibalism hypothesis (White 1976; Hausman and Ostriker 1978), orbital decay and merger times are sufficiently short that a superluminous galaxy will naturally form at the center of any rich relaxed cluster after about 10^{10} years. In its most extreme form (Hausman and Ostriker 1978), the "strong" hypothesis states that the sequence of cluster types, as described by Bautz and Morgan (1970) or Oemler (1974), is one of increasing dynamical evolution, the rate of evolution being fixed by the central orbital decay time.

At present these is no concensus about which (if either) of these hypotheses is correct. Numerous attempts have been made to simulate the evolution of a rich cluster after its formation (e.g. Richstone and Malumuth 1983; Malumuth and Richstone 1984; Merritt 1984a, 1985), but with very different results. The disagreement stems mostly from the fact that a galaxy cluster is a very inhomogeneous system, containing matter with a wide range of densities; thus a self–consistent approach is essentially impossible with existing computers. Instead, one is forced to represent cluster galaxies by a small set of parameters (e.g. mass, radius, orbital energy and angular momentum), which vary discontinously as a result of interactions with other galaxies and with the dark matter. The cross sections and efficiencies of the various physical processes (e.g. tidal truncation, collisional mass loss, mergers) must be specified at the outset as a function of these parameters. It is only relatively recently that accurate cross–sections for these processes have become available; most published simulations of cluster evolution have been based on extrapolations from a handful of relatively crude N–body simulations.

In the absence of a good theoretical understanding of cluster evolution, it is reasonable to ask what constraints the observations place on the cannibalism hypothesis. By far the strongest evidence for ongoing merging in rich clusters comes from the large numbers of first–ranked galaxies which are observed to contain two or more "nuclei" within a single envelope. The extra "nuclei" have traditionally been interpreted as cluster galaxies that are gravitationally bound to the central giant and in the process of merging with it. Recent studies (Hoessel 1980; Schneider, Gunn and Hoessel 1983) have shown that the probability of finding a second "nucleus" very near the center of a first–ranked galaxy is two or three times greater than would be expected from random projection against a uniform cluster core. Converting the observed overdensity into a merger rate is difficult, however, without a detailed understanding of the merger process. Figure 7 illustrates the problem with a simple model. That figure shows the evolution, due to orbital decay, of the same initially–isothermal cluster of Figure 6, except that now a central galaxy of mass $\sim 1.5 \times 10^{12} M_\odot$ has been added at the center. At early times, $t \lesssim \tau$, the number of galaxies seen close to the cluster center is enhanced, due to the additional frictional force from the central giant. At later times, however, the rate of infall of galaxies from outside the core is nearly matched by the rate at which they are "eaten", with the result that their density near the giant remains nearly constant. Although the "cannibal" continues to grow, this growth is not accompanied by an increase in the density of galaxies near to it. Also, since low–velocity galaxies are quickly accreted, the orbital velocities of the yet–unaccreted galaxies remain high, comparable to that of the other galaxies in the cluster. This simple model is consistent with the multiple–nucleus observations, both in terms of the observed overdensity, and the velocity dispersion of the nuclei (Tonry 1984). However it is clear from Figure 7 that we cannot hope to use statistics of the "nuclei" to derive the current merger rate, since the two are nearly uncoupled once a moderate amount of orbital decay has taken place.

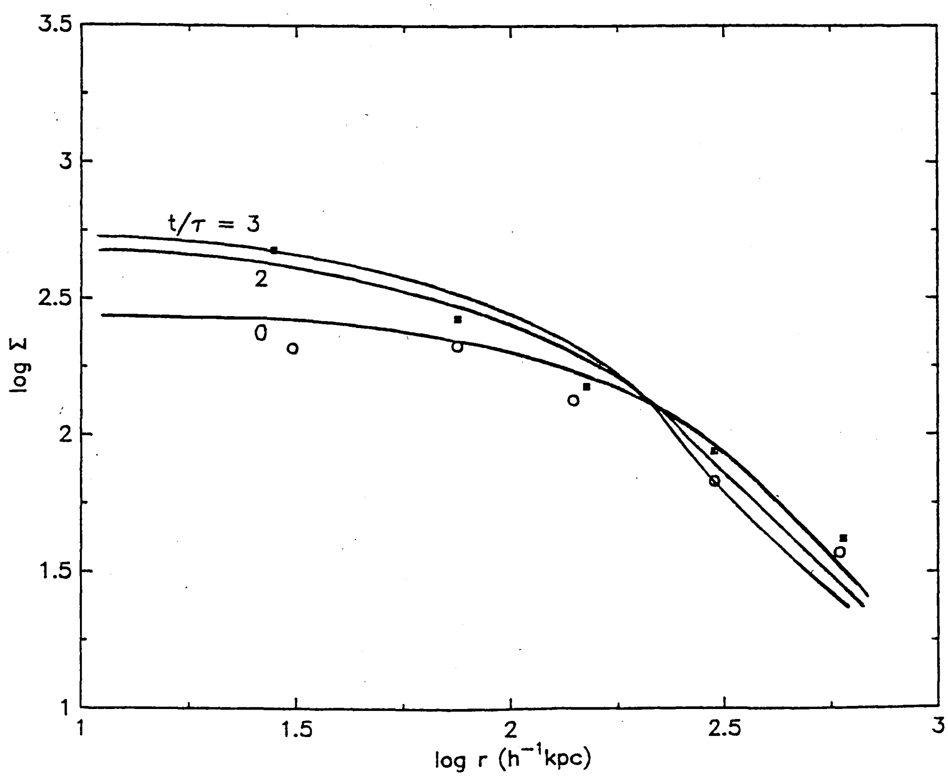

Fig. 7 Like Fig. 6, except that a $1.5 \times 10^{12} M_\odot$ galaxy has been added to the center of the model cluster.

Recently Lauer (1988) has shown a way around this problem. He observed the morphology of a set of 16 multiple–nucleus galaxies, and attempted to model each one as a superposition of normal "secondaries" seen against the brighter "primary". Roughly half of the systems could be convincingly modelled in this way; the other half showed evidence for distortions that presumably indicate tidal perturbations. Assuming that all of the "nuclei" in the latter group are in the process of being captured by their primaries, and that the typical merger time is $\sim 3 \times 10^8$yr, Lauer estimated that the current rate of growth of an average first–ranked cluster galaxy is $\sim 2L^*/5 \times 10^9$yr. This growth rate is insufficient to produce a superluminous ($L \approx 10L^*$) galaxy in a cluster that did not contain one initially, especially since, in the simple model for orbital decay outlined above, the *current* infall rate is always larger than the *time averaged* rate. Taken at face value, therefore, Lauer's result suggests that the "weak" cannibalism hypothesis is more correct, i.e., that cD galaxies grow by only moderate amounts over a cluster lifetime.

It follows that, if mergers were important in the formation of cD galaxies– and this seems *a priori* very likely, if only because cD's are so large and bright– then these mergers must have taken place at a time when their environments were very different than they are now. In fact it has been shown via N–body simulations (e.g. Barnes 1985) that galaxies in small dense groups tend to merge very quickly. Since rich clusters probably formed from the amalgamation of smaller groups (e.g. Layzer 1954), it may well be that cD galaxies acquired most of their mass and luminosity during the epoch preceding cluster formation, rather than later.

REFERENCES

Aguilar, L. A., and White, S.D.M. 1985, *Ap. J.*, **295**, 374.
Barnes, J. 1985, *M.N.R.A.S.*, **215**, 517.
Bautz, L. P., and Morgan, W. W. 1970, *Ap. J.(Letters)*, **182**, L149.
Beers, T. C., and Geller, M. J. 1983, *Ap. J.*, **274**, 491.
Beers, T. C., and Tonry, J. L. 1986, *Ap. J.*, **300**, 557.
Chandrasekhar, S. 1943, *Ap. J.*, **97**, 255.
De Lapparent, V., Geller, M.J., and Huchra, J. P. 1986, *Ap. J.(Letters)*, **302**, L1.
Gallagher, J. S., and Ostriker, J. P. 1972, *A. J.*, **77**, 288.
Hausman, M., and Ostriker, J. P. 1978, *Ap. J.*, **224**, 320.
Hoessel, J. G. 1980, *Ap. J.*, **241**, 493.
Hughes, J. P. 1988, preprint.
Kent, S.M., and Gunn, J. E. 1982, *A. J.*, **87**, 945.
Lauer, T. 1988, *Ap. J.*, **325**, 49.
Layzer, D. 1954, *A. J.*, **59**, 170.
Limber, D.N. and Mathews, W. G. 1960, *Ap. J.*, **132**, 286.
Malumuth, E. M., and Richstone, D. O. 1984, *Ap. J.*, **276**, 413.
Merritt, D. 1984a, *Ap. J.*, **276**, 26.
Merritt, D. 1984b, *Ap. J.(Letters)*, **280**, L5.
Merritt, D. 1985, *Ap. J.*, **289**, 18.
Merritt, D. 1987, *Ap. J.*, **313**, 121.

Merritt, D., and White, S.D.M. 1987, in *IAU 117, Dark Matter in the Universe*, ed. J. Kormendy and G. R. Knapp (Dordrecht: Reidel), p. 283.
Noonan, T. W. 1970, *Pub. A. S. P.*, **82**, 821.
O'Dea, C. P., Sarazin, C. L., and Owen, F. N. 1987, *Ap. J.*, **316**, 113.
Oemler, A. 1974, *Ap. J.*, **209**, 693.
Ostriker, J.P., and Tremaine, S. D. 1975, *Ap. J.(Letters)*, **202**, L113.
Pryor, C., and Geller, M. J. 1984, *Ap. J.*, **278**, 457.
Quintana, H., and Lawrie, D. W. 1982, *A. J.*, **87**, 1.
Richstone, D. O. 1976, *Ap. J.*, **204**, 642.
Richstone, D. O., and Malumuth, E. M. 1983, *Ap. J.*, **268**, 30.
Rood, H. J. 1965, Ph.D. thesis, University of Michigan.
Rood, H.J., Page, T.L., Kintner, E.C., and King, I.R. 1972, *Ap. J.*, **175**, 627.
Schechter, P. 1976, *Ap. J.*, **203**, 297.
Schneider, D. P., Gunn, J. E., and Hoessel, J. G. 1983, *Ap. J.*, **268**, 476.
The, L.S., and White, S.D.M. 1986, *A. J.*, **92**, 1248.
Thuan, T. X., and Kormendy, J. 1977, *Pub. A. S. P.*, **89**, 466.
Tonry, J. L. 1984, *Ap. J.*, **279**, 13.
Tonry, J. L. 1987, in *IAU 127, Structure and Dynamics of Elliptical Galaxies*, ed. T. de Zeeuw (Dordrecht: Reidel), p. 89.
White, S. D. M. 1976, *M.N.R.A.S.*, **174**, 19.
White, S. D. M. 1977, *M.N.R.A.S.*, **179**, 33.
White, S. D. M. 1983, *Ap. J.*, **274**, 53.
Wilson, C. P. 1975, *A. J.*, **80**, 175.
Zwicky, F. 1933, *Helvetica Physica Acta*, **6**, 110.

EXERCISE

1) Derive the virial theorem from the Jeans (or "stellar hydrodynamic") equation, as follows:
For a spherical nonrotating system (e.g., Coma cluster), the Jeans equation is

$$\frac{d(n\sigma_r^2)}{dr} + \frac{2n}{r}(\sigma_r^2 - \sigma_t^2) = -n\frac{d\Phi}{dr}$$

a) Multiply both sides by $4\pi r^3$, and integrate from zero to infinity.

b) Show that the first term becomes $-3N\langle\sigma_r^2\rangle$
Show that the second term becomes $+2N[\langle\sigma_r^2\rangle - \langle\sigma_t^2\rangle]$
Show that the third term becomes $-N\langle r d\Phi/dr\rangle$

c) Show that: $\langle\sigma_r^2 + 2\sigma_t^2\rangle = \langle v^2\rangle$ and

$$r\frac{d\Phi}{dr} = \frac{GM(r)}{r}, \qquad \text{so that}$$

$$\langle v^2\rangle = \left\langle\frac{GM(r)}{r}\right\rangle, \qquad \text{(the virial theorem).}$$

2) Use this equation to derive estimates for the <u>total mass</u> of a system, based on a sample of objects (stars, galaxies) with known $\langle v^2\rangle$ and known spatial distribution $n(r)$. Assume:

a) The mass is all in a central object, around which the "tracer" objects orbit; or

b) The mass is distributed with a constant density ρ_o out to some radius r_{max}.

3) Compute the ratio of these two estimated masses, assuming that the "tracer" population has density law

$$n(r) \propto r^{-1}, \qquad r < r_{max}$$
$$= 0 \qquad r > r_{max}$$

Comment on the usefulness of the virial theorem when nothing is known <u>a priori</u> about the form of the matter distribution.