GALAXY MERGERS

Scott Tremaine

Institute for Advanced Study,
Princeton

1. INTRODUCTION

It is difficult to improve on Holmberg's (1940) introduction to this subject:

> The average space separation of extragalactic objects is rather small, compared with the dimensions of single objects. In a stationary universe we must expect a large number of encounters. Every close encounter between two objects will create large tidal disturbances, and the resulting loss of kinetic energy may be sufficient to effect a capture, i.e., to change the hyperbolic orbits of the objects into elliptical ones. Immediately after the capture, the elliptical orbits may be assumed to have rather large eccentricities. Every subsequent passage of a component through the pericenter of the relative orbit will, however, create new tidal effects and thus tend to decrease the eccentricity. The general result will be a gradual contraction of the relative orbit, which may continue until the two components form practically one object.

Holmberg (1941) also obtained quantitative estimates of the tidal energy loss in a close passage of two disk galaxies. He used an analog computer based on 74 movable lamps representing stars. The gravitational force on each star was determined from the net flux, which Holmberg measured by temporarily replacing the lamp with a photocell.

Despite Holmberg's pioneering work there was only a trickle of papers on mergers in the next thirty years. The main reason for the lack of interest was that close encounters and mergers seemed to be rare events. Suppose that any encounter with impact parameter less than R leads to a merger. The probability that a galaxy undergoes merger within a time T is

$$p = \pi R^2 \langle v_{rel} \rangle NT, \qquad (1)$$

where N is the number density of galaxies and $\langle v_{r e \ell} \rangle$ is the mean relative velocity. Let the Hubble constant be $H = 100h$ km s^{-1} Mpc^{-1}; then

$$p = 2\times10^{-4}\left(\frac{Nh^{-3}}{0.05\ \mathrm{Mpc}^{-3}}\right)\left(\frac{Rh}{20\ \mathrm{kpc}}\right)^{2}\left(\frac{\langle v_{re\ell}\rangle}{300\ \mathrm{km\ s}^{-1}}\right)(HT)\ . \qquad (2)$$

We see that if the capture radius R is identified with the visible size of galaxies, then the merger rate at present is very small. The probability p is independent of the Hubble constant since observations yield Nh^{-3} and Rh, and $HT \sim O(1)$. Thus, the conclusion that the merger probability is small was already reached by Holmberg (1940), even using a Hubble constant which was badly in error.

In the last decade three separate lines of argument have suggested that mergers are a common and important event in galactic evolution. In 1974 Ostriker, Peebles & Yahil argued that "there are reasons, increasing in number and quality, to believe that the masses of ordinary galaxies may have been underestimated by a factor of 10 or more." The evidence has continued to mount since then (see Faber & Gallagher 1979 for a review), and it now seems likely that most giant galaxies have massive halos which may extend 100 kpc or more in radius. In this case the capture radius R and merger probability p in (2) must be increased by large and uncertain amounts; moreover, the drag exerted by this halo on bound companion galaxies (see section 3) will lead to the orbital decay and merger of many bound systems.

Second, Toomre & Toomre (1972) and Toomre (1977) have argued that if galaxies are formed without large peculiar velocities, then it is natural to suppose that in many cases a galaxy and its nearest neighbor will form a bound pair. In the absence of tidal torquing, the galaxies will fall together and collide; if the collision is close enough to head-on then merging will occur. One argument that this scenario is common is that our Galaxy appears to form a bound pair with M31. Also, Aarseth & Fall (1980) found many mergers of loosely bound binaries in their cosmological N-body simulations.

The third and most striking argument is that the observed merger rate appears greatly to exceed the value given by equation (2). Toomre (1977) identifies ten pairs of interacting galaxies "in fairly advanced throes of merger" from $\sim$ 4000 NGC galaxies. He argues that the tails by which he identifies his candidates cannot last much more than 5×10^{8} yr; assuming a uniform rate throughout the past 10^{10} yr yields a merger probability $p = 0.05$. A reasonable extrapolation to higher rates in the past would yield an even higher value of p.

In this review I will discuss only mergers occurring after the process of galaxy formation is complete. Mergers between proto-galaxies may also be an important process (e.g. White & Rees 1978) but the details are less certain and will not be discussed here.

2. MERGERS OF GALAXIES OF COMPARABLE MASS

Consider a perturbing mass m_p which passes a galaxy at impact parameter p with a large velocity $\vec{v} = v\hat{z}$. The center of the galaxy is at the origin and the perturber's orbit is in the xz plane. For $|\vec{r}| << p$ the perturbing potential is

$$U(\vec{r}) = \frac{Gm_p}{r_p^3}\left[\frac{1}{2}r^2 - \frac{3}{2}\frac{(\vec{r}\cdot\vec{r}_p)^2}{r_p^2}\right]. \qquad (3)$$

Thus, in the impulse approximation, a star at $\vec{r}$ receives a velocity increment

$$\Delta\vec{v} = -\int \nabla U dt = \frac{2Gm_p}{vp^2}(x,-y,0). \qquad (4)$$

Assuming $\Delta\vec{v}$ is uncorrelated with $\vec{v}$ the change in energy per unit mass is

$$\Delta E = \frac{2G^2m_p^2}{v^2p^4}(x^2+y^2). \qquad (5)$$

Averaging over a spherical galaxy of mass m_g gives

$$\Delta E = \frac{4G^2m_p^2m_g}{3v^2p^4}\langle r^2\rangle . \qquad (6)$$

Thus, in the encounter of two identical galaxies, we get

$$\Delta E = \frac{8G^2m_g^3\langle r^2\rangle}{3v^2p^4} . \qquad (7)$$

The orbital energy is $E = \frac{1}{4} m_g v^2$ so merging occurs if $\Delta E > E$ or

$$pv < \left[\frac{32}{3} G^2m_g^2 \langle r^2\rangle\right]^{1/4}. \qquad (8)$$

The derivation above follows closely the argument of Spitzer (1958), who was concerned with the disruption of open clusters by interstellar clouds. Spitzer also recognized that the derivation failed when $v/p \lesssim \Omega$, where Ω is a typical angular velocity of a star in the galaxy or open cluster. In this limit the approximation that the stars are stationary during the encounter fails; for $v/p << \Omega$ the stars orbit many times during the encounter and, by adiabatic invariance, their energy change is small.

We can also treat another limit, the head-on collision at large velocity. Consider a single star passing a spherical galaxy with mass distribution M(r). If the star has impact parameter p, the perpendicular impulse which it receives is

$$\Delta v_\perp = \frac{2Gp}{v} \int_p^\infty \frac{drM(r)}{r^2\sqrt{r^2-p^2}} \quad . \tag{9}$$

Now if we let $\Sigma(r)$ be the surface density of the galaxy, and if $\mathcal{M}(r) = 2\pi\int_0^r \Sigma r dr$ is the projected mass inside a cylinder of radius r, the formula above can be written as

$$\Delta v_\perp = \frac{2G\mathcal{M}(p)}{pv} \quad . \tag{10}$$

If the star is regarded as a member of a second identical galaxy colliding head-on, then the total change in energy in both galaxies during the collision is

$$\begin{aligned} \Delta E &= 2\cdot 2\pi \int_0^\infty \tfrac{1}{2}\Delta v_\perp^{\,2}\, \Sigma r dr \\ &= \frac{8}{3}\frac{G^2}{v^2}\int_0^\infty \left[\frac{\mathcal{M}(p)}{p}\right]^3 dp. \end{aligned} \tag{11}$$

Merger occurs if $\Delta E > E$ or

$$v^4 < \frac{32G^2}{3M} \int_0^\infty \left[\frac{\mathcal{M}(p)}{p}\right]^3 dp. \tag{12}$$

For example, a modified Hubble law $\Sigma = \Sigma_0(1+r^2/a^2)^{-1}$ (Rood, Page, Kintner & King 1972) has $\mathcal{M}(r) = \pi\Sigma_0 a^2 \ln(1+r^2/a^2)$ and $\Delta E = 298.2\, G^2\Sigma_0^3 a^4/v^2$.

The head-on formula (12) complements the tidal formula (8), and with a smooth interpolation between them they should yield fairly accurate results for the energy loss in high-velocity collisions. For low velocity collisions v and p may be replaced by the velocity and distance at closest approach (where closest approach is computed from the galaxy orbits on the assumption that they are rigid, possibly interpenetrating spheres). The usefulness and accuracy of the impulsive approximation have been stressed primarily by Alladin and his coworkers (Alladin 1965, Sastry & Alladin 1970, Sastry 1972, Alladin, Potdar & Sastry 1975); Toomre (1977) and Dekel, Lecar & Shaham (1980) also point out that the impulsive formulae are surprisingly accurate for low-velocity collisions.

The simple capture criteria (8) and (12) can be checked by numerical simulations. However, even with our crude arguments the main uncertainty is in the observations because the galactic mass distribution is so poorly known (e.g. Faber & Gallagher 1979). Thus, the main purpose of numerical calculations is to determine the properties of the merger remnant (rotation speed, ellipticity, velocity dispersion, etc.).

There are two main approaches to numerical simulations of collisions and mergers of galaxies of comparable mass:

(1) Full N-body calculations have been carried out by Holmberg (1941), with 37 bodies per galaxy, by White (1978, 1979a) with 250 bodies per galaxy, by Roos & Norman (1979) with $\sim$ 30 bodies per galaxy, and by Dekel et al. (1980) with a 250 body galaxy perturbed by a single body of comparable total mass. White concentrates on mergers from bound or parabolic orbits while Dekel et al. examine hyperbolic collisions.

The disadvantages of N-body calculations are that statistical fluctuations are large because of the small number of bodies and that two-body relaxation effects cause evolution of the galaxies even when they are isolated (White 1978). For head-on collisions N rings can be used instead of N point masses (Toomre 1977).

(2) An alternative approach is to expand the potential in a complete set of harmonic functions and keep only the lowest orders of the expansion. The advantage of this method is that if K orders are kept and N particles are followed, only $\sim$ NK operations are needed per step, whereas an N-body code needs $\sim N^2$. Thus more particles can be used and statistical errors are reduced; however, all features on small scales are washed out. van Albada & van Gorkom (1977) studied the head-on collisions of two galaxies with $N = 1000$ particles per galaxy; because the system was axially symmetric they expanded the potential using Legendre polynomials up to index $\ell = 4$. Villumsen (1980) used $N = 600$ particles per galaxy and expanded the potential in tesseral harmonics up to $\ell = 4$. Miller & Smith (1980) use $N = 50{,}000$ particles per galaxy, work within a cube of side L, and keep the lowest $(64)^3$ Fourier coefficients of the potential.

A useful trick introduced by van Albada & van Gorkom is to use two separate coordinate systems in the potential calculations, one centered on each galaxy. This reduces the higher-order potential components during the initial approach.

Some of the results from these calculations include:

(1) The cross-section for mergers between rotating galaxies is strongly enhanced if the spin and orbital angular momenta are

aligned, and sharply reduced if the spins are opposite to the orbital angular momentum (White 1979a). This is not surprising since the velocity of the perturber relative to a prograde star is small so the perturbation acts for a longer time.

(2) The merger remnant usually has both a higher central density and a more extended outer envelope than its progenitors (e.g. Dekel et al. (1980) find that the radius containing half the mass shrinks but the radius containing 90% of the mass grows).

(3) Escaping stars carry away relatively little mass or energy, although they may carry away significant amounts of angular momentum in off-center collisions.

(4) Head-on collisions lead to prolate galaxies elongated along the line of the initial trajectory; off-center collisions lead to oblate galaxies flattened along the initial orbital plane.

All of the work done so far has been on mergers of spheroidal systems. These calculations are mainly useful for determining the properties of merger remnants at the centers of rich clusters (see section 4). It is at least as interesting to try to understand the merging of two disk systems, since spiral-spiral mergers are likely to be much more common than elliptical-elliptical mergers in the field. However, there has been relatively little progress in this area because, among other reasons, it is difficult to construct isolated stable disk systems, and the number of parameters to be studied is so large.

3. MERGERS OF SATELLITE GALAXIES

3.1 Dynamical friction

In the impulsive calculations of the previous section the energy change due to a perturber of mass m_p is $\propto m_p^2$. Thus the frictional effects which lead to mergers are due to second-order perturbations. Since the orbital kinetic energy is $\propto m_p$, a sufficiently low mass perturber will not be captured from an unbound orbit. However, a satellite galaxy of mass m_s in a bound orbit around a galaxy of mass $m_g >> m_s$ may still merge because its orbital energy is lost gradually over many orbits.

This process may be regarded as a manifestation of dynamical friction (Chandrasekhar 1960). A galaxy of mass m_s moving at velocity $\vec{v}$ through an infinite homogeneous background of stars with density ρ and a Gaussian velocity distribution with one-dimensional dispersion σ suffers a drag

$$\frac{dv}{dt} = - 4\pi G^2 m_s \rho v^{-2} [\phi(x) - x\phi'(x)] \quad \ell n \, \Lambda \, , \tag{13}$$

where $x = 2^{-\frac{1}{2}}v/\sigma$, ϕ is the error function and $\Lambda = p_{max}/p_{min}$, where p_{max} and p_{min} are the maximum and minimum impact parameters considered. Usually $p_{min} = \max(r_s, Gm_s/v^2)$, where r_s is the size of the satellite, and p_{max} is the scale size of the background. Notice that dynamical friction is a second-order effect in m_s since the force on m_s is $\propto m_s^2$.

One example of the use of equation (13) is Tremaine's (1976) estimate of the decay rate of the orbit of the Magellanic Clouds in the halo of our Galaxy. Alar Toomre has pointed out that Cox (1972) also treated the decay of a satellite orbit in a spherical galaxy using equation (13).

The frequency of mergers of satellite galaxies is relatively small unless the central galaxies have extended massive halos. If the halo mass distribution is an isothermal sphere with a very small core radius, then $\rho(r) \cong \sigma^2/(2\pi G r^2)$, and a satellite in a circular orbit has $v = \sqrt{2}\sigma$. From equation (13) the orbital radius evolves as

$$r^2(t) = r^2(0) - 0.605 \frac{Gm_s t}{\sigma} \ln \Lambda, \qquad (14)$$

or

$$r(t) = \left[r^2(0) - (52\mathrm{kpc})^2 (t/10^{10}\mathrm{yr})(m_s/10^{10}M_\odot)(100\mathrm{km\ s}^{-1}/\sigma)\ln \Lambda\right]^{\frac{1}{2}}. \qquad (15)$$

Now suppose that initially the number density of satellites is $n(r)$. The flux through a given radius is $\propto r^2 n(r) dr/dt \propto rn(r)$. Thus if $n(r) \propto r^{-\gamma}$ we expect depletion of bright (i.e. massive) satellite galaxies at small radii for $\gamma > 1$ and an overabundance of bright close satellites for $\gamma < 1$. In fact $\gamma \simeq 1.8$ (Peebles 1980) so there should be depletion; however, the amount of depletion is uncertain and small because tidal stripping may reduce m_s and thus $|dr/dt|$ as the galaxy spirals in. Ostriker & Turner (1979) point out that depletion may also be masked by brightening of incoming galaxies due to tidal shocks which induce star formation. The observational evidence for depletion has been discussed by Ostriker & Turner (1979), White & Valdes (1980), and White (1980). The interpretation of the observations is uncertain but there is no strong evidence for or against the rapid decay rate and short lifetimes of nearby satellites predicted by the dynamical friction formula if massive halos are present.

How much has a typical galaxy eaten? Peebles (1980) writes the number density of galaxies of mass m_s at separation r from a given galaxy as $n(r,m_s)dm_s = n_0(m_s)dm_s\ (r/r_0)^{-\gamma}$ where $n_0(m_s)dm_s$ is the field density, $\gamma \cong 1.8$, and $r_0 \cong 3h^{-1}$Mpc (Davis, Geller & Huchra 1978). This expression is valid for $r \ll r_0$. We obtain the field density from Schechter's (1976) luminosity function, assuming constant M/L: $n_0(m_s)dm_s = n^*(m_s/m^*)^{-\alpha} \exp(-m_s/m^*)dm_s/m^*$, where $n^* = 0.04\ h^3\ \mathrm{Mpc}^{-3}$, $m^* = 6.8\times10^9\ h^{-2}(M/L)M_\odot$, $\alpha = 1.25$. Using equation (14) to determine r(0) for r(t) = 0, we find that the accreted mass is

$$m_{acc}/m^* = \int_0^\infty n_0(m_s)(m_s/m^*)dm_s \int_0^{r(0)} 4\pi(r/r_0)^{-\gamma}r^2dr$$

$$= 0.06\ [\ (H_0t)\ \ell n\ \Lambda \quad (M/hL)(100\ \mathrm{km\ s^{-1}}/\sigma)]^{0.6}. \tag{16}$$

For the luminous central parts of galaxies $M/hL \simeq 15$ (Faber & Gallagher 1979), and if we adopt $\ell n\ \Lambda \simeq 2$, $H_0t \simeq 1$, then $m_{acc}/m^* \simeq 0.5\ (100\ \mathrm{km\ s^{-1}}/\sigma)^{0.6}$ independent of h. Within the large uncertainties a typical giant galaxy has eaten somewhat less than its own mass.

3.2 The validity of the dynamical friction formula

The dynamical friction formula (13) is appealing because of its simplicity. However, it is not rigorous; in particular, the formula is difficult to generalize from an infinite homogeneous background to a spherical system. To see this, consider (for example) a star in a galaxy which is initially on a circular orbit in the plane of the satellite orbit. Suppose the satellite orbit is also circular with radius r_s. The potential of the central galaxy may be written U(r), and the angular speed of an object in a circular orbit of radius r is Ω, where $\Omega^2 = r^{-1}dU/dr$. The potential from the satellite may be written as a series of terms of the form $\phi_m(r) \cos m(\theta-\Omega_s t)$, where $\Omega_s = \Omega(r_s)$, m is an integer, and θ is the azimuthal angle in the orbital plane. Then the equations of motion of the star due to a single term in the perturbing potential are

$$\frac{d^2r}{dt^2} - \frac{J^2}{r^3} = -\frac{dU}{dr} - \frac{d\phi_m}{dr}\cos m(\theta-\Omega_s t),\ \frac{dJ}{dt} = +\ m\phi_m \sin m(\theta-\Omega_s t) \tag{17}$$

where $J = r^2 d\phi/dt$ is the angular momentum. If the perturbation is turned on at t = 0 then to first order in ϕ_m the solution is

$$J_1 = -\left(\frac{\phi_m}{\Omega-\Omega_s}\right)_{r_0} \{\cos[m(\Omega_0-\Omega_s)t+m\theta_0] - \cos m\theta_0\}\ ,$$

$$r_1 = -\frac{\left[d\phi_m/dr + 2\Omega\phi_m/r(\Omega-\Omega_s)\right]_{r_0}}{\kappa_0^2 - m^2(\Omega_0-\Omega_s)^2} \times$$

$$\{\cos\left[m(\Omega_0-\Omega_s)t + m\theta_0\right] - \cos\kappa_0 t\cos m\theta_0 + \frac{m(\Omega_0-\Omega_s)}{\kappa_0}\sin\kappa_0 t\sin m\theta_0\}$$

$$+ \frac{2\Omega_0\phi_m(r_0)}{r_0\kappa_0^2(\Omega_0-\Omega_s)}\cos m\theta_0(1-\cos\kappa_0 t) , \qquad (18)$$

where the star is at (r_0, θ_0) at $t = 0$, $\Omega_0 = \Omega(r_0)$ and the epicyclic frequency κ_0 is given by $\kappa_0^2 = (d^2U/dr^2 + 3r^{-1}dU/dr)_{r_0}$. There is no secular torque on the star to first order. The calculation can also be carried to second order and there is still no secular torque. There is also no secular increase in the star's energy E since $dE/dt = \Omega_s dJ/dt$ by Jacobi's integral, and consequently no drag on the satellite analogous to dynamical friction, which is also second order in the perturbing potential. A similar conclusion was reached by Kalnajs (1972) in a more artificial but exactly soluble system.

There are two aspects to the resolution of this problem. First, consider only relatively close encounters (say, by setting p_{max} in the Chandrasekhar formula equal to $2p_{min}$). For these encounters the approximation of an infinite homogeneous background is valid since the impact parameter is much less than the scale size and a given star is unlikely to have more than one close encounter. However, this restriction decreases Chandrasekhar's drag force only by a factor $\sim \ln \Lambda$. Thus, a conservative estimate of the drag force is given by considering only close encounters, i.e. by formula (13) with $\ln \Lambda \sim 1$. The difference of a factor of $\ln \Lambda$ is not generally large enough to have observable consequences.

Second, and more important, the first order perturbations J_1 and r_1 in equation (18) diverge at resonances where $\Omega = \Omega_s$ (corotation resonance) and $\kappa_0^2 = m^2(\Omega_0 - \Omega_s)^2$ (Lindblad resonance). Thus, perturbations near these resonances are large. Moreover, the formula for the second order torque dJ_2/dt contains periodic terms of long period near these resonances. Let us restrict ourselves for the moment to the vicinity of the Lindblad resonance; the corotation resonance can be handled by similar techniques. The second-order torque near the Lindblad resonance at $\kappa = m(\Omega_0 - \Omega_s)$ due to divergent long-period terms is

$$\frac{dJ_2}{dt} = -\frac{m}{4\kappa_L r_L^2}\left(r\frac{d\phi_m}{dr} + \frac{2m\Omega}{\kappa}\phi_m\right)^2_{r_L} \frac{\sin\Delta t}{\Delta} \qquad (19)$$

where $\Delta = \kappa - m(\Omega_0-\Omega_s)$. The resonance radius r_L is defined by

$\Delta(r_L) = 0$ and we have dropped the distinction between r_L and r_0 except in Δ. The torque on a given star (fixed Δ) grows like t until it drops out of resonance at $|\Delta t| \sim \pi$. As time goes on the number of stars in resonance decreases like t^{-1} but each one feels a torque $\propto t$. Thus the total torque is independent of time. The existence of secular torques at resonances was recognized by Lynden-Bell & Kalnajs (1972), who also derived a more general form of (19) for resonances in an arbitrary flat axisymmetric system. Eventually equation (19) fails because the perturbations on the stars which are still in resonance become non-linear. However, evolution of the satellite orbit will generally bring fresh stars into resonance so that a secular torque continues to be present.

In a real galaxy, with eccentric satellite and star orbits, the resonance structure is much more complicated than a single Lindblad resonance. The purpose of the simple example presented here is to show that near-resonant stars in a galaxy <u>can</u> exert secular torques on a satellite. These torques are the analogs, for spherical or axisymmetric systems, of the drag force described by Chandrasekhar's dynamical friction formula for infinite homogeneous systems. (To see this consider a resonance with azimuthal wavenumber m of order unity. Then $\phi_m \sim Gm_s/r$, where m_s is the satellite mass, and r is the orbital radius. We may set $\kappa \sim \Omega$, and the number of field stars in an interval $d\Delta$ is $\sim \rho r^2 d\Delta/\Omega$ where ρ is the mean galaxy density. Then integrating (19) over Δ we have $dJ_2/dt \sim G^2 m_s^2 \rho/\Omega^2 r$. The classical formula (13) yields the same result, $dJ_2/dt \sim m_s r dv/dt \sim G^2 m_s^2 \rho r v^{-2} \sim G^2 m_s^2 \rho/\Omega^2 r$, using $v \sim \Omega r$.) In principle it is possible to compute exactly the frictional force on a satellite in a specified orbit in a specified axisymmetric galaxy.

There are at least two situations where this kind of calculation will yield important new results. First, we can determine how fast frictional forces fall off beyond the outer edge of a galaxy; if the galactic halo ends at 50 kpc, what is the fate of a satellite in orbit at 100 kpc? Second, we can determine the orbital evolution of close companions of spiral galaxies which are influenced by the disk component of the central galaxy. For example, the disk may slow or halt merging of companions in prograde orbits by adding angular momentum to the orbit as fast as the halo removes it.

3.3 Numerical techniques

Numerical calculations of the merging of bound companions can be done using the methods described in the previous section. These techniques are best suited for systems of two galaxies of comparable mass although White's (1978) N-body experiments include a test of Chandrasekhar's dynamical friction formula for a satellite with 0.1 times the mass of the main galaxy; the formula works quite well.

If the mass of the satellite is much less than the mass of the central galaxy, then many dynamical times are needed for merger to occur. In this case N-body experiments are both expensive and subject to numerical error and two-body relaxation effects. An alternative approach has been proposed independently by Borne (1980) and by Lin & Tremaine (1980). They investigate a system containing three types of objects: (1) a central galaxy described by a potential U(r); Lin & Tremaine use a point mass potential $U(r) = -GM/r$ while Borne uses a more realistic potential, (2) N stars of mass m which orbit the central galaxy, (3) a satellite galaxy of mass m_s, also orbiting the central galaxy. The stars interact gravitationally with the central galaxy and the satellite; the difference from a standard N-body program is that the stars do not interact with each other. Turning off the star-star attraction in this way eliminates two-body relaxation, increases numerical accuracy and greatly increases computational speed ($\sim$ N calculations per step instead of N^2). The only sacrifice is that the self-gravity of the outer parts of the central galaxy has been neglected.

I will describe two sample results using this technique. To investigate frictional effects on a satellite orbiting beyond the outer edge of a galaxy, Lin and I constructed a galaxy with central mass M = 1 (in units with G = 1) surrounded by a halo of N = 450 stars with total mass Nm = 1. The stars were initially distributed with a phase space density f depending only on energy $\varepsilon = M/r - \frac{1}{2}v^2$. We chose $f(\varepsilon) \propto \varepsilon^{2.5}$ for $0.5 < \varepsilon < 2.5$ and $f(\varepsilon) = 0$ for $\varepsilon < 0.5$ or $\varepsilon > 2.5$. Thus the maximum radius which a star could reach in its unperturbed orbit was r = 2; however, over 95% of the stars have r < 1 and we will call $r = r_0 = 1$ the "edge" of the halo. The satellite galaxy mass was $m_s = 0.1$; it was placed in a circular orbit of radius $r_s = 2.28$, over a factor of two larger than the halo edge at r_0. Its initial orbital period was $2^{\frac{1}{2}}\pi \cdot (2.28)^{3/2} = 15.30$ units. (Recall that m_s is attracted by an effective mass M+Nm = 2, since it interacts with both the stars and the central mass.) The evolution of the angular momentum J_s of the satellite is shown in Figure 1. The angular momentum changes slowly at first but with increasing speed as the orbit decays. An arrow marks the point corresponding to a circular orbit at the edge of the halo $r_0 = 1$; notice that the orbital decay is very fast beyond this point as the satellite is within the halo itself. These results show that strong frictional effects are present even in a satellite orbiting well outside the radius of most of the stars.

Our second experiment began with the same initial conditions, but the satellite galaxy was frozen into its initial circular orbit at $r_s = 2.28$ for 120 time units ($\sim$ 8 orbital times) before it was permitted to decay. An initial freezing period may correspond more closely to the way the satellite galaxy was actually formed. In this case the decay time for the satellite to reach the edge of the halo was 138 units after release, whereas in the first experiment the decay time was only 51 units. The late stages of the decay are very similar in both cases, suggesting that the system loses all memory of the initial 'freezing' once the decay has begun.

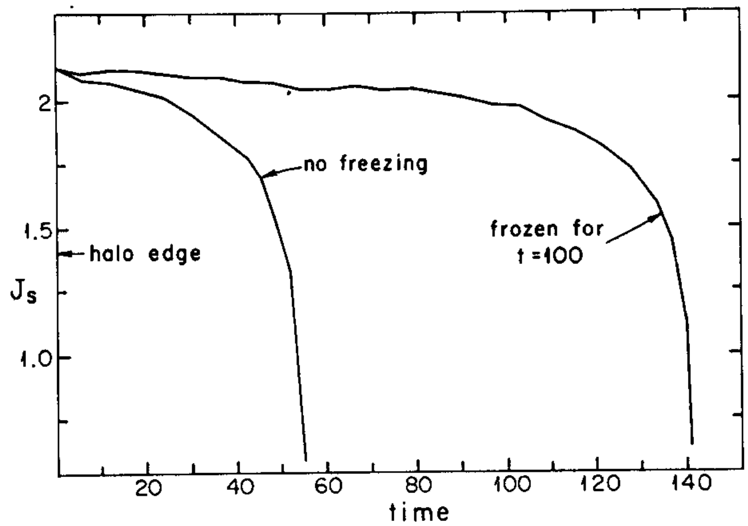

Fig. 1. Decay of the orbital angular momentum J_s of a satellite of mass $m_s = 0.1$ in orbit around a point mass $M = 1$ surrounded by a halo of $N = 450$ stars of mass $m = 1/450$. In one case the satellite was frozen into its initial circular orbit for 120 time steps before it was released.

4. MERGERS IN RICH CLUSTERS

The one-dimensional velocity dispersion in a typical rich cluster is $\sigma_{c\ell} \sim 1000$ km s^{-1}. Thus encounters between cluster galaxies usually take place at high relative velocity and mergers do not occur. However, mergers can occur by a slightly more subtle process. Dynamical friction causes the orbits of cluster galaxies to decay, just like the orbits of satellite galaxies in a halo (Lecar 1975). From equation (15), substituting $\ell n\ \Lambda \sim 1$ (a lower limit, see section 3) and $\sigma \sim \sigma_{c\ell}$, we see that the orbit of a typical giant galaxy with $m_s \sim 10^{12} M_\odot$ will decay all the way to the cluster center in 10^{10} yr if $r(0) \lesssim 150$ kpc. As galaxies accumulate in the cluster center they merge by the processes we have already discussed. Thus a single large merger remnant may be built up (Ostriker & Tremaine 1975, White 1976).

There is strong evidence that these remnants correspond to the cD galaxies identified by Morgan and his colleagues by morphological examination of photographic plates. These are defined by Matthews, Morgan & Schmidt (1964) and Morgan, Kayser & White (1975) as supergiant galaxies having an elliptical-like nucleus surrounded by an extensive envelope. Most known cD's are in rich clusters though some are found in groups (Morgan et al. 1975). It is attractive to identify the cD's in rich clusters as merger remnants because:

(1) The cD is often much more luminous than any other galaxy in the cluster (e.g. Dressler 1978). This statement is not very precise as it stands since cD's are already defined to be very luminous ('supergiant'). More accurately, we can say that the brightest members of clusters of galaxies are often more luminous than we would expect from the statistics of the luminosity function and that these exceptionally bright galaxies usually have extended halos (Oemler 1976, Tremaine & Richstone 1977, Dressler 1979). Both characteristics are necessary for a cD, and both are consistent with cD's being merger products: the observed luminosities are roughly equal to the total luminosity of the galaxies which spiral to the center in a Hubble time (White 1976) and the extended halos are observed to form in N-body simulations of mergers (cf. section 2).

(2) The cD galaxy is usually found near the cluster center (Morgan & Lesh 1965, Leir & van den Bergh 1977, R.A. White 1978), consistent with the buildup of a remnant at the bottom of the cluster's potential well.

(3) Many cD's have double or multiple nuclei (Morgan & Lesh 1965). A good example is NGC 6166, the cD in A2199 (Minkowski 1961). These may be tidally stripped remnants of galaxies which have spiralled into the cD. Hoessel (1980) finds that the fraction of first-ranked cluster galaxies with multiple nuclei separated by $< 10h^{-1}$ kpc (about 1 in 4 from a sample of 100 galaxies) is consistent with theoretical estimates of merger rates. Similarly, Rood & Leir (1979) point out that in $\sim 25\%$ of Bautz-Morgan Type I clusters (clusters containing a central cD, Bautz & Morgan 1970), the first-ranked galaxy is part of a 'dumb-bell' system containing two galaxies differing by $\lesssim 1$ magnitude. Contrary to Rood & Leir's conclusion, I believe that the observed fraction is roughly consistent with theoretical merger rates. The mean spatial separation of Rood & Leir's pairs is $\sim 20h^{-1}$kpc. From Jenner (1974) and Dressler (1979) the velocity dispersion in a typical cD envelope is ~ 500 km s^{-1}, corresponding to a circular velocity of ~ 700 km s^{-1}. The resulting orbital time is $\sim 1.8\times10^8\ h^{-1}$yr. The decay time for equal mass galaxies is about one orbital time; since the second galaxy is perhaps a factor of two fainter than the cD, a rough estimate for the decay time is $4\times10^8\ h^{-1}$yr. Since 25% of the cD's are in dumbbells we conclude that a dumbbell is formed about every (1 to 2)$\times10^9\ h^{-1}$yr. For want of better information we take the typical cluster age to be $\frac{1}{2}H_0^{-1} = 5\times10^9\ h^{-1}$yr; thus Rood & Leir's Bautz-Morgan Type I clusters have gone through 2-5 dumbbell stages on average. This is roughly the rate predicted by merger simulations

(Hausman & Ostriker 1978), and is also roughly the rate required to produce a cD galaxy one or two magnitudes brighter than the other galaxies in the cluster.

(4) Galaxies get fat when they eat. Ostriker & Hausman (1977) and Hausman & Ostriker (1978) have stressed that this effect should be detectable as a correlation between the luminosity of the first ranked galaxy and the structure parameter $\alpha = (d\ell nL/d\ell nR)_{R_0}$, where L is the luminosity inside an aperture of radius R. Hoessel (1980) has measured α for 90 first ranked galaxies at $R_0 = 10\ h^{-1}$ kpc and finds a strong correlation which follows the predictions of Hausman & Ostriker.

(5) The major axes of cD galaxies are aligned with the major axes of their clusters (Sastry 1968, Carter & Metcalfe 1980). This alignment is natural if the cD has grown by accreting cluster members.

Many further details of all these tests are given in the references. In the last five years an impressive array of evidence has been accumulated in support of the hypothesis that many first ranked cluster galaxies (and possibly all cD galaxies) are merger remnants. We must remember, however, that the center of a rich cluster is a complex environment, and many other processes may play a role in the evolution of galaxies there (e.g. Cowie & Binney 1977).

5. DO MERGED SPIRALS MAKE ELLIPTICALS?

Toomre & Toomre (1972) and Toomre (1977) have speculated that most or all elliptical galaxies may be the remnants of merged spirals. The original motivation for this suggestion was the calculation described in section 1, which showed that the observed merger rate (times a factor of three extrapolation to account for a higher rate in the past) could produce the observed fraction of ellipticals ($\sim$ 15%) in a Hubble time.

Since elliptical galaxies have virtually no interstellar gas, the merger process must remove $\sim 10^{10} M_\odot$ of gas in the collision of two normal spiral galaxies. This does not appear to present serious difficulties. The most plausible mechanism is that a burst of star formation driven by tidal or collisional shocks produces enough supernovae to drive a wind which expels the gas from the merger remnant. The peculiar colors of interacting galaxies suggest that star formation bursts do occur in galaxy collisions (Larson & Tinsley 1978). The average supernova rate in our Galaxy already supplies a large fraction of the energy input to the interstellar medium, so it is plausible that a burst of supernovae could sweep the gas out.

A second potential problem was raised by White (1979b), who pointed out that if mergers occur between unbound systems of comparable mass, then the merged remnant may have more angular momentum than observed elliptical galaxies. This is not a serious problem since most mergers occur between initially bound systems, as described in section 1 and mentioned by White. In fact, N-body simulations by Jones & Efstathiou (1979) and Aarseth & Fall (1980) yield quantitative agreement with observations. The latter authors expressed their results in terms of the dimensionless number $\lambda = J|E|^{1/2} G^{-1} M^{-5/2}$, where J, E and M are the angular momentum, energy and mass of the merged remnant. They found $\langle\lambda\rangle = 0.07 \pm 0.02$, and estimated $\langle\lambda\rangle \approx 0.08$ from observations, in good agreement. This result is encouraging but does not provide a strong test of the merger hypothesis since tidal torques also yield $\langle\lambda\rangle = 0.07 \pm 0.03$ (Efstathiou & Jones 1979). A more stringent test is to ask whether mergers produce the observed relation between flattening and V/σ (the ratio of rotation speed to velocity dispersion). This test requires accurate numerical experiments on the merging of disk systems.

However, there are several serious problems with the hypothesis that ellipticals are merged spirals:

(1) Ostriker (1980) has pointed out that the fractional abundance of ellipticals in the centers of cD clusters is high (e.g. Oemler 1974), but, as pointed out in section 4, mergers cannot occur in this environment because the encounter velocities are too high to lead to capture (cf. equation 12). Merging can only occur during the collapse of the cluster, when the velocity dispersion is low, but at this stage the density enhancement is also small so one would not expect an enhanced merger rate. This argument can and should be checked by N-body simulations since it is possible that merging may occur in sub-clumps during the collapse; existing simulations (Aarseth & Fall 1980) are inconclusive because they do not produce rich clusters (structures with velocity dispersion ~ 1000 km s^{-1}).

(2) Ostriker (1980) also points out that ellipticals satisfy both a color-luminosity relation (Visvanathan & Sandage 1977) and a metallicity-luminosity relation (Faber 1977) while spirals do not. This is difficult to explain if ellipticals are made from merged spirals.

(3) Dwarf elliptical satellite galaxies cannot form by mergers with other satellites since their relative velocities are too high. Also, it is difficult to imagine what they merged from: there are ~ 10 dwarf ellipticals within a few hundred kpc of our Galaxy and M31 and no spiral or irregular galaxies except for the Magellanic Clouds and M33, which are much more massive. Moreover, the properties of dwarf ellipticals (metallicity, number density, radius, etc.) appear to join smoothly onto those of giant ellipticals,

which suggests a common formation process.

(4) Elliptical galaxies have more globular clusters per unit luminosity than spirals, by a factor of ~ 3 (Harris & Racine 1979). However, the number of globulars per unit of luminosity in the spheroidal component is about the same for both types, as might be expected since the globulars and the spheroid are both Population II. This result is not expected in the merger picture, since in a merger the light of the disk is added to the spheroidal luminosity while the number of globulars remains the same. The problem could in principle be reduced if new globulars were produced in the merger process in the right amount, but there is no evidence that this process has occurred.

Let me make one final objection (though I will not dignify it with a number). In many ways the bulges of spiral galaxies look very similar to ellipticals (surface brightness distribution, metallicity, ellipticity, velocity dispersion, etc.). Occam's razor suggests that both should be made the same way, and spiral bulges are certainly not made by merging spirals.

Many of the arguments in this section can be sharpened (or dulled) considerably by more careful thought. At the moment there seem to be grave doubts that most ellipticals can be made from merged spirals. However, there is strong evidence, described in section 1, that several hundred NGC galaxies have undergone mergers in the past. Then where are they, and what do they look like? Are <u>some</u> ellipticals merger remnants? If so, can they be recognized in any way? And if not, as Toomre (1977) remarked, "Where else have they possibly gone?"

6. CONCLUSIONS

The work on mergers which I have described is perhaps most important as part of a fundamental change in our conception of galaxies. Over the last few years we have finally discarded the idea of galaxies as "island universes" which are born and die in splendid isolation. The replacement is a richer and more complex picture, only partly drawn, which (we hope) will lead to a marked improvement in our understanding of the observations.

I have benefited enormously from conversations with S.M. Fall, J.P. Ostriker and Alar Toomre. I am grateful to John Bahcall, James Binney, and Herb Rood for comments on a draft manuscript. This work was supported in part by NSF Grant PHY 79-19884.

REFERENCES

Aarseth, S.J. & Fall, S.M. (1980). Astrophys.J., 236, 43.
van Albada, T.S. & van Gorkom, J.H. (1977). Astron.Astrophys., 54, 121.
Alladin, S.M. (1965). Astrophys.J., 141, 768.
Alladin, S.M., Potdar, A. & Sastry, K. (1975). In Dynamics of Stellar Systems, ed. A. Hayli, p. 167. Reidel, Dordrecht.
Bautz, L. & Morgan, W.W. (1970). Astrophys.J.Lett., 162, L149.
Borne, K. (1980). Ph.D. Thesis, California Institute of Technology.
Carter, D. & Metcalfe, N. (1980). Mon.Not.R.Astr.Soc., 191, 325.
Chandrasekhar, S. (1960). Principles of Stellar Dynamics. Dover, New York.
Cowie, L. & Binney, J. (1977). Astrophys.J., 215, 723.
Cox, L.P. (1972). Unpublished.
Davis, M., Geller, M.J. & Huchra, J. (1978). Astrophys.J., 221, 1.
Dekel, A., Lecar, M. & Shaham, J. (1980). Astrophys.J., in press.
Dressler, A. (1978). Astrophys.J., 222, 23.
Dressler, A. (1979). Astrophys.J., 231, 659.
Efstathiou, G. & Jones, B.J.T. (1979). Mon.Not.R.Astr.Soc., 186, 133.
Faber, S.M. (1977). In The Evolution of Galaxies and Stellar Populations, ed. B.M. Tinsley and R.B. Larson, p. 157. Yale University Press.
Faber, S.M. & Gallagher, J.S. (1979). Ann.Rev.Astron.Astrophys., 17, 135.
Harris, W. & Racine, R. (1979). Ann.Rev.Astron.Astrophys., 17, 241.
Hausman, M.A. & Ostriker, J.P. (1978). Astrophys.J., 224, 320.
Hoessel, J.G. (1980). Astrophys.J., in press.
Holmberg, E. (1940). Astrophys.J., 92, 200.
Holmberg, E. (1941). Astrophys.J., 94, 385.
Jenner, D. (1974). Astrophys.J., 191, 55.
Jones, B.J.T. & Efstathiou, G. (1979). Mon.Not.R.Astr.Soc., 189, 27.
Kalnajs, A. (1972). In Gravitational N-Body Problem, ed. M.Lecar, p. 13. Reidel, Dordrecht.
Larson, R. & Tinsley, B. (1978). Astrophys.J., 219, 46.
Lecar, M. (1975). In Dynamics of Stellar Systems, ed. A. Hayli, p. 161. Reidel, Dordrecht.
Leir, A. & Van den Bergh, S. (1977). Astrophys.J.Suppl., 34, 381.
Lin, D.N.C. & Tremaine, S. (1980). In preparation.
Lynden-Bell, D. & Kalnajs, A.J. (1972). Mon.Not.R.Astr.Soc., 157, 1.
Matthews, T.A., Morgan, W.W. & Schmidt, M. (1964). Astrophys.J., 140, 35.
Miller, R.H. & Smith, B.F. (1980). Astrophys.J., 235, 421.
Minkowski, R. (1961). Astron.J., 60, 558.
Morgan, W.W., Kayser, S. & White, R. (1975). Astrophys.J., 199, 545.
Morgan, W.W. & Lesh, J. (1965). Astrophys.J., 142, 1364.
Oemler, A. (1974). Astrophys.J., 194, 1.
Oemler, A. (1976). Astrophys.J., 209, 693.

Ostriker, J.P. (1980). Comm.Astrophys., 8, 177.
Ostriker, J.P. & Hausman, M.A. (1977). Astrophys.J.Lett., 217, L125.
Ostriker, J.P., Peebles, P.J.E. & Yahil, A. (1974). Astrophys.J. Lett., 193, L1.
Ostriker, J.P. & Tremaine, S. (1975). Astrophys.J.Lett., 202, L113.
Ostriker, J.P. & Turner, E.L. (1979). Astrophys.J., 234, 785.
Peebles, P.J.E. (1980). The Large Scale Structure of the Universe. Princeton University Press, Princeton, N.J.
Rood, H. & Leir, A. (1979). Astrophys.J.Lett., 231, L3.
Rood, H., Page, T., Kintner, E. & King, I. (1972). Astrophys.J., 175, 627.
Roos, N. & Norman, C.A. (1979). Astron.Astrophys., 76, 75.
Sastry, G.N. (1968). Publ.Astr.Soc.Pac., 80, 252.
Sastry, K.S. (1972). Astrophys.Sp.Sci., 16, 284.
Sastry, K.S. & Alladin, S.M. (1970). Astrophys.Sp.Sci., 7, 261.
Schechter, P. (1976). Astrophys.J., 203, 297.
Spitzer, L. (1958). Astrophys.J., 127, 17.
Toomre, A. (1977). In The Evolution of Galaxies and Stellar Populations, ed. B.M. Tinsley and R.B. Larson, p. 401. Yale University Press.
Toomre, A. & Toomre, J. (1972). Astrophys.J., 178, 623.
Tremaine, S. (1976). Astrophys.J., 203, 72.
Tremaine, S. & Richstone, D. (1977). Astrophys.J., 212, 311.
Villumsen, J.V. (1980). Mon.Not.R.Astr.Soc., submitted.
Visvanathan, N. & Sandage, A. (1977). Astrophys.J., 216, 214.
White, R.A. (1978). Astrophys.J., 226, 591.
White, S.D.M. (1976). Mon.Not.R.Astr.Soc., 174, 19.
White, S.D.M. (1978). Mon.Not.R.Astr.Soc., 184, 185.
White, S.D.M. (1979a). Mon.Not.R.Astr.Soc., 189, 831.
White, S.D.M. (1979b). Astrophys.J.Lett., 229, L9.
White, S.D.M. (1980). Preprint.
White, S.D.M. & Rees, M.J. (1978). Mon.Not.R.Astr.Soc., 183, 341.
White, S.D.M. & Valdes, F. (1980). Mon.Not.R.Astr.Soc., 190, 55.