The Origin of Central Cluster Galaxies

S. Tremaine

Canadian Institute for Theoretical Astrophysics, University of Toronto,
60 St. George Street, Toronto, Ontario M5S 1A1, Canada

Abstract   The galaxies located at the centers of clusters of galaxies are the most luminous stellar systems in the universe, and explaining the origin of these remarkable objects is a major challenge for any comprehensive theory of galaxy formation. One possibility is that central galaxies are made by a series of mergers of cluster galaxies that have spiraled to the bottom of the cluster potential well through dynamical friction. However, the resulting rate of accumulation of luminosity at the cluster center appears to be too slow—by more than a factor of two—to produce central galaxies with the luminosities observed, at least in virialized clusters similar to those we see today. A related possibility is that central galaxies are formed before the cluster virializes, during the early stages of hierarchical clustering when the relative velocities of the galaxies are low enough that merging can occur in pairwise galaxy encounters. However, so far there are no convincing numerical models of hierarchical clustering in which the most massive stellar systems formed resemble observed central galaxies. The evidence for recent merging in central galaxies is still difficult to interpret: multiple nuclei in central galaxies do not provide direct evidence for merging, as most are simply projected cluster members, but there is strong circumstantial evidence that dumbbell galaxies are the ex-central galaxies from two recently merged clusters that will themselves merge in about 0.2 Hubble times.

1. Introduction

Wo viel licht is, ist starker Schatten. (Goethe, *Götz von Berlichingen*)

The bottoms of potential wells are sometimes interesting places. For example, the potential minima at the centers of galaxies are the sites of active galactic nuclei and may host supermassive black holes. The centers of star clusters are less spectacular, but sometimes exhibit cusps in the distribution of starlight that are believed to arise from core collapse. On the other hand, the centers of open clusters or of normal stars are not believed to contain any unusual features.

The potential wells associated with clusters of galaxies are $\gtrsim 10$ times deeper than the wells associated with galaxies, and it is natural to ask whether interesting phenomena occur at the bottoms of these potential wells. Unfortunately, it is difficult to locate the potential minimum in a given cluster, for at least two reasons: (i) Clusters are dynamically young and in many cases are still accreting new material; as a result their density distribution is often clumpy and irregular (Geller and Beers 1982, Dressler and Shectman 1988), and the potential surface may have several drainage basins, each with its own local minimum. (ii) The surface density is measured by counting galaxies, and bright galaxies are so rare that the density distribution can only be determined with limited accuracy. Because of the limited statistics, observers have traditionally assumed that the surface density distribution was azimuthally symmetric, so that the cluster center could be determined using strip counts and then the density distribution could be determined using ring counts; however, many clusters are so clumpy that this procedure does not yield an accurate picture of the cluster geometry.

Our understanding of the spatial structure of clusters has improved dramatically in the last decade. The *Einstein* satellite has mapped the distribution of X-ray emitting gas in some 50 clusters, and the density distribution of this gas can be used to trace the cluster potential (Jones and Forman 1984). Meanwhile, optical observers have compiled extensive databases of galaxy positions, photometry, and radial velocities for many clusters. The locations of potential minima can be determined by finding the sites of maximum X-ray brightness or galaxy surface density (which coincide, in cases where comparison can be made; see, e.g. Gioia et al. 1982).

Given that we can locate the potential minima in clusters of galaxies, what do we find there? Jones and Forman (1984) show that in most cases, the potential minimum as measured by the maximum X-ray surface brightness coincides with the location of the brightest galaxy in the cluster core. Similarly, Beers and Geller (1983) find that the potential minimum as marked by the local maximum in galaxy surface density is often occupied by a bright galaxy. Moreover, the galaxies found at the potential minima are unusual in that they have extended luminous halos or envelopes (morphological type D or cD; see Tonry 1987 for a review of these galaxies). Over 80% of the D and cD galaxies in the sample examined by Beers and Geller are located at surface density maxima; and over 80% are the brightest galaxies in their cores. However, morphological type is a better predictor than brightness, since faint D and cD galaxies are still preferentially located at density maxima, while bright galaxies of other types are not (Beers and Geller 1983). Kinematic data confirm that D and cD galaxies tend to be located at potential minima, since their rms velocity with respect to the cluster mean is much smaller than that of typical galaxies (Quintana and Lawrie 1982, Smith et al. 1985).

According to the original definition, D galaxies are characterized by "an elliptical-like nucleus surrounded by an extensive envelope...The very large D galaxies observed in clusters are given the prefix 'c' " (Matthews, Morgan and Schmidt 1964). Unfortunately, there has never been universal agreement on the proper use of the terms D and cD. Some observers, starting with Morgan and Lesh (1965) modify the definition of a cD to include secondary properties (e.g., the cD is the brightest cluster member and is centrally located in the cluster), but strictly, these properties are neither necessary nor sufficient to identify a cD (Dressler 1984). An additional source of confusion is that most galaxies originally classified as D's turn out to be S0's (see discussion following Tonry 1987).

Despite the unsettled notation, the astronomical facts are clear, and perhaps they can best be stated by banishing the term "cD" wherever possible in this review and using the term "central galaxy" to denote a galaxy that is located at the bottom of a potential basin (as measured by maxima in the surface density of galaxies or X-ray surface brightness). Then: (i) most galaxies with extended envelopes (i.e. surface brightness profiles with relatively flat logarithmic slopes) are central galaxies; (ii) central galaxies are brighter than typical galaxies, and are usually brighter than any other galaxy in the local basin.

These striking correlations demand an explanation, and many theories for the formation of central galaxies have been advanced (see Dressler 1984 and Sarazin 1986 for reviews):

(i) The cooling time for gas in some clusters may be short enough that $\sim 10^{12}\,M_\odot$ of gas may accumulate at the cluster center over a Hubble time. Efficient conversion of the gas into stars could create a bright central galaxy. However, the colors of most central galaxies are inconsistent with a substantial present star formation rate unless most of the stars are low-mass and therefore undetectable; and in this case the mass-to-light ratio should be high, whereas in fact central galaxies have the same mass-to-light ratios as normal ellipticals.

(ii) Some special aspect of the galaxy formation process may produce bright galaxies with extended envelopes only in the high-density regions characteristic of cluster centers. A special formation process might also stimulate the production of globular clusters, which appear to be several times more numerous per unit luminosity in central galaxies relative to normal ellipticals (Harris 1987). This enhanced globular cluster population is difficult to explain in other models for the formation of central galaxies.

(iii) The central galaxy may consist of debris stripped from cluster galaxies by the cluster tidal field and by encounters with other galaxies. Stripped material may comprise the envelopes of central galaxies but the body of the galaxy must have a different origin since its velocity dispersion ($\sim 300\,\mathrm{km\,s^{-1}}$) is much smaller than the cluster dispersion ($\sim 800\,\mathrm{km\,s^{-1}}$).

(iv) The central galaxy may be made by a series of mergers of cluster galaxies that have spiraled to the bottom of the cluster potential well through dynamical friction.

These models are not necessarily mutually exclusive, since the cores and envelopes of central galaxies may be formed in different ways; also, not all central galaxies are necessarily formed by the same process. Nevertheless, for the sake of brevity, in this review I will discuss only the last model in the list, in which central galaxies are made by mergers.

2. Simple merger models

The idea that central galaxies are made by mergers dates back to papers by Lecar (1975), Ostriker and Tremaine (1975), and S. White (1976a). Although sophisticated numerical models of cluster evolution have been constructed more recently (notably by Merritt 1984a, 1985, Richstone and Malumuth 1983, and Malumuth and Richstone 1984), analytic models still provide a good introduction to the theory. In deriving these, I shall refer to the book by Binney and Tremaine (1987, hereafter BT) for some of the more standard results.

I first consider the dynamical evolution of galaxies after the cluster core has reached a quasi-steady state following collapse and virialization; evolution during cluster collapse is discussed in the next section.

A massive galaxy orbiting in a cluster is subject to a drag force arising from its gravitational interactions with the other mass in the cluster. This drag force is called "dynamical friction" (Chandrasekhar 1943; see also BT). For a galaxy of mass m travelling at speed $\mathbf{v}$ through an infinite homogeneous medium of density ρ, the drag force per unit mass is

$$\mathbf{a}_d \equiv \frac{d\mathbf{v}}{dt} = -\frac{4\pi G^2 m \rho \ln \Lambda F(v)}{v^3}\mathbf{v}. \tag{1}$$

Here $F(v)$ is the fractional mass of objects in the medium with speeds less than $v = |\mathbf{v}|$, and $\Lambda = b_{\max}/b_{\min}$, where $b_{\max}$ and $b_{\min}$ are the maximum and minimum impact parameters of encounters contributing to the drag. The nature of the objects in the medium is irrelevant so long as they are both collisionless and of mass much less than m (so that there is no stochastic acceleration of the galaxy due to individual encounters).

Although (1) was derived for a homogeneous medium, it can also be used to compute the decay of the orbit of a galaxy in a cluster, with ρ now denoting the local density at the galaxy's position (I assume that most of the cluster mass is in objects with mass much less than a galaxy's). In this context, (1) has been tested against many numerical simulations and usually gives remarkably accurate predictions for the rate of orbital decay (S. White 1978, Lin and Tremaine 1983, Bontekoe and van Albada 1987, Zaritsky and S. White 1988).

I will model the cluster density distribution as an isothermal sphere (see BT), since this provides a satisfactory fit to the number density distribution in the inner parts of rich clusters. The isothermal sphere is specified by two parameters, the one-dimensional rms velocity dispersion σ, which is independent of radius, and the core radius r_c, which measures the extent of the constant-density core region.

The median dispersion of clusters containing a D or cD galaxy is $650\,\mathrm{km\,s^{-1}}$, similar to the dispersion for rich clusters generally (Geller 1988), but since I am concerned with the central rather than the average dispersion over the cluster I shall use the somewhat higher value $\sigma = 800\,\mathrm{km\,s^{-1}}$. The core radius r_c is less certain. Bahcall (1975) estimates $r_c = 125\pm20h^{-1}$ kpc (h is the Hubble constant H_0 in units of $100\,\mathrm{km\,s^{-1}}\,\mathrm{Mpc^{-1}}$), while Dressler (1978) estimates $r_c = 240 \pm 60h^{-1}$ kpc, almost a factor of two larger. Much of the uncertainty arises because it

is difficult to locate the exact cluster center. Beers and Tonry (1986; see also Yahil 1974) stress that centering errors usually lead to a measured core radius that is too large. They argue that in most clusters there is no evidence for any non-zero core radius, and show that in clusters whose centers are well-defined, by the presence of a cD galaxy or a local maximum in the X-ray surface brightness, the surface density of galaxies increases as r^{-1} right into the cluster center. (Indirect evidence from the frequency of multiple nuclei in central galaxies [see Lauer 1989 and §3] and from models of clusters as gravitational lenses [Grossman and Narayan 1989] also suggests that the core radius is much smaller than the Bahcall or Dressler estimates.)

The power-law density profiles found by Beers and Tonry and by Yahil closely resemble the profile of the singular (i.e. $r_c = 0$) isothermal sphere. Since the singular isothermal has the additional benefit of analytic simplicity, I shall adopt it as a crude model of the mass distribution in a cluster. The density, potential, enclosed mass, and fraction of objects with speed $< v$ are given by (see BT):

$$\rho(r) = \frac{\sigma^2}{2\pi G r^2}, \quad \Phi(r) = 2\sigma^2 \ln r + \text{const}, \quad M(r) = \frac{2\sigma^2 r}{G}, \quad F(v) = \text{erf}(x) - x\,\text{erf}'(x), \quad (2)$$

where $x = v/\sqrt{2}\sigma$ and erf denotes the error function.

I assume for simplicity that the galaxies move on circular orbits. Then the rate of change of orbital radius is related to the specific angular momentum $J(r) = \sqrt{rd\Phi/dr}$ and the drag force (eq. 1) by $dr/dt = dJ/dt(dJ/dr)^{-1} = -r|\mathbf{a}_d|(dJ/dr)^{-1}$. Specializing to the singular isothermal sphere and using (1) and (2) yields

$$\frac{dr}{dt} = -\frac{\text{erf}(1) - \text{erf}'(1)}{\sqrt{2}} \cdot \frac{Gm}{\sigma r} \ln \Lambda = -0.302 \frac{Gm}{\sigma r} \ln \Lambda. \quad (3)$$

The decay rate is proportional to the galaxy mass and hence is affected by the extent of massive halos in the cluster galaxies. I shall examine two limiting cases and compute for each the rate at which galaxies spiral into the center of the cluster:

(i) No massive halo   I assume that the galaxy mass-to-light ratio Υ is independent of radius and equal to its value in the core of a typical nearby elliptical, $\Upsilon = 12h\Upsilon_\odot$ in the V band (cf. Lauer 1985). Tidal stripping by the cluster potential is small since the galaxy's mass is concentrated at small radii; thus I assume m is time-independent. I shall also neglect changes in $\ln \Lambda$ as the galaxy spirals in. Then (3) yields the initial orbital radius of galaxies that spiral into the cluster center in a time t:

$$r^2(t) = 0.604\frac{Gmt}{\sigma} \ln \Lambda = (63h^{-1} \text{ kpc})^2 \left(\frac{800 \text{ km s}^{-1}}{\sigma}\right) \left(\frac{\Upsilon}{12h\Upsilon_\odot}\right) \left(\frac{L}{L_*}\right) (H_0 t) \ln \Lambda, \quad (4)$$

where I have written the galaxy luminosity L in units of the characteristic luminosity $L_* = 1.0 \times 10^{10} h^{-2} L_\odot$.

Taking a Schechter luminosity function, and assuming that the density distribution of the galaxies is the same as that of the cluster mass, I can write the number density of galaxies per unit luminosity and volume in the form

$$n(L,r)dLdr = n_*(r)dr \exp\left(-\frac{L}{L_*}\right) \left(\frac{L}{L_*}\right)^\alpha \frac{dL}{L_*}, \quad (5)$$

where $\alpha = -1.25$ (e.g., Colless 1989). I assume that the cluster mass-to-light ratio Υ_c (only the light in galaxies is counted) is independent of position, so that $n_*(r) \propto \rho(r)$; thus

$$\frac{\rho(r)}{n_*(r)} = \Upsilon_c L_*(\alpha + 1)!. \quad (6)$$

Combining (2), (4), (5) and (6) gives the total luminosity accreted to the cluster center:

$$\frac{L_{acc}}{L_\star} = 2\frac{(\alpha + \frac{3}{2})!}{(\alpha + 1)!}\Upsilon_c^{-1}\left(\frac{0.604\sigma^3\Upsilon t \ln\Lambda}{GL_\star}\right)^{1/2}$$

$$= 3.1\left(\frac{360h\Upsilon_\odot}{\Upsilon_c}\right)\left(\frac{\Upsilon}{12h\Upsilon_\odot}\right)^{1/2}\left(\frac{\sigma}{800\,\mathrm{km\,s^{-1}}}\right)^{3/2}\left(\tfrac{3}{2}H_0t\right)^{1/2}(\ln\Lambda)^{1/2}. \tag{7}$$

The comparison value $360h\Upsilon_\odot$ used for the cluster mass-to-light ratio is the value obtained by Kent and Gunn (see Kent and Sargent 1983) for the Coma cluster, which is typical of rich clusters. The factor $\tfrac{3}{2}H_0t$ is unity if the cluster formed shortly after the Big Bang in a critical ($\Omega = 1$) Friedmann universe.

The appropriate value of $\ln\Lambda = \ln(b_{max}/b_{min})$ may be estimated by setting b_{max} to the typical orbital radius, here $60h^{-1}$ kpc by (4), and b_{min} to the median radius of a typical elliptical, $r_h \simeq 5h^{-1}$ kpc (see Bontekoe and van Albada 1987). (These values are approximate but only enter (7) in the square root of a logarithm.) Thus $\ln\Lambda = 2.5$ and I find

$$\frac{L_{acc}}{L_\star} = 4.9\left(\frac{360h\Upsilon_\odot}{\Upsilon_c}\right)\left(\frac{\Upsilon}{12h\Upsilon_\odot}\right)^{1/2}\left(\frac{\sigma}{800\,\mathrm{km\,s^{-1}}}\right)^{3/2}\left(\tfrac{3}{2}H_0t\right)^{1/2}. \tag{8}$$

(ii) Massive halo Tidal forces from the cluster place a strong constraint on the extent of the massive halo (Gunn 1977, Merritt 1984a). Following Merritt, I locate the tidal radius r_t at the collinear Lagrange points (the locations where the effective potential—cluster plus galaxy plus centrifugal potential from the orbital motion—has a saddle point). For a galaxy of mass m on a circular orbit of radius r in a cluster potential given by (2) I find

$$r_t^3 = \frac{Gmr^2}{4\sigma^2}. \tag{9}$$

I approximate the galaxy as a singular isothermal sphere (2) with dispersion σ_g, which I identify with the central line-of-sight dispersion in the galaxy. I then write the galaxy mass as a fraction $f < 1$ of the mass contained in the isothermal sphere within the tidal radius r_t. (This is a plausible approximation since the dispersion is likely to decrease outwards, so that the actual galaxy mass inside any radius will be less than the mass in the isothermal sphere.) Thus

$$m = \frac{2\sigma_g^2 f r_t}{G}. \tag{10}$$

The fraction f is uncertain but a plausible value is $f = 0.5$. Then (9) and (10) yield

$$m = 2^{1/2}f^{3/2}\frac{\sigma_g^3 r}{G\sigma}, \qquad r_t = 2^{-1/2}f^{1/2}\frac{\sigma_g}{\sigma}r. \tag{11}$$

Then (3) yields the initial radius of galaxies that spiral into the cluster center in a time t:

$$r(t) = 0.428f^{3/2}\frac{\sigma_g^3 \ln\Lambda t}{\sigma^2} = 25h^{-1}\,\mathrm{kpc}\left(\frac{800\,\mathrm{km\,s^{-1}}}{\sigma}\right)^2\left(\frac{L}{L_\star}\right)^{0.75}\left(\frac{f}{0.5}\right)^{3/2}(H_0t)\ln\Lambda, \tag{12}$$

where I have used the empirical Faber-Jackson law,

$$\sigma_g = \sigma_\star\left(\frac{L}{L_\star}\right)^{0.25}, \qquad \text{with} \quad \sigma_\star = 220\,\mathrm{km\,s^{-1}}. \tag{13}$$

Using (2), (5), and (6), I find the accreted luminosity to be

$$\frac{L_{\rm acc}}{L_\star} = 0.855 f^{3/2} \frac{(\alpha + 0.75)!}{(\alpha + 1)!} \Upsilon_c^{-1} \left(\frac{\sigma_\star^3 t \ln \Lambda}{G L_\star} \right)$$

$$= 2.0 \left(\frac{f}{0.5} \right)^{3/2} \left(\frac{360 h \Upsilon_\odot}{\Upsilon_c} \right) \left(\tfrac{3}{2} H_0 t \right) \left(\frac{\sigma_\star}{220 \, {\rm km \, s^{-1}}} \right)^3 \ln \Lambda. \tag{14}$$

I shall take $\ln \Lambda = \ln(r/0.5 r_t) = \ln(2^{1/2} \sigma / 0.5 f^{1/2} \sigma_g)$ since the median radius is probably about half the tidal radius; for the standard values of the parameters $\ln \Lambda = 2.7$ and I have

$$\frac{L_{\rm acc}}{L_\star} = 5.4 \left(\frac{f}{0.5} \right)^{3/2} \left(\frac{360 h \Upsilon_\odot}{\Upsilon_c} \right) \left(\tfrac{3}{2} H_0 t \right) \left(\frac{\sigma_\star}{220 \, {\rm km \, s^{-1}}} \right)^3. \tag{15}$$

There are a number of uncertainties in the estimates (8) and (15) of the central luminosity produced by dynamical friction.

(i) The assumption of circular orbits is unrealistic. Our estimates of $L_{\rm acc}$ should probably be increased by about 10% to account for the fact that the galaxies are on isotropic rather than circular orbits. On the other hand, the assumption of circular orbits minimizes the mass stripping due to tidal forces, which are strongest at pericenter.

(ii) I have assumed that the luminous mass in the galaxy is concentrated at sufficiently small radii that it is not stripped by tidal forces. To check this, I use (9) to evaluate the orbital radius r_s at which the tidal radius r_t equals the median radius r_h for a galaxy with mass $m = \Upsilon L_\star$, $\Upsilon = 12 h \Upsilon_\odot$, and dispersion $\sigma_\star = 220 \, {\rm km \, s^{-1}}$. I find $r_s = 7 h^{-1}$ kpc, small enough that stripping of the luminous mass should be unimportant. However, some stripping may have occurred—Strom and Strom (1979) find galaxies near cluster centers to be smaller than field galaxies—which would reduce the accreted luminosity.

(iii) Mass loss due to galaxy-galaxy collisions is probably unimportant (for rate estimates see Richstone 1975, Merritt 1983, 1984a), but galaxies on elongated orbits may suffer mass loss at pericenter due to impulsive tidal shocks from the central part of the cluster mass distribution (Aguilar, Hut and Ostriker 1988).

(iv) The fraction f in (10) is uncertain and may be less than 0.5; for example Merritt (1984a) argues for $f = 0.25$ (his α is $2\sqrt{f}$).

(v) In the model described here, central galaxies are formed from a small number of accreted cluster galaxies, and hence there is a statistical uncertainty in the luminosity of the central galaxy expected in a given cluster. Assuming that the luminosities and initial positions of the cluster galaxies are statistically independent, it is straightforward to show that the rms fractional uncertainty in $L_{\rm acc}$ is $[(\alpha + \tfrac{5}{2}) L_\star / L_{\rm acc}]^{1/2}$ if no massive halos are present, and $[(\alpha + 2.75) L_\star / L_{\rm acc}]^{1/2}$ if massive halos are present. Using the nominal parameters in (8) and (15) respectively, the fractional uncertainties are 0.5 in both cases.

(vi) Probably the main uncertainty in $L_{\rm acc}$ arises from the uncertain cluster core radius. Galaxy counts (Beers and Tonry 1986) yield core radii that are consistent with zero, but indirect arguments based on the frequency of multiple nuclei in central galaxies (see Lauer 1989 and §3 below) suggest that $r_c \gtrsim 20 h^{-1}$ kpc, much smaller than the estimates of Bahcall (1975) and Dressler (1978) ($r_c = 0.13$ and $0.24 h^{-1}$ kpc respectively), but still large enough to reduce substantially the number of bright galaxies accreted (eqs. 4 and 12). Moreover, it is likely that the core radius of the dark matter is larger than that of the galaxies, since it is not subject to dissipation or dynamical evolution; hence its density will be lower than in the singular isothermal and the drag will be reduced.

The conclusions from these simple models are consistent with more careful calculations made by Merritt (1984a, 1985). Using parameters similar to those in our "massive halo" calculation, except that the cluster core radius is $r_c = 120$ kpc, Merritt (1984a, Table 2) found that the luminosity accumulated in the core due to dynamical friction was only about $0.1 L_\star$.

This is substantially smaller than our estimate of $5L_*$ (eq. 15), and confirms that the accreted luminosity is a strongly decreasing function of core radius.

The typical luminosity of a cD galaxy is about $13L_*$ (this is the median luminosity of a sample of 27 cD's from Schombert 1988, with the luminosity of the extended envelope removed, since it may have a different origin from the main body of the galaxy). Therefore the estimates (8) and (15) of the luminosity accreted onto the cluster center are more than a factor of two smaller than the actual luminosity of central galaxies, and the most important of the uncertainties described above are likely to make more realistic estimates smaller still. A similarly small estimate, $L_{acc} \simeq 2L_*$, has been obtained by Lauer (1988) using the fraction of close neighbors to the central galaxy that exhibit evidence of tidal interactions.

Thus, it appears that dynamical evolution of galaxies in a quasi-static, virialized cluster core similar to those we see today is not strong enough to produce central galaxies with luminosities comparable to those observed. This conclusion is originally due to Merritt (1984a, 1985), whose models are much more sophisticated than the ones here; the principal contribution of the present calculations is to show that Merritt's conclusion remains valid whatever the core radius may be.

3. Mergers during cluster formation

The formation of galaxies and clusters is likely to occur through hierarchical clustering (S. White and Rees 1978, Peebles 1980). The early universe contains density fluctuations on a wide range of scales. As the universe expands, bound density fluctuations eventually "break away" from the Hubble flow and collapse. The strongest density perturbations are on small scales, so that small systems generally collapse first; larger bound systems have lower overdensities and collapse only after the smaller systems they contain. Thus, small systems (e.g. galaxies, or, more properly, dark matter halos of galaxies) develop before larger ones (groups and clusters of galaxies). An important feature of hierarchical clustering is that most or all of these small subunits are generally disrupted by tidal forces during the collapse of the next larger structure in the hierarchy; thus, for example, the dark matter in a group or cluster of galaxies will be smoothly distributed rather than concentrated in distinct galaxy halos.

In hierarchical clustering models, the luminous material in galaxies forms at the bottoms of the potential wells of individual galaxy halos. We must therefore ask why the individual galaxies in a cluster have survived, when the halos have been disrupted. One possibility is that dissipation concentrates the luminous material so that its density is high enough to survive disruption (S. White and Rees 1978). Alternatively, it may be possible for the dense cores of the individual halos to survive, with luminous material at the center of each dark core. (These two alternatives correspond crudely to the two simple models—no massive halo and tidally limited halo—that were described in the previous section.)

Merritt (1985) has pointed out that since cD galaxies are found in groups as well as rich clusters (R. White 1978 and references therein) it is natural to ask whether the cD's in rich clusters might have formed while the cluster members were still distributed in groups.

The simple merger models of the previous section do not suggest that groups are more likely to form cD's than clusters: groups have lower velocity dispersion than clusters, and the accreted luminosity L_{acc} varies as $\sigma^{3/2}$ in models with no massive halos (eq. 8) and is independent of σ in models with halos (eq. 15). However, when the cluster dispersion σ becomes as small as the internal galaxy dispersion σ_g, a new effect appears: mergers can occur during galaxy-galaxy collisions at any location in the cluster. (The reason follows from the impulse approximation. In a high-speed head-on collision between two galaxies, the energy converted from relative motion of the two galaxies to internal energy is of order $\Delta E \approx \sigma_g^4/\sigma^2$ per unit mass, while the relative kinetic energy before the encounter is of order $E \approx \sigma^2$ per unit mass. Mergers can occur once $|\Delta E| \gtrsim E$, which requires $\sigma \lesssim \sigma_g$; of course in this limit the impulse approximation is not really valid, but it still gives the correct order-of-magnitude result.) A large merger remnant

can be built up more rapidly by direct mergers between galaxies throughout the cluster than by merging only galaxies that have spiraled in to the cluster center.

Numerical simulations provide dramatic confirmation of the importance of merging in small groups (e.g. Ishizawa et al. 1983, Barnes 1989). These simulations typically follow a group of 5–10 spherical or disk galaxies. They find that within a few group crossing times most or all of the galaxies have merged to form a single amorphous system. Moreover, the material at the centers of the individual galaxies is concentrated at the center of the remnant system, suggesting that the luminous material will collect in a single structure at the center of the merged halos.

Thus, mergers in the early stages of hierarchical clustering can produce single galaxies that resemble observed central galaxies. However, the initial conditions for the simulations described above, though plausible, are not derived self-consistently by following the growth of linear perturbations with a given fluctuation spectrum. The challenge to numerical models of hierarchical clustering and galaxy formation is to yield enough mergers at early stages to produce $10 - 15L_*$ galaxies and yet not so much that an entire rich cluster merges into a single enormous starpile.

4. Multiple nuclei and dumbbell galaxies

A remarkably large fraction of central galaxies have multiple nuclei or superimposed secondary galaxies, a fact already noted by Morgan and Lesh (1965). Hoessel and Schneider (1985) found that about 50% of brightest cluster galaxies had multiple nuclei brighter than about $0.05L_*$ within an aperture of $10h^{-1}$ kpc radius; in contrast, the frequency of multiple nuclei in second or third brightest cluster galaxies is about an order of magnitude smaller (Schneider, Gunn and Hoessel 1983). The luminosity function of the secondary galaxies is similar to the luminosity function of cluster galaxies in general (Hoessel and Schneider 1985, Lauer 1989).

This concentration of galaxies near the central galaxy has often been cited as evidence that the central galaxy is presently consuming cluster galaxies at a rapid rate. However, I have never found this argument convincing. Consider a model in which the galaxies move on circular orbits, the galaxy number density is $n(r, t)$, and the rate of orbital decay is $dr/dt < 0$. Then the continuity equation implies that orbital decay leads to an increase in the number density ($\partial n/\partial t > 0$) if and only if $r^2 n |dr/dt|$ is increasing outwards. In a singular isothermal, $n \propto r^{-2}$, and $|dr/dt| \propto r^{-1}$ (for constant mass galaxies); hence the number density is depleted by friction. In a central core, $n \propto const$ and $|dr/dt| \propto r$ (see Merritt 1985); thus the number density is enhanced. However, the central galaxy is a strong source of drag (its density is higher and velocity dispersion lower than the cluster's) so that near the center the drag increases rapidly. If the increase is rapid enough so that $|dr/dt|$ increases inwards faster than r^{-2}, then the density will be depleted. Thus we expect an enhancement in the cluster galaxy density inside the core due to friction, but it is much less certain that there will be an enhancement in density within the central galaxy itself.

Recent observational evidence is consistent with this argument, in that it now appears that most multiple nuclei are not associated with the central galaxy at all, but are simply cluster galaxies projected in front of the central galaxy:

(i) The velocity dispersion of the nuclei is similar to that of the cluster (Tonry 1985, Smith et al. 1985 and references therein), which is expected if their orbits are those of typical cluster members, and inconsistent with orbits bound to the central galaxy, which has a much lower dispersion (e.g. Tonry 1984).

(ii) The numbers of multiple nuclei are about four times the number expected from chance projections for an assumed core radius of $150h^{-1}$ kpc (Schneider et al. 1983). However, Beers and Tonry (1986) argue that the actual core radii of clusters are quite uncertain and may be much smaller; assuming a simple model for the surface density

$$\Sigma(r) = \frac{\Sigma_0}{[1 + (r/r_c)^2]^\gamma}, \tag{16}$$

then for a given asymptotic surface density profile the central surface density varies as $\Sigma_0 \propto r_c^{-2\gamma}$. Thus, a core radius as small as $r_c = 75h^{-1}$ kpc ($\gamma = 1$) or $40h^{-1}$ kpc ($\gamma = \frac{1}{2}$) would imply that all the multiple nuclei could be chance projections.

Lauer (1989) has compared the total galaxy luminosities enclosed within $10h^{-1}$ kpc and $50h^{-1}$ kpc of the central galaxy center (excluding the central galaxy itself). If the multiple nuclei are due to chance projections and the core radius $r_c \gg 50h^{-1}$ kpc then the ratio of these luminosities should be 25. The ratios found by Lauer were 23 ± 8 (richness class 2), 13 ± 3 (class 1), and 8 ± 2 (class 0). These values are consistent with the hypothesis that all the multiple nuclei are chance projections if $r_c \gg 50h^{-1}$ kpc (richness class 2), $\simeq 30h^{-1}$ kpc (class 1), and $\simeq 20h^{-1}$ kpc (class 0). (I assume a surface density given by [16] with $\gamma = 1$.)

(iii) Lauer (1988) finds that roughly half of the nuclei show morphological disturbances not seen in isolated galaxies, implying recent direct interactions. To see whether this fraction is consistent with the assumption that the secondaries are simply typical cluster members, I use the surface brightness distribution (16) with $\gamma = 1$. This implies a space density distribution $n(r) \propto [1 + (r/r_c)^2]^{3/2}$, which implies that the fraction of galaxies projected on a central aperture of radius $r_{\rm ap} \ll r_c$ that are actually within a distance $r_{\rm int}$ small enough to be interacting is $f = r_{\rm int}/\sqrt{r_c^2 + r_{\rm int}^2}$. Assuming that the tidal disturbances damp on a crossing time $r_{\rm ap}/(\sqrt{3}\sigma_g)$, and that the typical encounter speed is $\sqrt{3}\sigma$, then $r_{\rm int} \approx r_{\rm ap}(\sigma/\sigma_g)$. Using the nominal parameters $\sigma_g = 220 \, {\rm km \, s^{-1}}$, $\sigma = 800 \, {\rm km \, s^{-1}}$, $r_{\rm ap} = 10h^{-1}$ kpc, $r_c = 75h^{-1}$ kpc, I find $r_{\rm int} = 36h^{-1}$ kpc and $f = 0.4$, consistent with the observed value of about 0.5.

(iv) Merritt (1984b) has pointed out that small core radii may be a natural consequence of dynamical friction, which tends to shrink the core radius. This may also provide a natural explanation for the smaller core radii deduced in (ii) for richness class 0 or 1 compared to richness class 2: the rate of shrinkage is likely to be faster in clusters of lower richness because their dispersion is lower and $dr/dt \propto \sigma^{-1}$ (eq. 3).

Thus multiple nuclei in central galaxies provide no direct support for the hypothesis that these galaxies have been made by mergers. The nuclei can be explained as typical cluster galaxies that happen to be projected on top of the central galaxy.

A related but distinct class is the binary supergiant or "dumbbell" (db) galaxies. Matthews et al. (1964) described the dumbbells as "a group allied to the D galaxies, in which two separated, approximately equal, nuclei are observed in a common envelope." In most dumbbells both components are very luminous—often brighter than any other galaxies in the cluster—and appear to belong to the class of D or cD galaxies.

There are no systematic surveys for dumbbell galaxies, although Leir and van den Bergh's (1977) study of the Abell catalog of rich clusters noted all brightest cluster galaxies that were "double or multiple", and Rood and Leir (1979) state that most of the galaxies classified as "multiple" in this way are dumbbells. They also stress that the fraction of clusters containing a multiple first-ranked galaxy is remarkably high—23% for Bautz-Morgan type I and I–II clusters.

Valentijn and Casertano (1988) have compiled a list of 44 clusters containing dumbbell galaxies. For most of these the magnitude difference between the components was less than 1 magnitude. (Rood 1988 estimates a median difference of 0.5 magnitude in the Rood and Leir sample.) The median separation of the centers of the two components is $10h^{-1}$ kpc and only 3 of the dumbbells have separations $\gtrsim 25h^{-1}$ kpc. The high frequency of dumbbells with such small separations is a striking result, although the absence of dumbbells with large separations is largely an artifact—by definition, "dumbbells" cannot have large separations.

The rms velocity difference between the dumbbell components in the Valentijn-Casertano sample is $650\,\mathrm{km\,s^{-1}}$, implying a single-object dispersion of $650/\sqrt{2} = 460\,\mathrm{km\,s^{-1}}$ if the velocities of the two components are uncorrelated. This is substantially smaller than the typical cluster dispersion of $800\,\mathrm{km\,s^{-1}}$, and smaller than the dispersion of multiple nuclei, which is similar to the cluster dispersion. The mean velocity difference increases with separation.

There is a simple picture for the formation of dumbbell galaxies that is consistent with much of this data. Consider two clusters that collapse and virialize, then fall towards each other and merge into a single cluster. Such structures arise commonly in cosmological simulations (e.g. S. White 1976b, Cavaliere et al. 1986) and strongly resemble structures that can be identified in both X-ray observations ("double clusters", see Forman et al. 1981) and galaxy counts (Geller and Beers 1982; see Geller 1988 for a review).

Now suppose that each of the two merging clusters had a central galaxy. The central galaxies will remain surrounded by their respective cores until the two cores merge. Thus the central galaxies, like most of the other collisionless material in the two original cores, will end up orbiting in or near the core of the merged cluster. (In other words, violent relaxation during a merger of this kind is not complete: material with high binding energy in the merging systems will have high binding energy in the merger remnant.)

This argument suggests that clusters containing two D or cD galaxies are the product of a merger of two clusters in which these were once the central galaxies, and that in the merged cluster the two galaxies will be separated by at most a couple of core radii, that is, by $\lesssim 200 h^{-1}$ kpc. It remains to be shown why in most observed clusters the two D's or cD's are separated by a much smaller distance, $\lesssim 20 h^{-1}$ kpc.

Once the ex-central galaxies are orbiting in the core of the merger remnant, dynamical friction will drag them toward the center of the remnant, where they will eventually merge to form a new central galaxy. The radius r of a circular orbit within the core decays exponentially, $r \propto \exp(-t/t_d)$, where (Merritt 1985)

$$t_d = \frac{\sqrt{2\pi}\, r_c^2 \sigma}{3Gm\ln\Lambda} = 0.44 \left(\frac{2}{3H_0}\right)\left(\frac{r_c}{75h^{-1}\,\mathrm{kpc}}\right)^2\left(\frac{\sigma}{800\,\mathrm{km\,s^{-1}}}\right)\left(\frac{L_*}{L}\right)\left(\frac{12h\Upsilon_\odot}{\Upsilon}\right), \qquad (17)$$

and I have assumed $\ln\Lambda = 2.5$. I shall concentrate on the less luminous ex-central galaxy, since the more luminous (and presumably more massive) of the pair will settle to the center first and await its partner. Assuming a typical luminosity ratio $L_>/L_< = 1.6$ (i.e. 0.5 magnitude) and $L_< = 5L_*$ (this gives a merged central galaxy with a typical cD luminosity $L_> + L_< = 13L_*$), the e-folding time $t_d \simeq 0.1t_u$, where $t_u = \frac{2}{3}H_0^{-1}$ is the age of the universe. Probably of order one or two e-foldings are required before the ex-central galaxies reach the separation $r_{\mathrm{crit}} \approx 15h^{-1}$ kpc where their mutual drag begins to dominate the drag from the cluster core. Once this happens the two galaxies will rapidly merge. Since we have overestimated the drag somewhat by using the central density (even at one core radius the density is only 0.35 of the central value), it is likely that the two galaxies merge only after a total time $t_{\mathrm{tot}} \simeq 0.3t_u$.

This model explains many of the features of dumbbell galaxies. (i) The dumbbell components are D or cD galaxies because they used to be central galaxies. (ii) The strong concentration towards small separations arises because the frictional decay rate $dr/dt \propto r$ inside the core but outside r_{crit}; in a steady state this implies that the number density $n(r) \propto r^{-3}$ for $r \gtrsim r_{\mathrm{crit}}$. (iii) The rms velocity is lower than the cluster value because the dumbbell components are orbiting within the cluster core; within the core $v \propto r$, which explains why the velocities increase with separation. (iv) To estimate whether the numbers of dumbbell galaxies are consistent with the model, assume that there is a steady flux F of ex-central galaxies across radius r; then $F = 4\pi r^2 n(r)|dr/dt|$. Then the total number of such galaxies at a given time is $N = Ft_{\mathrm{tot}}$ and the total number of mergers over the age of the universe is $N_m = Ft_u$, assuming that the flux has been constant. Taking $t_{\mathrm{tot}} = 0.3t_u$, I find $N_m \approx 3N$. This is probably an underestimate,

since the flux was surely higher in the past. For comparison, Rood and Leir (1979) estimate that 23% of Bautz-Morgan type I and I-II clusters have brightest galaxies that are multiple, and most of these are dumbbells; if these clusters have undergone an average of one transient dumbbell stage then $N_m = 4N$, consistent with the crude theoretical prediction.

Using similar arguments, Rood (1988) has reached a quite different conclusion: that the number of dumbbells is at least five times that predicted by simple hierarchical clustering models. The reason for the discrepancy is that Rood uses a much shorter decay time, $t_{tot} \simeq 3 \times 10^8 h^{-1}$ yr $= 0.05 t_u$. This is the decay time inside r_{crit}, where the drag is dominated by the galaxies; however, the steady-state population of dumbbells is mostly determined by the slowest part of the decay process, the decay through the cluster core outside r_{crit}, for which the much longer decay time of $0.3 t_u$ used above is the appropriate one.

The Coma cluster may be an example of a cluster that will eventually contain a dumbbell. Coma has two galaxies with many cD characteristics (Schombert 1988), with a present separation of $180 h^{-1}$ kpc, and the distribution of galaxies both on the sky and in velocity space suggest that it consists of two merging subclusters (Fitchett and Webster 1987 and references therein).

It has often been argued that the presence of multiple nuclei in central galaxies provides strong evidence that merging is an important and ongoing process, while the existence of dumbbell galaxies suggests that dynamical friction and merging are less effective than theory would lead us to believe. It is interesting and a little ironic that present observations suggest strongly that multiple nuclei can be explained as nothing more than a projection effect, while it is the dumbbell galaxies that provide the most direct evidence that merging of giant galaxies is a common process in the centers of clusters.

5. Summary

The principal conclusions are summarized in the abstract.

References

Aguilar, L., Hut, P., and Ostriker, J. P. 1988. *Astrophys. J.* **335**, 720.

Bahcall, N. 1975. *Astrophys. J.* **198**, 249.

Barnes, J. 1989. *Nature* **338**, 123.

Beers, T. C., and Geller, M. J. 1983. *Astrophys. J.* **274**, 491.

Beers, T. C., and Tonry, J. L. 1986. *Astrophys. J.* **300**, 557.

Binney, J., and Tremaine, S. 1987. *Galactic Dynamics* (Princeton: Princeton University Press).

Bontekoe, Tj. R., and van Albada, T. S. 1987. *Mon. Not. Roy. Astron. Soc.* **224**, 349.

Cavaliere, A., Santangelo, P., Tarquini, G., and Vittorio, N. 1986. *Astrophys. J.* **305**, 651.

Chandrasekhar, S. 1943. *Astrophys. J.* **97**, 255.

Colless, M. 1989. *Mon. Not. Roy. Astron. Soc.* **237**, 799.

Dressler, A. 1978. *Astrophys. J.* **226**, 55.

Dressler, A. 1984. *Ann. Rev. Astron. Astrophys.* **22**, 185.

Dressler, A., and Shectman, S. A. 1988. *Astron. J.* **95**, 985.

Fitchett, M., and Webster, R. 1987. *Astrophys. J.* **317**, 653.

Forman, W., Bechtold, J., Blair, W., Giacconi, R., Van Speybroeck, L., and Jones, C. 1981. *Astrophys. J. Lett.* **243**, L133.

Geller, M. J. 1988. In *Large-Scale Structure in the Universe*, 17th Advanced Course of the Swiss Society of Astronomy and Astrophysics, ed. L. Martinet (to be published).

Geller, M. J., and Beers, T. C. 1982. *P.A.S.P.* **94**, 421.

Gioia, I. M., Geller, M. J., Huchra, J. P., Maccacaro, T., Steiner, J. E., and Stocke, J. 1982. *Astrophys. J. Lett.* **255**, L17.

Grossman, S. A., and Narayan, R. 1989. *Astrophys. J.* xx, xx.

Gunn, J. E. 1977. *Astrophys. J.* **218**, 592.

Harris, W. E. 1987. *Astrophys. J. Lett.* **315**, L29.

Hoessel, J. G., and Schneider, D. P. 1985. *Astron. J.* **90**, 1648.

Ishizawa, T., Matsumoto, R., Tajima, T., Kageyama, H., and Sakai, H. 1983. *P.A.S.Japan* **35**, 61.

Jones, C., and Forman, W. 1984. *Astrophys. J.* **276**, 38.

Kent, S. M., and Sargent, W. L. W. 1983. *Astron. J.* **88**, 697.

Lauer, T. R. 1985. *Astrophys. J.* **292**, 104.

Lauer, T. R. 1988. *Astrophys. J.* **325**, 49.

Lauer, T. R. 1989. Preprint.

Leir, A. A., and van den Bergh, S. 1977. *Ap. J. Supp.* **34**, 381.

Lecar, M. 1975. In *Dynamics of Stellar Systems*, IAU Symposium No. 69, ed. A. Hayli (Dordrecht: Reidel), 161.

Lin, D. N. C., and Tremaine, S. 1983. *Astrophys. J.* **264**, 364.

Malumuth, E. M., and Richstone, D. O. 1984. *Astrophys. J.* **276**, 413.

Matthews, T. A., Morgan, W. W., and Schmidt, M. 1964. *Astrophys. J.* **140**, 35.

Merritt, D. 1983. *Astrophys. J.* **264**, 24.

Merritt, D. 1984a. *Astrophys. J.* **276**, 26.

Merritt, D. 1984b. *Astrophys. J. Lett.* **280**, L5.

Merritt, D. 1985. *Astrophys. J.* **289**, 18.

Morgan, W. W., and Lesh, J. R. 1964. *Astrophys. J.* **142**, 1364.

Ostriker, J. P., and Tremaine, S. 1975. *Astrophys. J. Lett.* **202**, L113.

Peebles, P. J. E. 1980. *The Large-Scale Structure of the Universe* (Princeton: Princeton University Press).

Quintana, H., and Lawrie, D. G. 1982. *Astron. J.* **87**, 1.

Richstone, D. O. 1975. *Astrophys. J.* **200**, 535.

Richstone, D. O., and Malumuth, E. M. 1983. *Astrophys. J.* **268**, 30.

Rood, H. J. 1988. *P.A.S.P.* **100**, 1242.

Rood, H. J., and Leir, A. A. 1979. *Astrophys. J. Lett.* **231**, L13.

Sarazin, C. L. 1986. *Rev. Mod. Phys.* **58**, 1.

Schneider, D. P., Gunn, J. E., and Hoessel, J. G. 1983. *Astrophys. J.* **268**, 476.

Schombert, J. M. 1988. *Astrophys. J.* **328**, 475.

Smith, R. M., Efstathiou, G., Ellis, R. S., Frenk, C. S., and Valentijn, E. A. 1985. *Mon. Not. Roy. Astron. Soc.* **216**, 71P.

Strom, S. E., and Strom, K. M. 1979. *Astron. J.* **84**, 1091.

Tonry, J. L. 1984. *Astrophys. J.* **279**, 13.

Tonry, J. L. 1985. *Astron. J.* **90**, 2431.

Tonry, J. L. 1987. In *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, ed. T. de Zeeuw (Dordrecht: Reidel), 89.

Valentijn, E. A., and Casertano, S. 1988. *Astron. Astrophys.* **206**, 27.

White, R. A. 1978. *Astrophys. J.* **226**, 591.

White, S. D. M. 1976a. *Mon. Not. Roy. Astron. Soc.* **174**, 19.

White, S. D. M. 1976b. *Mon. Not. Roy. Astron. Soc.* **177**, 717.

White, S. D. M. 1978. *Mon. Not. Roy. Astron. Soc.* **184**, 185.

White, S. D. M., and Rees, M. J. 1978. *Mon. Not. Roy. Astron. Soc.* **183**, 341.

Yahil, A. 1974. *Astrophys. J.* **191**, 623.

Zaritsky, D., and White, S. D. M. 1988. *Mon. Not. Roy. Astron. Soc.* **235**, 289.