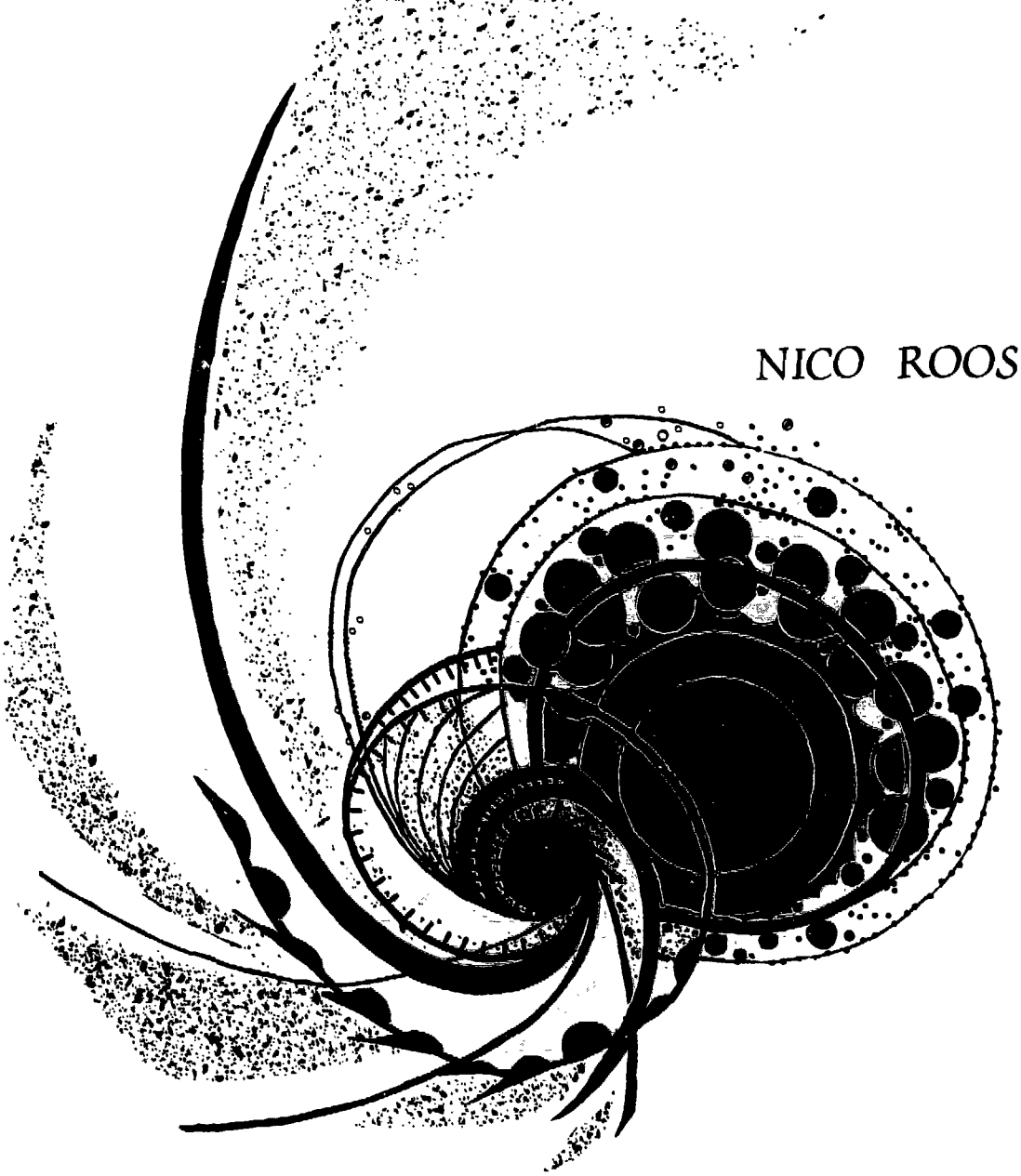
GALAXY MERGERS

Implications for the Evolution, Clustering and Central Activity of Galaxies



NICO ROOS

GALAXY MERGERS

IMPLICATIONS FOR THE EVOLUTION, CLUSTERING AND CENTRAL ACTIVITY OF GALAXIES

2 juni 1981
RU Leiden

Aan mijn ouders:
Piet Roos en Netty Groeneveld

CONTENTS

Stellingen behorende bij het proefschrift: "Galaxy Mergers" van Nico Roos

1. Het merendeel van de "fatale" tegenwerpingen die Ostriker tegen het *merger* model voor de vorming van elliptische sterrenstelsels inbrengt, kunnen gebruikt worden als argumenten ten gunste van dit model.
   Ostriker, J.P. 1980, Comments on Astrophysics 8, 177.

2. De relatie tussen de bindingsenergie en de massa van stersystemen zoals die voor elliptische stelsels bepaald is door Fish, heeft een veel algemenere geldigheid dan meestal wordt aangenomen.
   Fish, R. 1964, Astrophysical Journal 139, 284.

3. Met behulp van een geschikt waarneemprogramma kan op korte termijn bepaald worden in hoeverre activiteit in de kernen van sterrenstelsels beinvloed wordt door de omgeving.

4. Het anthropisch principe is pas dan als een fysisch principe te beschouwen als aangetoond zou kunnen worden dat de waarden van fysische constanten noodzakelijke voorwaarden zijn voor het ontstaan van levende wezens zoals de mens.

5. Het probleem van het solipsisme bij Nabokov komt onder andere tot uiting in het veelvuldig voorkomen van het solus rex thema.

6. De Oost-West onderhandelingen over ontwapening hebben de afgelopen decennia geen effect van betekenis op de (kern)wapenwedloop gehad. In een dergelijke kritieke situatie getuigt het niet van politieke realiteitszin eenzijdige ontwapeningsinitiatieven te blijven afwijzen.

7. Het "Presidential Directive 59" is een legitimatie op voorhand van een offensief gebruik van kernwapens in Europa.

Leiden, 2 juni 1981.

CHAPTER I

INTRODUCTION AND SUMMARY

This thesis contains a series of four papers (Chapter II to V) dealing with the effects of interactions among galaxies during the epoch of cluster formation. In Chapter II galaxy interactions are investigated and the results incorporated in numerical simulations of the formation of groups and clusters of galaxies. In Chapter III the role of galaxy interactions is analysed in the more general context of simulations of an expanding universe. In Chapter IV the evolution of galaxies in rich clusters is discussed. The results of the investigations presented in Chapters II to IV and their relation to other work done in the field are briefly reviewed in section I.2. In Chapter V an attempt is made to link galaxy mergers to the occurrence of activity in galactic nuclei. A short outline of this chapter is given in section I.3.

The formation of galaxies and clusters of galaxies is by no means a matter of consensus among astronomers. Even the hypothesis adopted in this thesis, that galaxies were formed before clusters, is still a point of debate. It seems therefore worthwhile to give here a brief description of the general theoretical background of the investigations presented in the chapters which follow. For a comprehensive review of the theoretical and observational developments (since about 1930) on the subject of the large-scale distribution of matter in galaxies and clusters of galaxies the reader is referred to a recently published book by Peebles (1980a).

I.1 Hierarchical Clustering

The distribution of galaxies and clusters of galaxies in the universe appears to be remarkably homogeneous on scales larger than about 100 Mpc. On smaller scales, galaxies tend to occur in associations of a few to thousands of members, in which the mean density increases with decreasing size. It is generally assumed that galaxies and clusters of galaxies are formed via gravitational instability from small density fluctuations in the early universe. The spectrum of initial density fluctuations is still unknown, but the simplest assumption, suggested

by observations is that the spectrum of density fluctuations obeys some power law $\partial\rho/\rho \propto M^{-\alpha}$, where M is the mass contained in the fluctuation and $\alpha \sim 1/3$ to 2/3 (e.g. Gott, 1979; Rees, 1979). The lack of knowledge about the origin and evolution of these fluctuations before the epoch of recombination has lead to two opposing scenarios for the formation of galaxies and clusters of galaxies. Doroshkevich et al. (1974) have proposed and explored a scenario in which only fluctuations having a mass $10^{13} - 10^{14}$ $M_\odot$ (see, however, Press and Vishniac, 1980) are able to grow before recombination. These protoclusters contract, cool and fragment into galaxies and stars. An attractive feature of such a scenario, in which galaxies form directly from gas clouds, is that a mass and radius which are characteristic for galaxies follow from equating the cooling and free fall times of gas clouds (Silk, 1977; Rees and Ostriker, 1977). In the alternative scenario it is assumed that soon after recombination gas dissipation is completed and galaxy interactions are purely gravitational. Fluctuations in the mass distribution will then lead to the formation of bound and virialised systems on progressively larger scales (Peebles and Dicke, 1968; Press and Schechter, 1974). This hierarchical clustering process accounts in a natural way for a number of basic properties of galaxies and clusters:

(i) The luminosity function of galaxies in rich clusters does not differ significantly (except perhaps at the bright end) from the luminosity function in regions of low galaxy density (Schechter, 1974)

(ii) The form of the luminosity function of galaxies is very similar to the form of the self-similar mass function arising in a hierarchical clustering process

(iii) The luminosity function of galaxies joins smoothly to the luminosity (or multiplicity) function of groups and clusters of galaxies (Gott and Turner, 1976; Bahcall, 1979)

(iv) The clustering of galaxies has no preferred scale: the observed two-point correlation function is well described by a single power law for galaxy separations ranging from $\sim$ 10 kpc to $\sim$ 10 Mpc[+]. Another interesting point noted by Peebles (1980a) is that the

[+]A value of $H_o = 50$ km sec^{-1} Mpc^{-1} for the Hubble constant is adopted throughout this thesis.

galaxy distribution around galaxies, as derived from the two point correlation function, joins smoothly to the mass distribution inside galaxies if the mean mass density of the universe equals the critical density.

(v) Several characteristic properties of galaxy clustering such as the two- and three-point correlation functions and the multiplicity function can be reproduced in cosmological N-body simulations (Miyoshi and Kihara, 1975; Aarseth et al. 1979).

The maximum scale of galaxy clustering is larger than indicated by the galaxy-galaxy correlation function alone. It has been shown by Kiang and Saslaw (1969) and by Hauser and Peebles (1973) that the distribution of Abell clusters is non-random on a maximum scale of 80 - 100 Mpc. Recent observations of the large-scale distribution of galaxies using samples for which complete redshift information is available have dramatically confirmed the existence of such large structures, often called superclusters (Gregory and Thompson, 1978; Einasto et al., 1980). Clusters and superclusters appear to be separated by large regions of space containing very few or no galaxies. According to Einasto et al. (1980) the galaxies outside rich clusters are arranged in thin sheet-like structures connecting the clusters. They prefer an interpretation of superclusters within the context of the scenario proposed by Doroshkovich et al. (1974). Large voids between clusters and superclusters do not seem inconsistent with the hierarchical clustering picture and large empty regions have indeed been found in N-body simulations (Aarseth et al. 1979), although the maximum scale covered in these simulations is smaller than about 25 Mpc. Non-spherically symmetric galaxy distributions are expected as well in these large structures for which the crossing time exceeds the Hubble time. However, systematic sheet-like arrangements of galaxies outside rich clusters would admittedly be inconsistent with the hierarchical clustering picture.

Another problem posed by superclusters has been noted by Peebles (1980a, §78. He points out that the mean excess density of clusters within about 80 Mpc of the centre of a rich cluster is larger than the excess density of galaxies. This is not expected in the hierarchical clustering picture and he concludes that "galaxies and clusters cannot both be good tracers of the large-scale mass distribution". Another interpretation is suggested by the result, found in Chapter III of this

thesis, that the probability for two neighbouring masses to form a bound pair is proportional to their combined mass. This implies that the ratio of the mean number of clusters over the mean number of galaxies around a cluster centre is given by

$$N_{cc}/N_{cg} = \frac{2M_c}{M_c + M_g} \bar{n}_c/\bar{n}_g ,$$

which is about $2\bar{n}_c/\bar{n}_g$ since the cluster mass M_c is much larger than the galaxy mass M_g. Here $\bar{n}_c$ and $\bar{n}_g$ are the mean number densities of clusters and galaxies respectively. This is consistent with the observed densities given by Peebles (1980a) in his Table 78.2 from which we find $N_{cc}/N_{cg} = 2.1 \pm 0.6\ \bar{n}_c/\bar{n}_g$.

Additional support for the hierarchical clustering picture comes from recent X-ray observations of distant clusters indicating that the gas in distant clusters is cooler than in nearby clusters (Perrenod and Henry, 1981).

In a purely dissipationless hierarchical clustering process substructure is destroyed during the collapse of bound systems and clusters would not contain galaxies. Dissipation must have played an important role in the formation of the luminous parts of galaxies. There is much evidence that luminous galaxies are embedded in massive, dark halos (cf. Peebles, 1980b). In an attractive model White and Rees (1978) propose that these halos are formed by hierarchical clustering while the visible galaxies are formed dissipatively from gas falling onto the potential wells of the halos. The luminous cores may survive in the subsequent hierarchical clustering process due to their smaller cross section. Thus, in this scenario, the epoch of cluster formation sets in when the dissipational formation of galaxies is completed. There are indications that this happened within about 10^9 yr after the Big Bang (van Albada et al. 1981). The characteristic size and mass of galaxies will now also be determined by gas dynamical processes (see above). Moreover the characteristic angular momentum of luminous disk galaxies can be explained as the result of tidal torques if they have contracted by a factor of about 10 (Efstathiou and Jones, 1979).

We may conclude that a combination of a hierarchical clustering model for the dominant mass component in the universe with dissipational formation of the visible galaxies offers a good starting point for further investigations of the formation and evolution of galaxies and clusters of galaxies.

I.2 Galaxy Interactions and the Origin of Morphological Types

The flatness of the disks of spiral galaxies suggests that they were formed directly from cooling gas contracting mainly along its rotation axis. The difference in structure and rotation velocity (Illingworth, 1977; Binney, 1974) between elliptical galaxies and spiral galaxies must mean that they were formed in different ways. Gott (1977) remarks in his review of recent theories of galaxy formation that "dissipationless collapse models can give a natural account of the dynamical properties of elliptical galaxies (and spheroidal components of spiral galaxies) relying on stellar dynamics alone". Thus, elliptical galaxies and the spheroidal components of spiral galaxies may have formed after gas dissipation was completed, that is during the epoch of cluster formation. Toomre and Toomre (1972; see also Toomre, 1977) have argued that encounters between galaxies at low relative speed are very inelastic and may lead to coalenscence of the two galaxies to form a stellar system very similar to an elliptical galaxy. Estimating the fraction of merger products from the observed fraction at tidally interacting galaxies they found a number consistent with the over-all fraction of elliptical galaxies. In rich clusters the relative velocities of the galaxies is high, and Spitzer and Baade (1951) showed, using the impulsive approximation, that collisions can hardly have affected the stellar distributions of the galaxies in these clusters. One of the objections raised against Toomre and Toomre's model was that the fraction of ellipticals in the central regions of rich clusters was higher than outside clusters.

In the first part of Chapter II the results of a systematic numerical study of the effects of collisions between galaxies (consisting of 28 to 94 particles) are presented. It is found that (identical) colliding galaxies will merge if the relative velocities are smaller than their internal velocity dispersion and the impact parameters are smaller than about twice their radius. In the second part of Chapter II collisional effects are incorporated in simulations of the collapse of some groups and clusters of (initially identical) galaxies (N = 10-60). Most mergers occur during the collapse of the clusters and the final fraction of merger products is consistent with the observed fraction of elliptical galaxies. Also, the fraction of merger products is highest in the

central region of the clusters, in agreement with observations for elliptical galaxies.

These results are confirmed in a series of papers on N-body simulations of an expanding universe in which the (initially identical) particles were allowed to merge (Jones and Efstathiou, 1979; Aarseth and Fall, 1979; Chapter III). An important result of the first two papers is that the angular momentum of merger products is consistent with the characteristic angular momentum of elliptical galaxies. In Chapter III much attention is payed to the clustering properties of the merged particles in the simulations. It is shown that the merging probability of two galaxies is roughly proportional to their combined mass, in agreement with the infall model by Gunn and Gott (1972). This probability law also leads asymptotically to a self-similar mass function which is identical to the mass function obtained by Press and Schechter (1974) for self-similar hierarchical clustering in the case of a Poissonian initial density fluctuation spectrum. It is found that the observed clustering properties of elliptical galaxies can be explained using this probability law. The clustering properties of 50 galaxies indicate that they form a class intermediate between spirals and ellipticals. In the second part of Chapter III a model is investigated in which galaxies are formed at $z \gtrsim 10$ as predominantly stellar, late-type, disk galaxies having a mass function very similar to that at the present epoch. During the epoch of cluster formation, growth of a spheroidal component of galaxies due to infall of smaller galaxies, leads to an evolution along the sequence (Sc - Sb - Sa) - SO - E. The ratio (Δ) of the infallen mass over the original mass should be comparable to the bulge to disk ratio of galaxies in this model. The model prediction that the fraction of early-type galaxies should increase with galaxy mass is found to be consistent with observations.

Results of a numerical simulation of a rich cluster of galaxies, starting with a realistic initial mass function, are presented in Chapter IV. Mass segregation, segregation of morphological types (as defined above) and tidal stripping of galaxies are discussed and compared with recent observational data.

Merging of galaxies occurs inevitably in the hierarchical clustering picture. It appears that the merger model can explain in a natural manner several observed properties of elliptical galaxies such as their

frequency, mass function, clustering properties, characteristic rotation and density profile (White, 1979). If ellipticals are the result of mergers it seems difficult to escape the conclusion that the spheroidal components of disk galaxies were formed by the same process.

Dissipation must have played some role in the mergers that took place during the transition from the epoch of galaxy formation to the epoch of cluster formation. A merger model for the formation of ellipticals and the spheroidal components of disk galaxies including dissipation has recently been proposed by Tinsley and Larson (1979) and by Norman and Silk (1981). They argue that observed properties such as the color-magnitude relation and the existence of metallicity gradients require a dissipational galaxy formation process and they propose that spheroidal systems are formed in mergers of gaseous systems. Disks would then have formed late during the epoch of cluster formation, and in rich clusters disk formation would be inhibited, resulting in a high fraction of elliptical galaxies. It is clear that the benefits of the dissipationless merger model, such as a clear distinction between disk and spheroidal components, are not necessarily conserved in such a model. Also, the observational evidence seems to favour a fast collapse ($\lesssim 10^9$ yr) of the Milky Way (van Albada et al., 1981). It may therefore be useful to investigate to what extent mergers between mainly stellar disk systems (which are themselves formed dissipationally) can account for or at least conserve those properties of ellipticals and the spheroidal components of disk galaxies, which seem the result of dissipation.

I.3 Galaxy Mergers and Active Nuclei

Accretion of mass by a massive black hole seems a very promising mechanism to explain the quantity and characteristics of the radiation emitted from the nuclei of active galaxies (Rees, 1978). Accretion rates of the order of 1 $M_\odot$ yr^{-1} are required to account for the energy output of the most luminous active galaxies at the present epoch. This mass could be provided by infall of intergalactic gas clouds (Gunn, 1979) or gaseous dwarf galaxies (Silk and Norman, 1979), by stellar mass loss processes, by star-star collisions in a dense stellar nucleus (Spitzer, 1971) or by tidal disruption of stars in the vicinity of the massive

black hole (Hills, 1975). The latter process is particularly attractive because it occurs very close to the Schwarzschild radius of a massive hole and a large fraction of the gas may be swallowed by the hole.

Merging of galaxies is expected to affect the conditions in the nuclei of galaxies (cf. Begelman et al. 1980). In Chapter V the effect of a merger on the stellar orbits in the central region of a galaxy, assumed to contain a massive black hole, is considered. It is argued that perturbation of the stellar orbits by the intruding galaxy will lead to an enhanced tidal disruption rate of stars near the black hole. Assuming stellar densities similar to those in the central part of the Milky Way (Oort, 1977), this process could yield luminosities comparable to those observed in Seyfert galaxies. The active phase may last 10^6 to 10^7 yr if the mass ratio of the merging galaxies is of the order of 10^{-2}.

This model allows an estimate of the fraction of galaxies that is expected to be active at some level of activity. Comparison with observations is complicated by our limited knowledge of the total luminosity function of active galaxies. In section V.4.2 an attempt is made to estimate this function from available radio, optical and X-ray luminosity functions. A standard radiation spectrum similar to that of radio-loud 'QSO's was used to convert the luminosities in the different wavebands to total luminosities. The resulting total luminosity functions are remarkably similar to each other, indicating that the different types of activity may be manifestations of a basic mode of underlying activity. The agreement between the model prediction and observed luminosity functions is encouraging.

The evolution of a binary black hole formed by mergers of two galaxies both containing a massive black hole is also discussed (see also Begelman et al. 1980). If a massive binary is embedded in an extended, low density core, it cannot lose its orbital energy via interactions with stars below a critical binary separation distance and a long-lived binary may be the result. In section V.2.3 it is shown that, for star densities similar to those in the central region of the Milky Way, a massive binary may reach a separation small enough for loss of orbital energy via emission of gravitational waves to become important, resulting in merging of the holes.

Chapter V is very much a preliminary study of the implications of

merging for active galaxies and the results are still tentative in several respects. On the theoretical side the perturbation of stellar orbits by an intruding galaxy should be investigated in greater detail. High resolution (spectro-) photometric observations as may be provided by the Space Telescope, could reveal the structure of galaxies on scales as small as a few parsecs.

I.4 Concluding Remarks

The dramatic effects of galaxy interactions are well illustrated by the photographs in Vorontsov-Velyaminov's (1959) "Atlas and Catalog of Interacting Galaxies" and by Arp's (1966) "Atlas of Peculiar Galaxies". Toomre and Toomre (1972) first showed that many of the features of these galaxies could be explained as the result of tidal interactions between galaxies. However, for several years their cautious suggestion that elliptical galaxies might be the endproducts of such violent encounters, remained a relatively isolated alternative among models for the origin of different galaxy types. In the last few years this situation has changed. Galaxies can no longer be regarded as isolated systems. Galaxy interactions and mergers between galaxies can play a significant and sometimes dominant role in determining the evolution, clustering and central activity of galaxies.

References

Aarseth, S.J. and Fall, S.M. 1979, Ap.J. 236, 43

Aarseth, S.J., Gott, J.R., Turner, E.L. 1979, Ap.J. 228, 664

Arp, H. 1966, *Atlas of Peculiar Galaxies* (Pasadena: California Institute of Technology); also Ap.J. Suppl. 14, No. 123

Bahcall, N.A. 1979, Ap.J. 232, 689

Begelman, M.C., Blandford, R.D. and Rees, M.J. 1980, Nature 287, 307

Binney, J. 1974, M.N.R.A.S. 168, 73

Doroshkevich, A.G., Sunyaev, R.A. and Zeldovich, Ya.B. 1974, In *"Confrontation of Cosmological Theory with Observations"*, ed. M.S. Longair (Reidel)

Efstathiou, G. and Jones, B.J.T. 1979, M.N.R.A.S. 186, 133
Einasto, J., Jõeveer, M. and Saar, E. 1980, M.N.R.A.S. 193, 353
Gott, J.R. and Turner, E.L. 1976, Ap.J. 209, 1
Gott, J.R. 1977, Ann. Rev. Astron. and Astroph. 15, 235
Gott, J.R. and Turner, E.L. 1979, Ap.J. 232, L79
Gott, J.R. 1980, In *"Physical Cosmology"*, eds. R. Balian, J. Audouze, D.N. Schramm, p. 564, (North Holland Publishing Company, Amsterdam, New York, Oxford)
Gott, J.R., Turner, E.L., Aarseth, S.J. 1980, Ap.J. 234, 13
Gregory, S.A. and Thompson, L.A. 1978, Ap.J. 222, 784
Gunn, J.E. and Gott, J.R. 1972, Ap.J. 176, 1
Gunn, J.E. 1979, In *"Active Galactic Nuclei"*, p.213, eds. C. Hazard, and S. Mitton (Cambridge University Press)
Hauser, M.G. and Peebles, P.J.E. 1973, Ap.J. 185, 757
Hills, J.G. 1975, Nature 254, 295
Illingworth, G. 1977, Ap.J. 218, L43
Jones, B.J.T. and Efstathiou, G. 1979, M.N.R.A.S. 189, 27
Kiang, T. and Saslaw, W.C. 1969, M.N.R.A.S. 143, 129
Miyoshi, K. and Kihara, T. 1975, Publ. Astron. Soc. of Japan, 27, 333
Norman, C.A. and Silk, J. 1981, Ap.J. (June 1)
Oort, J.H. 1977, Ann. Rev. of Astron. and Astroph. 15, 295
Peebles, P.J.E. and Dicke, R.H. 1968, Ap.J. 154, 891
Peebles, P.J.E. 1980a, In *"The Large-Scale Structure of the Universe"*, (Princeton University Press, Princeton, New Jersey)
Peebles, P.J.E. 1980b, In *"Physical Cosmology"*, eds. R. Balian, J. Audouze and D.M. Schramm, p. 213 (North Holland Publishing Company, Amsterdam, New York, Oxford)
Perrenod, S.C., Henry, J.P. 1981, Ap.J. Letters, submitted
Press, W.H. and Schechter, P. 1974, Ap.J. 187, 425
Press, W.H. and Vishniac, E.T. 1980, Ap.J. 236, 323
Rees, M.J. and Ostriker, J.P. 1977, M.N.R.A.S. 179, 541
Rees, M.J. 1978, Physica Scripta 17 (No.3), 193
Rees, M.J. 1979, In *"Observational Cosmology"*, eds. A. Maeder, L. Martinet, G. Tammann (Geneva Observatory, Ch-1290 Sauverny, Switzerland)
Schechter, P. 1976, Ap.J. 203, 297

Silk, J.I. 1977, Ap.J. 211, 638

Silk, J.I. and Norman, C.A. 1979, Ap.J. 234, 86

Spitzer, L. and Baade, W. 1951, Ap.J. 113, 413

Spitzer, L. 1971, In *"Galactic Nuclei"*, ed. D. O'Connell, p. 443 (North Holland Publishing Company, Amsterdam, New York, Oxford)

Tinsley, B.M. and Larson, R.B. 1979, M.N.R.A.S. 186, 503

Toomre, A and Toomre, J. 1972, Ap.J. 178, 662

Toomre, A. 1977, In *"The Evolution of Galaxies and Stellar Populations"*, p. 401, eds. B.M. Tinsley and R.B. Larson (New Haven: Yale Univ. Obs.)

Turner, E.L., Aarseth, S.J., Gott, J.R., Blanchard, N.T. and Matthieu, R. 1979, Ap.J. 228, 684

Van Albada, T.S., Dickens, R.J. and Wevers, B.M.H.R. 1981, preprint

Vorontsov-Velyaminov, B.A. 1958, *"Atlas and Catalog of Interacting Galaxies"* (Moscow: Sternberg State Astronomical Institute)

White, S.D.M. and Rees, M.J. 1978, M.N.R.A.S. 183, 341

White, S.D.M. 1979, M.N.R.A.S. 189, 831

CHAPTER II

Galaxy Collisions and Their Influence on the Dynamics and Evolution of Groups and Clusters of Galaxies

N. Roos and C. A. Norman

Sterrewacht, Huygens Laboratorium, Wassenaarseweg 78, NL-2300 RA Leiden, The Netherlands

Received July 6, 1978

Summary. We investigate the effect of collisions on the dynamics and evolution of groups and clusters of galaxies. It is assumed that galaxies are formed initially with massive spherical halos and their subsequent evolution is studied using N-body simulations. Firstly, individual galaxy-galaxy collisions are simulated for a large number of impact parameters, velocities and masses with approximately 30 point masses in each colliding galaxy. Loss of orbital kinetic energy, merging criteria, increase of internal energy and mass-loss from the galaxies are discussed. Secondly, the results of these calculations are used in calculating the energy change and mass loss of the galaxies in a cluster of 10–60 members. It is shown that, in particular during the collapse of a group or cluster of galaxies, a considerable fraction of the galaxies merge and that elliptical galaxies are good candidates for these merged systems. Observational consequences are discussed for a number of properties of groups of galaxies such as the occurrence of interacting galaxies, the relative number and spatial distribution of elliptical and disk galaxies, the formation of cD-galaxies, and the mass-to-light ratio of groups.

Key words: galaxy collisions – clusters – elliptical galaxies

1. Introduction

As early as 1951 Spitzer and Baade pointed out that collisions between galaxies must occur frequently in the central regions of rich clusters. Since then the possibility that galaxies are much larger than their visible component has revived interest in the galaxy collisions. Indications that galaxies have or have had massive halos are fairly convincing. Studies of individual galaxies (Roberts and Rots, 1973; Hartwick and Sargent, 1977; Bosma, 1978), binaries (Turner and Ostriker, 1977), instabilities in disk galaxies (Ostriker and Peebles, 1973) and of galaxy formation (Rees and Ostriker, 1977) give indirect evidence for the existence of dark massive halos surrounding the visible parts of galaxies. We can estimate an upper limit to the mass of these halos from the mass of rich clusters of galaxies derived by the virial theorem. This mass is about ten times larger than the mass present in the visible part of the galaxies. Recent studies (reviewed by Bahcall, 1977) indicate that this "missing mass" does not consist of neutral or ionized gas, but probably of some kind of clumpy material with a very high M/L-value (White, 1977). Another indication of the stellar nature of the missing mass comes from recent observations that the hot intracluster medium has approximately the solar abundance and a total mass equal to that present in the visible part of the galaxies (Malina et al., 1978). This suggests that not only the visible mass of the cluster, but the total virial mass was involved in producing this gas. Therefore we feel justified in calculating collisions of galaxies with massive halos and studying their effect on the dynamics and evolution of clusters of galaxies.

Send offprint requests to: N. Roos

2. Galaxy Collisions

2.1. Review of Previous Work

Both numerical and analytical approaches to the problem of collisions between galaxies consisting of a number of point masses have been utilized, but it was only a few years ago that the first selfconsistent N-body calculations of collisions between two spherical galaxies were reported. In earlier calculations it was assumed that the over-all structure of the galaxies is not changed during a collision thus restricting the applicability of these calculations to velocities that are much larger than the velocity dispersion of the stars in the galaxies. Analytic calculations of this kind were done by Alladin in 1965 (see also Sastry and Alladin, 1977) assuming that the internal motions of the stars within the galaxies may be neglected during a collision (the "impulsive approximation").

In the numerical work by Richstone (1977) collisions between truncated isothermal spheres as described by King (1966) were calculated. He calculated the orbits of the stars under the influence of the gravitational field of the colliding galaxies, which were assumed to be rigid. From his work, and also from the work of Biermann and Silk (1977), it is evident that collisions must be penetrating in order to give any significant changes in the internal energy. They find mass losses and energy changes less than about 10% in the velocity region for which their assumptions are valid. Structural changes of the galaxies during low impact velocity encounters make a self-consistent calculation of the stellar orbits necessary. Such calculations (Toomre, 1974, 1977; Lauberts, 1974) indicated that collisions can become very inelastic at low velocities.

An axially symmetric model for the head-on collisions of up to 2000 particle galaxies with a density distribution of an $n=3$ polytrope has been discussed by van Albada and van Gorkom (1976). For low velocities highly inelastic behaviour is found and in particular merging of the galaxies occurs at relative velocities below 3.2 σ where σ is the internal velocity dispersion of the galaxies (van Albada and van Gorkom, private communication). They find that loss of orbital kinetic energy decreases rapidly with increasing relative velocity. Their code cannot accurately describe

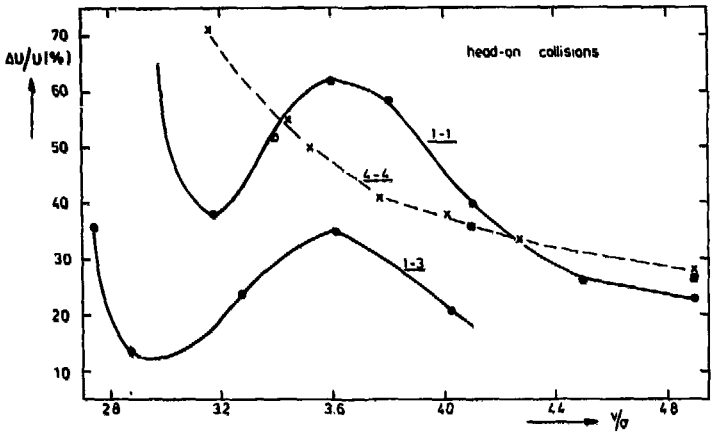


Fig. 1. *Fractional change of the internal* energy of a galaxy after head-on collisions between identical galaxies (1 – 1 and 4 – 4) and of galaxy 1 after collisions with galaxy 3 (1 – 3) as a function of the relative velocity at minimum separation in units of the internal velocity dispersion of the galaxy. 28 particles per galaxy were used in the calculation of the upper solid curve and 94 in the calculation of the dashed curve along the crosses. The squares are results obtained by van Albada and van Gorkom with identical $n=3$ polytropes ($N=1000$)

Table 1

Galaxy	M $10^{11}\ M_\odot$	$R_{1/2}$ 100 kpc	σ 300 km s^{-1}	N
1	14	0.4	0.80	28
2	10	0.3	0.75	29
3	7	0.22	0.75	29
4	15	0.4	0.8	94

non-zero impact parameter collisions because both galaxies are assumed to remain symmetric about the line connecting the two centres of gravity.

White's calculations (1978) show that collisions between two spherically symmetric galaxies in a parabolic orbit are strongly inelastic as long as the mass distributions of the galaxies overlap significantly at closest approach.

However systematic information on the effects of galaxy collisions, especially for low collision velocities, is still lacking. Clearly this information is necessary if one wants to estimate the effects of collisions on the dynamics and evolution of clusters of galaxies.

2.2. *Description of the N-body Simulation*

The numerical code is a three dimensional N-body code in which the integration of the orbits is done with the same predictor-corrector method as used by an Albada and van Gorkom (1977). As usual in this kind of calculation the Newtonian potential is modified from r^{-1} to $(r^2+\varepsilon^2)^{-1/2}$, where ε is called the softness parameter, in order to suppress large accelerations in close encounters. The time-step is adjusted during a calculation so that the changes in the velocities in one step will not exceed a chosen small value. In our calculations the total energy of the N-body system does not change by more than a few tenths of a percent during 2,000 time-steps. For this type of simulation time reversibility will not be satisfied because the deviations from the "true" orbits grow quickly but the overall statistical properties such as total energy, angular momentum, and total momentum will be conserved as discussed by Smith (1977). The calculation time was found to depend strongly on the value of ε because short range interactions and thus ε determine the time step. A PDP 11/45 was used for the calculations of the collisions.

We formed equilibrium galaxies by allowing an initial spherical distribution of about 40 equal mass particles to evolve towards dynamical equilibrium during several crossing times. The ε used was 10 kpc, about one quarter of the final radius of the smallest galaxy. The trial and error method was used to find initial conditions which formed galaxies with a range of masses from 0.5–2.0 $10^{12}\ M_\odot$ and an internal velocity dispersion of about 250 km s^{-1}. During galaxy formation those mass-points that had enough energy to escape to radii larger than 100 kpc were thrown out of the simulation. The final distribution is similar to a truncated isothermal sphere (King, 1966).

The relaxation time is not very long for a system consisting of only 30 particles. We find no changes in the mean structure of the galaxies during at least 10 crossing times which is about the time we need to calculate a collision. However even after 10 crossing times the galaxies with $N \simeq 30$ appeared to have fluctuations about the equilibrium state accounting for roughly 10% of the potential energy with a characteristic time of a crossing time. The amplitude of these apparently radial oscillations tend to decrease with increasing number of particles. In order to trace the effects of these fluctuations we performed collision calculations with an $N=94$ galaxy. The properties of the galaxies we used in the collision calculations are displayed in Table 1. Columns 3, 4, and 5 of this table gives the half-mass radius of the galaxies, the internal velocity dispersion and the number of point masses of the galaxy respectively.

These galaxies are then collided with one another for a large number of impact parameters and relative velocities with an initial separation of about 400 kpc. The calculation is continued after the collision until the galaxies have had time to reach equilibrium again. After the collision, the loss of orbital kinetic energy, the gain of internal energy, the mass loss and the change of internal angular momentum is calculated for each galaxy. We define the mass loss as the percentage of the total mass that can reach a radius >150 kpc, i.e. about half the mean intergalactic distance in a dense group of galaxies.

2.3. *Results and Interpretation*

In this section we will use the symbols U and σ for the internal energy and internal velocity dispersion of a galaxy, E, p, and r for

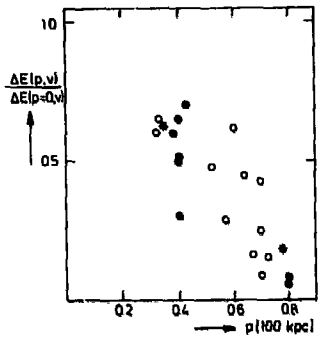


Fig. 2. The ratio of the loss of orbital energy in collisions with $p \neq 0$ over the loss of orbital energy in head-on collisions at the same velocity. Open circles denote collisions at low velocity, filled circles at high velocity. The crosses are results obtained with galaxy 4

the orbital kinetic energy, the impact parameter and relative velocity of the galaxies at minimum separation and E_0, p_0, and v_0 for these quantities measured at the initial separation of 400 kpc.

2.3.1. Change of Internal Energy of Colliding Galaxies

In a collision the internal energy of the galaxies increases due to mutual tidal interaction. The short range tidal force is most effective during the overlap of the two galaxies in a head-on collision. At maximum overlap of two equal galaxies the inward gravitational force on each particle becomes about twice as large, resulting in a radial contraction followed by an outward splash (Toomre, 1977). The radial oscillation that is excited will be damped out quickly by violent relaxation leaving the galaxy with a higher internal energy than before. The characteristic time scale of the potential variation during a collision is of the order of the collisional crossing time or overlap time. The characteristic response time of the galaxy to an extra gravitational impulse at its center is the free fall time. Thus we expect resonance-like behaviour to occur in head-on collisions at collision velocities of a few times the internal velocity dispersion of a galaxy.

The solid lines in Fig. 1 give the fractional change of the internal energy of galaxy 1 in head-on collisions with itself (1 – 1) and with galaxy 3 (1 – 3) as a function of the collision velocity in units of its internal velocity dispersion. The crosses along the dashed line are results obtained with galaxy 4 ($N=94$). The black squares are results obtained by van Albada and van Gorkom (1978). There are several things to note in this figure. First, there is a very close agreement between the dashed line and the calculations by van Albada and van Gorkom with $N=1000$ galaxies. The agreement between these results and the solid line 1–1 is best for higher velocities. At velocities $\gtrsim 3.5\,\sigma$ the dashed curve shows a fall off proportional to v^{-2} as predicted by the impulsive approximation.

The discrepancy between the solid line and the dashed one is caused by the fluctuations around the equilibrium state that galaxy 1 exhibits. Each collision calculation starts at the same initial separation and a galaxy always starts at the same initial phase of its radial oscillation. Hence, the phase at the moment of collision depends only on the initial velocity of the galaxies. When a galaxy is contracting at the moment of collision this contraction will make the induced contraction stronger and $\Delta U/U$ larger. Similarly, at different velocities a minimum will occur if the collision occurs in an expansion phase. In this way the oscillation of the solid line 1–1 around the dashed line reflects the oscillatory behaviour of galaxy 1.

The increase in the internal energy is accounted for by a loss of orbital kinetic energy of the colliding galaxies, and at velocities lower than $3.1\,\sigma$ the galaxies merge. Note that this corresponds to a relative velocity at infinity of about $1.2\,\sigma$ (Sect. 2.3.5).

The solid line 1 – 3 in Fig. 1 shows the change of internal energy that galaxy 1 undergoes as a result of a collision with galaxy 3 which is about half as massive and half as large. $\Delta U/U$ is now about half as large, as expected on the basis of the impulsive approximation. However galaxy 3 undergoes roughly the same change in internal energy as in a head-on collision with itself. The reason is that during a collision with galaxy 1, galaxy 3 feels the same amount of mass within its radius as in a head-on collision with itself. We adopt the following relation for different masses

$$(\Delta U/U)_1 = \frac{M_2}{M_1}(\Delta U/U)_2$$

or, using

$$U = -\tfrac{1}{2} M\sigma^2,$$

$$\Delta U_1 = \Delta U_2 \qquad (1)$$

in the case where $\sigma_1 = \sigma_2$ and $M_1 \cong M_2$, while

$$(\Delta U/U)_2 \simeq 0.5\left(\frac{3.5\,\sigma}{v}\right)^2 \qquad (2)$$

for

$$v \geqq 3.5\,\sigma.$$

We now turn to the $p \neq 0$ collisions. Since we will not have such a strong radial contraction in these cases the differences between the $N=94$ and $N \simeq 30$ galaxy collisions will become smaller. It is not simple to isolate the v- and p-dependences of the energy changes, especially since the variations with velocity are complicated by the radial oscillations of the $N \simeq 30$ galaxies. The ratio of the loss of orbital kinetic energy at a particular impact parameter p and velocity v to the loss at the same velocity with $p=0$ is plotted versus p on Fig. 2. The data cover a small range of collision velocities $3\sigma < v < 4.5\sigma$, but the collisions with high velocity (filled circles) and low velocity (open circles) do not have systematically different values of ΔE, which we assume has the form

$$\Delta E(p,v) = f(p)g(v).$$

A reasonable linear fit is given by

$$\Delta E(p,v) = \Delta E(p=0,v)\,(1-p). \qquad (3)$$

The dependence of ΔU on impact parameter now follows since ΔU and ΔE must have the same p-dependence (see 2.3.6).

The critical velocity, v_{crit}, the largest velocity for which merging occurs, decreases with increasing p. In Fig. 3 values of the velocities and impact parameters for some mergers are given. These points all lie slightly below $v_{crit}(p)$. In order to account for the small differences in internal velocity dispersion the velocities have been normalised to the average velocity dispersion of the two merging galaxies. The p-dependence of the critical velocity can be approximated by the following linear relation:

$$v_{crit}(p) = 3.1\,(1 - 0.4\,p), \quad p < R_g \qquad (4)$$

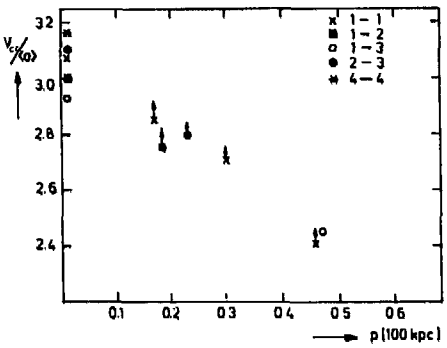


Fig. 3. Relative collision velocities versus smallest separation for some mergers. The velocities are divided by the average internal velocity dispersion of the colliding galaxies. These points are lower limits on the critical velocity for which merging occurs at some values of p

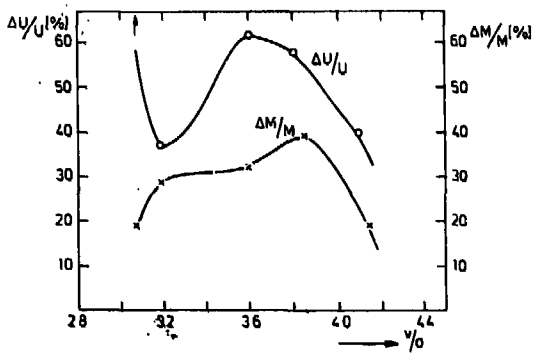


Fig. 4. Fractional increase of internal energy and mass-loss for galaxy 1 after collisions with an identical galaxy

where p is given in units of 100 kpc and R_g is the galactic radius. Note that there seems to be very little or no dependence on the mass-ratio of the colliding galaxies.

2.3.2. Mass-loss

The actual amount of mass that is lost from a galaxy after a collision depends on the average number density or the average separation of the galaxies in the neighbourhood of the collided galaxy. This average separation is approximately given by

$$S = R/N^{1/3} \tag{5}$$

where R is the radius of the system of galaxies, and N is the number of galaxies of the system. Since we want to apply our results to groups of galaxies with $R \sim 2$ Mpc and $N \sim 50$ we define mass-loss in a collision as the fraction of the stars of a galaxy that is capable of reaching a radius $> \frac{1}{2} R/N^{1/3} = 150$ kpc. With this definition we might underestimate the mass-loss from galaxies in a cluster through collisions because after a collision the galaxies expand and it will be easier to strip the outer parts of such a galaxy in a second collision.

In general $\Delta M/M$ and $\Delta U/U$ behave similarly as a function of the collision velocity (Fig. 4). There seems to be no significant dependence of $\langle \Delta M/M \rangle / \langle \Delta U/U \rangle$ on the velocity. Taking averages over different collision velocities it appears that $\langle \Delta M/M \rangle$ is slightly larger than $1/2 \langle \Delta U/U \rangle$ in the cases $1-1$ and $1-2$ and slightly below $1/2 \langle \Delta U/U \rangle$ in collisions between galaxies 2 and 3 which have smaller radii.

Sastry and Alladin (1977) have noted already that

$$\langle \Delta M/M \rangle / \langle \Delta U/U \rangle$$

might be larger for $p \neq 0$ than for head-on collisions. We find:

$$\langle \Delta M/M \rangle = (0.45 \pm 0.2) \langle \Delta U/U \rangle \quad \text{for} \quad 0 < p < 0.3\ R_g$$

and,

$$\langle \Delta M/M \rangle = (0.7 \pm 0.4) \langle \Delta U/U \rangle \quad \text{for} \quad p > 0.3\ R_g \tag{6}$$

Of course we have to treat mergers separately (see Fig. 4). For mergers we find roughly (with our definition of mass-loss)

$$\langle M_{\text{merger}} \rangle \simeq M_1 + \tfrac{1}{2} M_2, \quad \text{where} \quad M_1 \geqq M_2. \tag{7}$$

If $M_1 = M_2$ both galaxies will lose $\sim 1/4$ of their mass, while in the case $M_1 \gg M_2$ galaxy 1 will not be perturbed very much, but galaxy 2 will lose about half its mass.

Equation (7) is certainly a very rough estimate because the mass of the merger will depend on the relative velocity at infinity and probably on the impact parameter. The velocity dispersion of the merger is about the same as the velocity dispersion of the merging galaxies. When the energy of the lost mass is about zero we can write:

$$U_{\text{merger}} = U_1 + U_2 + \frac{1}{2} \frac{M_1 M_2}{M_1 + M_2} v_\infty^2$$

or

$$-\frac{1}{2} M_{\text{merger}} \sigma^2 = -\frac{1}{2} M_1 \sigma^2 - \frac{1}{2} M_2 \sigma^2 + \frac{1}{2} \frac{M_1 M_2}{M_1 + M_2} v_\infty^2,$$

or

$$M_{\text{merger}} = M_1 + M_2 - \frac{M_1 M_2}{M_1 + M_2} \frac{v_\infty^2}{\sigma^2}. \tag{8}$$

As we will see in Sect. 2.3.4 merging takes place at relative velocities at infinity lower than about $1.2\,\sigma$ so whenever a merger takes place the mass of the merger will lie between M_1 and $M_1 + M_2$ where $M_1 \geqq M_2$. This might explain the average value we find in Eq. (7). We could also use as a crude approximation of (6)

$$\Delta M/M = 0.5\ \Delta U/U,$$

which means with Eq. (1)

$$\Delta M_1 = \Delta M_2.$$

In the velocity region around $v \simeq 3.5\,\sigma$ (and small p) this will be about 1/4 of the mass of the smallest galaxy. This gives for the sum of the masses of the galaxies after the collision, $M_1 + 1/2\, M_2$, where $M_1 \geqq M_2$.

2.3.3. Structural Changes of the Galaxies

Figure 5 shows the total mass $M(R)$ inside radius R as a function of the radius for galaxy 1 before and after collisions with an identical galaxy. There are indications that the outer parts lose more mass than the central parts. Figure 6 shows the mass distributions for galaxy 1 and some mergers of galaxy 2 with its

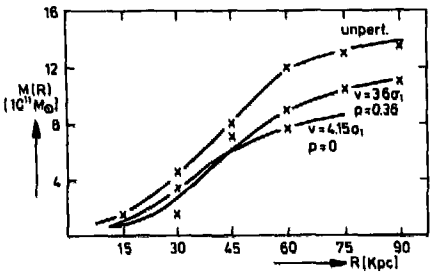


Fig. 5a. Mass-distribution of galaxy 1 before and after collision with its twin

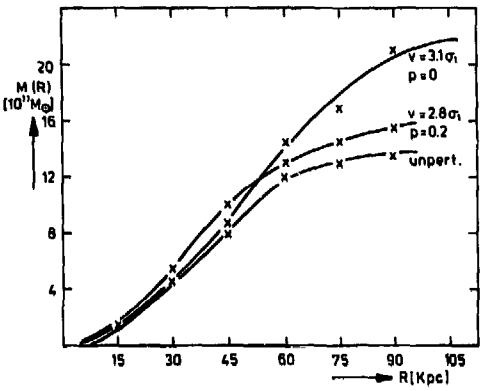
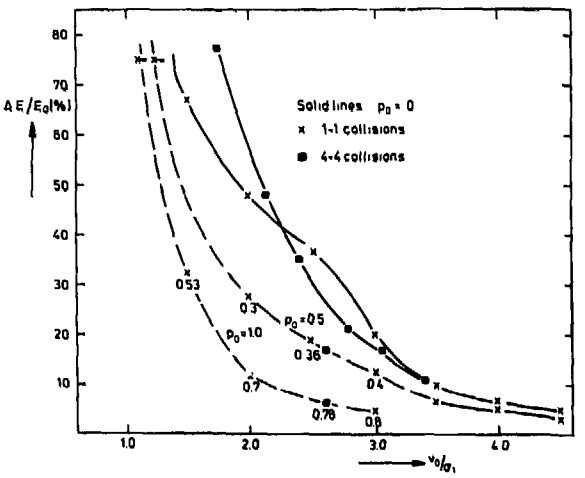


Fig. 5b. Mass-distribution of galaxy 1 and of some mergers of galaxy 1 with its twin



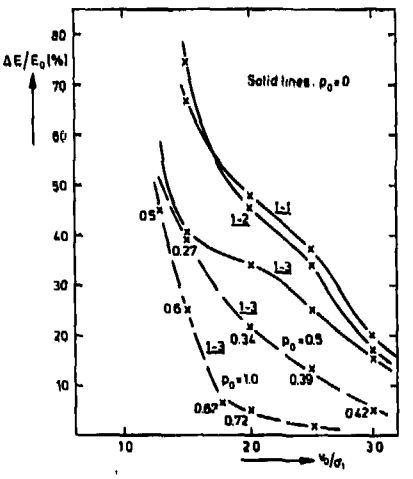


Fig. 6a and b. Loss of orbital kinetic energy calculated at a relative distance of 400 kpc (see 2.3.6). Values of p are given below the crosses

twin. Mergers tend to have a higher central density than the unperturbed galaxy, a result that is in agreement with White (1978).

2.3.4. Transfer of Angular Momentum

In collisions where the loss of orbital kinetic energy was not very large the relation

$$pv = p_\perp v_\perp \tag{9}$$

was valid to an accuracy better than 5%, which means that the loss of orbital energy and consequently the loss of orbital angular momentum occurs mainly after the two galaxies have reached minimum separation. Our small particle number galaxies do not allow us to make explicit statements on the rotational structure of the collided galaxies, but a large portion of the lost orbital angular momentum is carried away by the particles that are "lost" from the galaxies. Also in the case of merging, a large fraction of the orbital angular momentum is transferred into the outer parts of the galaxies. The central parts are brought together with much lower angular momentum per unit mass so that the angular momentum per unit mass of the final merger will increase with radius.

2.3.5. Relation between (p, v) and $(p, v)_\infty$

We need two relations connecting the collision parameters at infinity and at the moment of closest approach. The first one is given by (9). This relation breaks down when the inelasticity is very large. A second relation can be derived easily for head-on collisions if we assume that the galaxies do not change during the collision. This relation states that the kinetic energy at infinity is equal to the sum of the kinetic energy and the extra potential energy $\Delta\phi$ that the galaxies feel when they are overlapping. When the masses are equal we get

$$\tfrac{1}{4}Mv^2 = \tfrac{1}{4}Mv_\infty^2 + |\Delta\phi|$$

$$|\Delta\phi| = |\Delta\phi_1| + |\Delta\phi_2| = 2M\sigma^2$$

so

$$v^2 = 8\sigma^2 + v_\infty^2. \tag{10}$$

When two unequal mass galaxies are overlapping ($\sigma_1 = \sigma_2$, $M_1 \geqq M_2$) we have (see Sect. 2.3.1)

$$|\Delta\phi_1| = |\Delta\phi_2| = M_2\sigma^2.$$

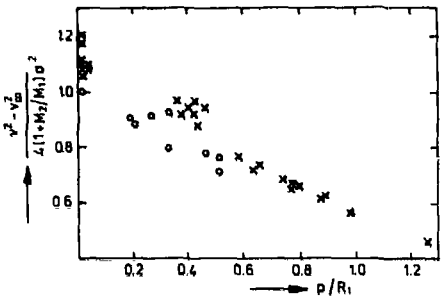


Fig. 7. The factor f defined in (12) plotted against p for a large number of collisions

Equation (10) then becomes

$$v^2 = 4(1 + M_2/M_1)\sigma^2 + v_\infty^2. \tag{11}$$

Let us assume that when $p \neq 0$ $\Delta\phi$ has to be multiplied by some function $f(p)$: $\Delta\phi(p) = \Delta\phi(p=0) \times f(p)$. This gives

$$v^2 = 4(1 + M_2/M_1)\sigma^2 f(p) + v_\infty^2. \tag{12}$$

The factor f is plotted versus p in Fig. 7 for a large number of collisions. First of all we see that in the case $p=0$, $f(p)$ is somewhat larger than expected on the basis of Eq. 11. The reason for this is that it was assumed in the derivation of this equation that the galaxies were rigid during a collision, $|\Delta\phi|$ and therefore v will be somewhat higher in those cases where the galaxies undergo a significant contraction during the collision. When v is close to the critical velocity (the circles in Fig. 8), the inelasticity is already very strong and reduces v. For large impact parameters the resonant behaviour disappears and we see that the scatter in Fig. 8 is much smaller in this case. A good fit to all the points in Fig. 8 is

$$f(p) = 1.1\,(1 - 0.5\,p/R_1), \quad v < 4\sigma, \quad p < R_1.$$

For large velocities ($v_{coll} > 4\sigma$) the function will probably look more like

$$f(p) = 1 - 0.4\,p/R_1.$$

2.3.6. Loss of Orbital Kinetic Energy

In Fig. 6a the dependence of the fractional change in orbital kinetic energy measured at a separation of 400 kpc on v_0 and p_0 is shown for collisions between identical galaxies. Figure 6b shows $\Delta E/E_0$ for head-on collisions of galaxy 1 with galaxies 2 and 3 and $p \neq 0$ collisions of galaxy 1 with galaxy 3. Velocities, fractional energy changes and impact parameters at infinity can be calculated using

$$v_\infty^2 = v_0^2 - 0.54\,\sigma^2(1 + M_2/M_1), \quad M_1 \geqq M_2.$$

Now let us analyse some of the characteristics of these figures. Because the total energy in the system of the two colliding galaxies is conserved we have for the decrease in the orbital kinetic energy:

$$\Delta E = \Delta U_1 + \Delta U_2. \tag{13}$$

This relation together with (1), (2), and (3) yields

$$\Delta E/E_{(0)} = \left(\frac{3.5\sigma}{v}\right)^2 (1 - 0.8\,p/R_1)\,(1 + M_2/M_1)\,\frac{\sigma^2}{v_{(0)}^2}, \tag{14}$$

with $M_1 \geqq M_2$, and R_1 the radius of galaxy 1. Note that (12) can be used to calculate v. At high velocities the dependences on the mass-ratio and impact parameter explicitly shown in (14) give a good explanation of the behaviour of $\Delta E/E_0$ shown in Fig. 6. At low velocities ($v_0 \lesssim 2\sigma$) the first term in (12) becomes important reducing the sensivity of $\Delta E/E_0$ to M_2/M_1 and p. Neglecting the second term in (12) and using $p = p_0/2$ we find for $\Delta E/E_0$ at velocities close to the critical velocity

$$\Delta E/E_0 = \frac{3\sigma}{v_0^2}\,\frac{1 - 0.4\,p_0/R_1}{1 - 0.2\,p_0/R_1}. \tag{15}$$

This explains why the critical velocity measured at the initial separation is only a very weak function of M_2/M_1 and p. Galaxies therefore have a large cross-section for merging.

Since $\Delta E/E_0$ will have the same value at the critical velocity, we find

$$v_0^{crit}(p) = 1.6\,\sigma \left\{\frac{1 - 0.4\,p_0/R_1}{1 - 0.2\,p_0/R_1}\right\}^{1/2}, \quad p_0 \lesssim 2R_1$$

Equation (14) suggests further that

$$v_{crit}(p) = v_{crit}(p=0) \left\{(1 - 0.8\,p/R_1)\left(\frac{1 + M_2/M_1}{2}\right)\right\}^{1.4} \tag{16}$$

This is not a bad representation of the data given in Fig. 3. Calculations with a larger number of particles suggest that a value of 3.2 σ for the critical collision velocity in head-on collisions between equal galaxies is probably better than 3.1 σ. This value corresponds to $v_0 = 1.6\sigma$, or $v_\infty = 1.2\sigma$. A smaller softness parameter might give a critical velocity that is even a few percent higher.

3. The Evolution of Groups and Clusters of Galaxies

3.1. Introduction

In Sect. 2 we have shown that at low relative velocities, collisions between galaxies will have strong effects. Such effects have not been included in previous work on the dynamics of clusters of galaxies. We will show here that collisions between galaxies might explain several observed features of groups and clusters of galaxies if galaxies are formed with massive halos. Initially all our galaxies have a mass of $1.4\ 10^{12}\ M_\odot$ and a radius of 80 kpc. Galaxies in groups and clusters will have about this mass if all the mass of the cluster is bound to the galaxies. We further assume that all galaxies in a cluster are formed before the region of the universe to which they belong reached maximum expansion. When the calculations are started the clusters do not have any systematic radial motion.

3.2. Description of the Calculation

The same code used for N-body collisions in Sect. 2 is used here but with a larger softness parameter. This is an important parameter in the code since our galaxies must not behave like point masses at small distances thus reaching high collision velocities. If ε were chosen to be too large the velocity at the moment of collision would be underestimated and the effects of the collision overestimated. We have adopted $\varepsilon = 25$ kpc. This value will give collision velocities for head-on collisions that are a bit too large, but it is a good approximation for larger impact parameter collisions which are probably the most numerous.

To use the results of the galaxy collisions in a cluster calculation we have made some model fits to the data. It is not necessary to

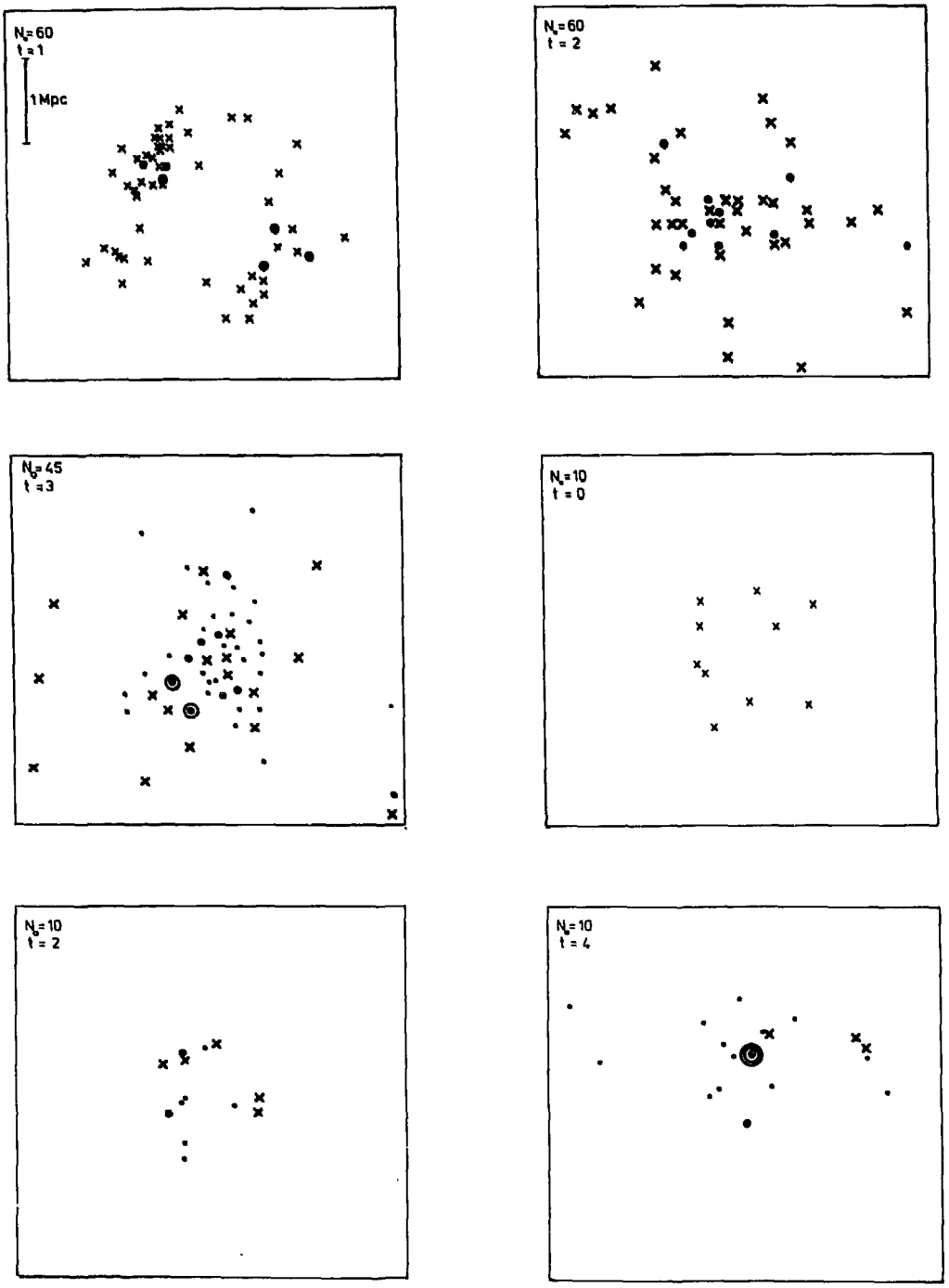


Fig. 8a–f. Projected distributions of the galaxies after the evolution of the models. The scale is the same for all pictures. (Fig. 12a) Crosses are galaxies. big dots are merged galaxies, big dots in circles are multiple mergers and small dots are background particles stripped from the galaxies. The time unit is 6.6 10^9 yr

Table 2

Model	N_0	R_0 mpc	$\langle\rho_0\rangle$ 10^{-28} g/cm^3	t 6.6 10^9 yr	$\langle R\rangle$ 100 kpc	$\langle R\rangle_M$	σ 300 km s^{-1}	σ_M	N_M %	N_{tot}	N_B
1	10	1.0	2.5	1	3.8		1.2		25	8	8
				2	10.7		0.6		25	8	8
				3	11.5		0.9		28	7	15
				4	18.4		0.67		33	6	16
2	45	2.0	1.4	1	5.1		2.1		10	41	18
				2	14.3		1.2		29	35	41
				3	19	12.5	1.0	0.9	30	33	49
3	60	2.5	0.96	1	9.7		1.9		11	54	22
				2	12.7	7.2	1.25	1.18	20	50	47
4	30	2.5	0.5	1	13.3		1.0		0	30	2
				2	6.8		1.5		4	28	8
				3	14.4		0.8		11	27	15
				4	16.8	13.1	0.9	0.85	11	27	15

use exact data for each collision since we are mainly interested in the over-all effects of collisions in a cluster calculation.

For the critical collision velocity at zero impact parameter we adopt

$$v_{crit}(p=0)=700 \text{ km s}^{-1}$$

For the dependence of the critical velocity on the impact parameter we use Eq. (4).

The loss of orbital kinetic energy is approximated by

$$\Delta E/E=0.1\,(1-p/80 \text{ kpc}), \qquad (17)$$

If $p>80$ kpc or $v>1000$ km s^{-1} the collision is inelastic. To calculate the mass-loss in the case of a merger we applied (7) and in the region $700<v<1000$ km s^{-1} we used $\Delta M/M=0.25$. The calculation is started by choosing random positions for the galaxies in a sphere of radius $1\leqq R_0\leqq 2.5$ Mpc according to a density distribution of the form $n(r)\propto r^2(1-r/R_0)$. This density distribution has a maximum in a spherical shell at a distance of 2/3 R_0 from the cluster centre. During the collapse this shell breaks up into subclumps (see Fig. 7a) which merge into one cluster. The initial velocity distribution is Maxwellian with a small velocity dispersion of 60 km s^{-1}. Initially galaxies are further than 300 kpc apart. In every collision the total momentum of the centre of mass of the two galaxies is conserved. The mass-loss that a galaxy undergoes is stored until it exceeds one-quarter of the initial mass of the galaxies. A background particle having this mass is then created and sent away radially from the parent galaxy in a random direction starting from a distance of 150 kpc with a velocity which we choose to be 60 km s^{-1}.

During the evolution of a cluster the total energy must be conserved i.e. the sum of the potential and kinetic energies of the point masses and the internal energies of the galaxies. We are unable to treat mass- and energy-losses in a completely consistent way because of our lack of knowledge of the structure of the galaxies after a few collisions. We realize that our choice of treating mass-loss is still rather crude and that a different choice might influence the final mass-distribution of the cluster.

The maximum number of point masses in these calculations is larger than in the collision calculations so the IBM 370/158 machine was used for this part of the work.

3.3. Results and Interpretation

A sequence of 4 models was run with an initial radius lying in the range 1–2.5 Mpc and with N galaxies, where $N=10, 30, 45$, and 60, randomly distributed in the initial volume as described in the previous section. The results are given at different epochs for these models in Table 2. The first four columns give the properties of the models at $t=0$. Columns 6 and 7 give the average distance to the cluster centre for all the galaxies and the merged galaxies. Columns 8 and 9 give the velocity dispersion for all the galaxies and for the merged ones. Columns 10, 11, and 12 give the percentage of mergers, the total number of galaxies and the number of background particles. For comparison, we did the same calculations for the case of elastic collisions. It is clear that in each simulation collapse has occurred. From previous N-body calculations by Peebles (1970) it is known that virialisation occurs after approximately twice the collapse time starting from maximal expansion. The collapse and virialisation times are proportional to $\rho_0^{-1/2}$ where ρ_0 is the initial mass density. Consequently model 1 is expected to reach virialisation first. At the end of the calculation models 1 and 2 have entered the virialised regime. Models 3 and 4 are approaching this virialised state. Projected galaxy distributions for model 1, 2, and 3 are given at various epochs in Fig. 8.

3.3.1. Number of Mergers and Strongly Interacting Binaries

The percentage of mergers as a function of time is given for the four models in Fig. 8 with timescales proportional to the collapse time of model 2. It can be seen that merging occurs predominantly during the collapse. After the collapse mergers are about 20% of the total number of galaxies. There is only a slow increase in the number of mergers after collapse because the velocity dispersion of the galaxies increases and thus the probability of mergers which depends on low velocity collisions, decreases. The model predicts a maximum in the number of interacting binaries in the collapse stage of a group or cluster while the resulting number of merged galaxies is still small. It may be possible to check this observationally, the decay of the orbit of the two galaxies that will eventually merge can take $\simeq 2\ 10^9$ yr, but the stage in which the two central parts are merging together will only last a few

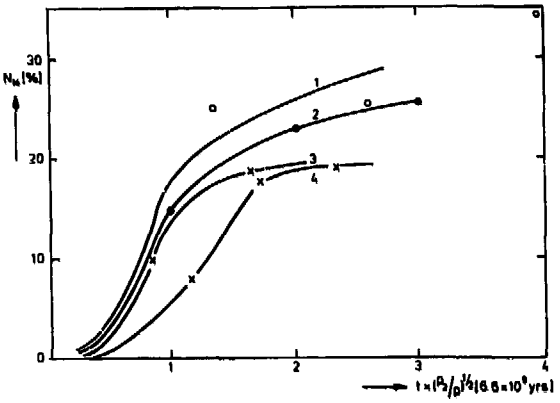


Fig. 9. Percentage of mergers as a function of time. The timescales are normalised in such a way that at each time the models are at the same stage of collapse as model 2

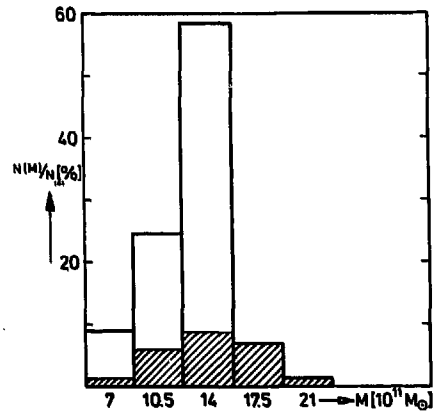


Fig. 10. Mass-distribution for the galaxies at the end of the calculation using models 2 and 3. The dashed area is the distribution for the merged galaxies

Table 3

	$N_0=60$		$N_0=45$	
	$\langle N_{tot}\rangle$	⟨% mergers⟩	$\langle N_{tot}\rangle$	⟨% mergers⟩
$0<R<0.7$ Mpc	10.5	45	11.7	38
$0.7<R<1.0$ Mpc	8.5	29	8	27
$1<R<3.0$ Mpc	30.5	6.5	15.7	9.3
Total		19%		24%

times 10^9 yr, so the percentage of interacting binaries that one can see might be very low. The percentage of mergers after collapse seems to depend only slightly on the density and agrees remarkably well with the percentage of ellipticals that is found in regions of different galaxy density (Bahcall, 1977). The results in Fig. 9 can be understood on the basis of the following simple estimation. The mergers are formed during collapse of the cluster while the dispersion of the galaxies in the cluster is still smaller than the internal velocity dispersion of the galaxies. Starting with a velocity dispersion of about zero the time needed to accelerate the galaxies to σ in a cluster mass M_{cl}, radius R_{cl}, density ρ_{cl} is

$$\Delta t=\sigma\frac{R_{cl}^2}{GM_{cl}}=\frac{3\sigma}{4\pi GR_{cl}\rho_{cl}}.$$

The total number of mergers formed in this interval divided by the total number of galaxies is approximately

$$\frac{N_M}{N}=\text{collision rate}\times\Delta t.$$

For the collision rate we can write

$$\text{collision rate}=\frac{\rho_{cl}}{M_g}\pi(2R_g)^2\sigma,$$

where $\pi(2R_g)^2$ is taken for the cross section of a galaxy near the critical velocity. This is based on the results of 2.3.6. Putting these estimates together and using for galaxies mass M_g

$$M_g=\frac{\sigma^2R_g}{G}$$

we get

$$\frac{N_M}{N}\simeq\frac{3R_g}{R_{cl}}. \tag{18}$$

This is of course a rather crude estimation but it shows that the percentage of mergers will be of the order of 20% after the collapse of a cluster and that this percentage does not depend very much on the density of the cluster. The results presented in this section show that it is certainly possible to explain the overall percentage of elliptical galaxies via the merging picture as first suggested by Toomre and Toomre (1972). The actual percentage we have found probably depends on our initial conditions but it is not possible to prevent a considerable fraction of the galaxies from merging during the formation of groups and clusters if galaxies are formed with massive halos. Elliptical galaxies are the most likely candidates for these merged systems. In a subsequent paper the dependence of our results on the initial conditions will be investigated.

One might think of a scenario for the evolution of galaxies in which the gas lost from the stars in the massive halo during stellar evolution falls to the centre, dissipates its energy and settles down in a rotating disk when the halo has just a small amount of rotation. When two such galaxies merge the central luminous disk will be brought together and will also merge. We speculate that the parameter which may most strongly determine the ellipticity of the final merger is the angle between the disks of the two merging galaxies while the rotation of the merged galaxies will be determined by the orbital angular momentum of the colliding galaxies and their spins. Most of the orbital angular momentum will be transferred into the outer parts of the merger so that it might be possible to explain the low rotational velocities of elliptical galaxies in the merging picture (Illingworth, 1977). It seems likely that the central luminous part of the merger has a lower gas content than the disks from which it was formed because the relative velocity of the two disks during a merger is high enough to heat the gas in the disks to temperatures of 10^7 K if cooling is not important. This hot gas will quickly expand into a much larger volume than it occupied before the collision.

3.3.2. Distribution of Galaxies of Different Morphological Type

From Table 2, but also by visual inspection of Fig. 8 we see that the distribution of merged galaxies has a smaller mean radius than the whole system of galaxies. This type of segregation is also observed in clusters of galaxies between elliptical galaxies and disk galaxies, despite the fact that the elliptical and SO populations are often taken together (van den Bergh, 1978). Table 3 shows how the composition of our models 3 and 4 changes with distance to the cluster centre. The averages are taken over a few models that have the same initial parameters but different random numbers for the initial velocities and positions.

Because ellipticals should occur mostly in collapsed regions and because their distribution will have a smaller characteristic radius the amount of clustering of ellipticals must be higher than that for spirals. Davis and Geller (1976) give supporting evidence for such an effect in the two-point correlation function for ellipticals. Besides the spatial segregation between ellipticals and other galaxies there is some indication of velocity segregation in the observations (van den Bergh, 1978; Moss and Dickens, 1977). In agreement with these observations the merged galaxies in our model tend to have a somewhat smaller velocity dispersion than the other galaxies.

We conclude that merged galaxies are more strongly bound in a cluster than the other galaxies. In our model this is not caused by two-body relaxation or dynamical friction. The average mass of the mergers is not much larger than the average mass of the other galaxies (see Fig. 10). Merging occurs among low velocity galaxies in high density regions where the collision rate is high.

Multiple mergers occur preferentially in the central regions of clusters and groups (see Fig. 8) and we associate such cases with giant ellipticals and cD galaxies. It is understandable that cD's can occur in groups (Albert et al., 1977) as well as in rich clusters. Because clusters are probably formed through merging of subclumps (White, 1977) which often have a multiple merger in their centre, cD's need not be found right in the centre of rich clusters.

Our prediction that interacting galaxies should occur predominantly outside collapsed regions where the percentage of ellipticals is low is in agreement with Thompson's observation (1977) that ring galaxies lie preferentially outside rich clusters (if one adopts the interpretation of ring galaxies by Lynds and Toomre, 1976).

3.3.3. Background Material

Table 2 shows that at the end of the calculation the stripped mass amounts to only 1/4 of the total mass. This stripped mass has about the same radial distribution as the unmerged galaxies, but is slightly more centrally concentrated. We have already noted in 2.3.1 that we have probably underestimated mass-loss. The distribution of the stripped mass in our model 3 is slightly more centrally concentrated than the distribution of the unmerged galaxies. The total mass within 0.5 Mpc of the centre is about 10% higher in the inelastic case than in the elastic case. This might have consequences for the X-ray morphology of clusters of galaxies.

In Fig. 10 the mass-spectrum at the end of the calculation is shown. The differences in mass have not become large. During the merging process just a small amount of mass is gained, but merged galaxies also lose mass through collisions with other galaxies. The model predicts that merged galaxies are embedded in considerable amounts of stripped mass.

3.3.4. M/L-values

In applying the virial theorem to poor clusters or groups, a large part of the stripped mass might be missed when the radial distribution of this matter is broader than the radial distribution of the galaxies. This would lead to a lower M/L-value for these systems. At present our results only give hints in that direction.

Several authors have noted an increase in the M/L-value going from groups of galaxies to rich clusters. Gott and Turner (1977) studied a sample of 39 groups of galaxies and discuss several important difficulties in extracting M/L-values for them such as: a) low number of galaxies with known redshift, b) contamination by foreground and background galaxies, c) projection effects, and d) non-equilibrium effects. They found a mass-to-light ratio for groups that is about a factor 2–3 smaller than for rich clusters. Rood and Dickel (1978) have completed an extensive study on the M/L-values for groups and they have found correlations of the form

$$M/L \propto V^{1.6} R.$$

But, as pointed out by Turner and Sargent (1974), such a relation can easily be created in an artifical way when the virial theorem is not applicable. Application of the virial theorem gives of course the wrong results if: spurious galaxies are counted as group members or the group has not yet collapsed or virialised. Turner and Gott (1976) point out that on the average the groups are just entering the virialised regime. Many of these groups might still be in their collapse stage. Application of the virial theorem to groups of galaxies especially to loose groups is therefore very uncertain. Gott and Turner (1977) found that $\langle M/L \rangle = 200$ for groups with $L > 10\,L^*$ and $\langle M/L \rangle = 65$ for groups with $L < 10\,L^*$, where $L^* = 3.4\,10^{10}\,L_\odot$. They note that this might correspond to the larger portion of spirals in the low-luminosity groups. This agrees with the difference they found between the M/L of the early type (E, SO) and the late type (S, Sb, Irr) population in their groups

$$\langle M/L \rangle_{E,SO} = 136 \pm 50$$

$$\langle M/L \rangle_{Sp,Irr} = 61 \pm 21.$$

This might of course mean that the M/L of individual galaxies in the two classes differs by a factor of about 2, but this would be in contradiction with the most recent information on M/L-values for galaxies of different types from Dickel and Rood (1978). They find no difference in the M/L-values for spirals and for ellipticals. Another interpretation is now possible on the basis of the results presented in this paper, namely, that the low-density spiral-rich regions such as loose groups tend to give low M/L-values because these regions have not yet undergone collapse.

4. Conclusions

Our principal conclusions are as follows:

4.1. Galaxy Collisions

(i) Galaxy collisions tend to be strongly inelastic at collision velocities smaller than 4.5 σ, and minimum separation smaller than the radius of the largest colliding galaxy.

(ii) The criterium for merging to take place between two galaxies with equal internal velocity dispersion σ is

$$v \leqq 3.1\,\sigma\,(1 - 0.3\,p/R_g)\left(\frac{1 + M_2/M_1}{2}\right)^{1/4}, \quad M_2 \leqq M_1, \quad p/R_g < 1$$

for the collision parameter at the moment of collision, or

$$v_\infty \leqq 1.1\,\sigma\,(1-0.25\,p_\infty/R_g), \quad p_\infty/R_g < 2$$

for the collision parameters measured at infinity. Here, R_g is the radius of the largest galaxy.

(iii) At higher velocities the loss of orbital kinetic energy is given by (14).

4.2. Evolution of Groups and Clusters of Galaxies

(i) During the collapse of clusters merging takes place which can result in $\approx 20\%$ merged galaxies after collapse. After collapse the percentage of merged galaxies rises only slowly.

(ii) Multiple mergers can be found in the subclumps that form before the collapse of the cluster and afterwards in the central region of the cluster.

(iii) There is a clear increase in the percentage of mergers towards the cluster center.

We regard points (i), (ii), and (iii) as strong support for the idea of Toomre and Toomre that elliptical galaxies might be the remnants of merged galaxies. Giant ellipticals and cD's might very well be the products of a multiple merging process.

Acknowledgements. It is a pleasure to thank T. van Albada, I. King, and J. Silk for stimulating discussions and encouragement, and L. Allamandola, J. Dickel, J. Oort, and A. Toomre for comments on the manuscript. In particular we thank Th. Vaarmeyer for his invaluable assistance with the P. D. P. computer.

References

Albert, C.E., White, R.A., Morgan, W.W.: 1977, *Astrophys. J.* **211**, 309
Alladin, S.M.: 1965, *Astrophys. J.* **141**, 768
Bahcall, N.A.: 1977, *Ann. Rev. Astron. Astrophys.* **15**, 505
Biermann, P., Silk, J.: 1976, *Astron. Astrophys.* **48**, 287
Bosma, A.: 1978, Thesis, Groningen
Davis, M., Geller, M.J.: 1976, *Astrophys. J.* **208**, 13
Dickel, J.R., Rood, H.J.: 1978, *Astrophys. J.* **223**, 391
Gott III, J.R., Turner, E.L.: 1977, *Astrophys. J.* **213**, 309
Hartwick, F.D.A., Sargent, W.L.W.: 1977 *Astrophys. J.* **221**, 512
Illingworth, G.: 1977, *Astrophys. J.* **218**, L43
King, I.R.: 1966, *Astron. J.* **71**, 64
Kormendy, J., Sargent, L.W.: 1974, *Astrophys. J.* **193**, 19
Lauberts, A.: 1974, *Astron. Astrophys.* **33**, 231
Lynds, R., Toomre, A.: 1976, *Astrophys. J.* **209**, 382
Malina, R.F., Lea, S.M., Lampton, M., Bowyer, C.S.: 1978, *Astrophys. J.* **219**, 796
Moss, C., Dickens, A.J.: 1977, *Monthly Notices Roy. Astron. Soc.* **178**, 701
Ostriker, J.P., Peebles, P.J.E.: 1973, *Astrophys. J.* **186**, 467
Ostriker, J.P., Hausman, A.: 1977, *Astrophys. J.* **217**, L125
Peebles, P.J.E.: 1970, *Astron. J.* **75**, 13
Rees, M.J., Ostriker, J.P.: 1977, *Monthly Notices Roy. Astron. Soc.* **179**, 541
Richstone, D.: 1975, *Astrophys. J.* **200**, 535
Roberts, M.S., Rots, A.H.: 1973, *Astron. Astrophys.* **26**, 486
Rood, H.J., Dickel, J.R.: 1978, *Astrophys. J.* **224**, 724
Sastry, K.S., Alladin, S.M.: 1977, *Astrophys. Space Sci.* **46**, 285
Smith, H.: 1977, *Astron. Astrophys.* **61**, 305
Spitzer, L., Baade, W.: 1951, *Astrophys. J.* **113**, 413
Thompson, L.P.: 1977, *Astrophys. J.* **211**, 684
Toomre, A., Toomre, J.: 1972, *Astrophys. J.* **178**, 662
Toomre, A.: 1974, IAU *Symp.* **58**, 347
Toomre, A.: 1977, in *The evolution of galaxies and stellar populations*, eds. B. M. Tinsley and R. B. Larson, New Haven: Yale Univ. Obs., p. 401
Turner, E.L., Sargent, W.L.W.: 1974, *Astrophys. J.* **194**, 587
Turner, F.L., Gott, J.R.: 1976, *Astrophys. J. Suppl.* **32**, 409
Turner, E.L., Ostriker, J.P.: 1977, *Astrophys. J.* **217**, 24
van Albada, T.S., van Gorkom, J.H.: 1977, *Astron. Astrophys.* **54**, 121
van den Bergh, S., de Roux, J.: 1978, *Astrophys. J.* **219**, 352
van den Bergh, S.: 1977, *Vistas Astron.* **21**, 71
White, S.D.M.: 1977, *Monthly Notices Roy. Astron. Soc.* **179**, 33
White, S.D.M.: 1978, *Monthly Notices Roy. Astron. Soc.* **184**, 185

CHAPTER III

Merging of Galaxies in an Expanding Universe

N. Roos

Sterrewacht, Huygens Laboratorium, P.O. Box 9513, 2300 RA Leiden, The Netherlands

Received May 30, 1979; accepted July 30, 1980

Summary. Expanding universe models containing 400 equal mass galaxies are numerically simulated and the effects of galaxy collisions are specifically included. Initially the galaxies are distributed at random with low velocity dispersion relative to the Hubble flow. The merging rate is insensitive to galactic parameters since most mergers occur between galaxies in bound orbits. The final fraction of merged galaxies is found to be approximately 10–40% in both the open and closed universe models. Both galaxy clustering and the evolution of the galaxy mass function seem consistent with a simple infall model. The clustering properties of merged systems in the simulations are in agreement with those observed for bright elliptical galaxies. The clustering properties of lenticular galaxies suggest that they form a class intermediate between ellipticals and spirals. The mass function evolution is investigated further by Monte Carlo simulations on a randomly distributed system of galaxies where the merging probability of a pair of galaxies is chosen to be proportional to their combined mass. Such a merging probability, consistent with the results of the expanding universe simulations, gives a mass spectrum that approaches a self-similar form identical to the Press and Schechter mass function in the case of random initial fluctuations. It is also consistent with the infall model. It is proposed to identify the bulge to disk ratio parameter with the ratio of the infallen mass to the original mass. It is shown that the observed frequencies of different galaxy types as a function of mass and local galaxy density can be modelled in this way.

Key words: cosmology – merging – galaxy types

1. Introduction

Galaxies can interact strongly and this fact must necessarily be incorporated into most investigations of their properties. This paper presents the results of numerical studies concerning the effects of galaxy collisions on the formation, evolution and clustering of galaxies. Our aim is to elucidate and extend current models of galaxy formation that invoke galaxy mergers in an expanding universe.

The general technique adopted here and in recent work (Jones and Efstathiou, 1979, hereafter JE and Aarseth and Fall, 1980, hereafter AF) is to utilize studies of individual galaxy-galaxy collisions (Van Albada and Van Gorkom, 1977; White, 1978; Roos and Norman, 1979, Paper I) and incorporate the relevant collisional cross-sections into standard expanding universe calculations (Miyoshi and Kihara, 1975; Aarseth et al., 1979; Turner et al., 1979; Gott et al., 1980).

The suggestion brought forward by Toomre and Toomre (1972, see also Toomre, 1977) that elliptical galaxies would be the result of mergers between galaxies of comparable size has gained considerable support recently in numerical N-body simulations which showed that (i) the final frequency and clustering properties of merged particles in a study of the evolution of individual groups and clusters (Paper I) and in expanding universe simulations (JE, AF) are very similar to those of bright elliptical galaxies, (ii) the characteristic angular momentum of merged particles in the cosmological experiments agrees with the low rotation velocities found in elliptical galaxies, (iii) merger products of spherically symmetric galaxies of comparable mass tend to have a Hubble density profile (White, 1979; Villumsen, 1979). The principal thrust of this paper is to investigate in detail the clustering properties and mass function evolution of the merged systems and to try to make the natural extension of the merger model of ellipticals to explain the origin of other morphological types as the result of mergers between galaxies of unequal mass.

The next section deals with the numerical experiments. In Sect. 3 the merging rate, clustering properties and mass function of the particles in the simulations are discussed. It is proposed that the latter two properties can be interpreted using a simple infall model. In Sect. 4 observed frequencies and clustering properties of elliptical and SO galaxies are compared with those of merged galaxies in the simulations. We consider briefly the evidence that galaxy mass is indeed correlated with local galaxy density as predicted by the infall model. A model is presented for the evolution of a realistic initial mass function during the epoch of cluster formation. Its implications for the origin of different galaxy types are discussed in Sect. 6.

2. Numerical Experiments

Before discussing the detailed calculations and results of the present work it seems useful to make a comparison with other recent studies. The merging criteria in our calculations are only slightly different from those used by JE and AF. Mass-loss and loss of orbital energy during collisions are additional effects considered here. The total orbital and internal energy of a merging pair was not strictly conserved. This is not important since in this paper we are not primarily interested in the structural properties of the merged systems. In fact the mass-radius relations found by AF, who do conserve total orbital and internal energy, is very similar to the one we have used. In our $\Omega=1$ models most of the de-

celeration of the universe is modelled by an extra force component due to some uniformly distributed hot constituent.

2.1. The Code

The numerical N-body code is basically the same as the one used by Aarseth et al. (1979, hereafter AGT) so we will merely mention the differences here. Collision effects were included. No particles were reflected at the boundary R of the exanding sphere. However, at the end of the run the density in the outer shell between 0.9 R and R was not significantly lower than the overall density. In the case of a critical universe an extra radial force component decelerating the expansion was added, mimicking the effect of a universal component. This force is given at time t by $F(t)=-\frac{1}{2}H^2(t)cr$, where c is the ratio of the density of the dynamically hot component to the critical density, $H(t)$ is the Hubble parameter and $|r|$ is the distance to the centre of the expanding sphere.

2.2. Initial Conditions

Initially all 400 galaxies have equal mass M_g. Following AGT, their luminosity L_g is chosen to be that of a typical bright galaxy, which means $L_g \simeq L^* \simeq 3\,10^{10}\,L_\odot$ (Turner and Gott, 1976; Schechter, 1976). The mass-to-light ratio of the galaxies and the clusters that are formed is then $100\ M_\odot/L_\odot$ ($M_g/3\,10^{12}\ M_\odot$). Adopting a luminosity density of $5\,10^7\ L_\odot\ \mathrm{Mpc}^{-3}$ the final number density of our galaxies should be $n_g \simeq 1.7\,10^{-3}\ \mathrm{Mpc}^{-3}$. The final value of the Hubble parameter in the simulations is $50\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. The initial radius of the expanding boundary sphere R_i is chosen to produce about the right amount of clustering in the final state as measured by the amplitude of the spatial covariance function.

The particles are placed at random inside the initial volume so that the density fluctuations follow a Poissonian distribution. The distribution of peculiar velocities is taken to be an isotropic Maxwellian with a dispersion of $50\ \mathrm{km\ s^{-1}}$, so that the ratio of kinetic energy in random motions to the kinetic energy in the expansion is initially much smaller than unity and the universe is cold as predicted by the standard big bang model. Although the initial velocity perturbations should be related to the density fluctuations in the galaxy distribution, the fluctuations in both the density and velocity distributions are small and therefore we neglect this effect. At the start of the simulations individual galaxies represent density fluctuations

$$(\delta\varrho/\varrho)_g = (1-c)(R_i/R_g)^3/N_i + c - 1$$

where R_g is the galaxy radius, and N_i is the initial number of galaxies. The initial density contrast is $(\delta\varrho/\varrho)_g \sim 1$ in the $\Omega=1$ models and $\sim 10^0$–10^4 in the open models. In some models the sensitivity of the final state to changes in the initial small scale distribution was tested.

2.3. Collision Parameters

As in Paper I, the galaxy potential was taken to be proportional to $(r^2+\varepsilon^2)^{-1/2}$. In a collision, the softness parameter ε has a strong effect on the velocity v and separation p at closest approach. The inelasticity of a collision is determined by v and p and so one has to be careful in choosing ε. The collision velocity measured in the numerical calculations of Paper I is, for $p\ R_1 \lesssim 1$, and $v_\infty \lesssim \sigma_g$, approximated by

$$v^2 = 8.8\,\sigma_g^2(1-0.4\,p/R_1)\left(\frac{1+M_2/M_1}{2}\right)+v_\infty^2, \qquad (1)$$

where $M_1 \geqq M_2$, v_∞ is the relative velocity at infinity, σ_g the internal velocity dispersion of the galaxies (assumed to be equal for the two galaxies) and R_1 is the radius of the largest galaxy with mass M_1. The radius R_1 is defined as twice the half-mass radius and σ_g^2 is proportionality to GM_g/R_g; the constant of proportionality is determined by the structure of the galaxies, which is assumed not to change in collisions. For the galaxies used in the collision calculations we had $\sigma_g^2 = 0.72\,GM_g/R_g$. In the present calculations where galaxies are treated as softened pointmasses the collision velocity is

$$v^2 = 4\,GM_1\left(\frac{1+M_2/M_1}{2}\right)(p^2+\varepsilon^2)^{-1/2}+v_\infty^2.$$

The collision calculations in Paper I indicate that σ_g does not change much in collisions, and so $R_g \propto M_g$. We find

$$v^2 = 5.6\,\sigma_g^2 R_1\left(\frac{1+R_2/R_1}{2}\right)(p^2+\varepsilon^2)^{-1/2}+v_\infty^2 \qquad (2)$$

In order to match (1) and (2) we need $\varepsilon/R_1 = 0.64$ in the case $p=0$. If $p=R_1$ we need $\varepsilon/R_1 = 0.3$. Most collisions occur between galaxies in bound orbits with small ellipticities more consistent with the value 0.6. In the simulations ε/R_1 ranged from 0.34 to 0.57. Although the velocities for the nearly head-on collisions will be overestimated with small ε/R_1, the loss of orbital energy is still large and generally the collision velocity was below v_{crit} at the second or third impact.

All collisions with $p > R_1$ were purely elastic. For the critical upper limit to the relative velocity between galaxies with mass M_1 and M_2 ($M_1 \geqq M_2$) for mergers to still occur we have used

$$v_{crit}(p) = 3.2\sigma_g(1-0.3\,p/R_1)\left(\frac{1+M_2/M_1}{2}\right)^{1/4}$$

(see Paper I). Particles may also lose orbital energy and mass in non-merging collisions. The energy loss is then effectuated by decreasing their relative momentum instantaneously at closest approach. The loss of orbital kinetic energy E is prescribed by

$$\Delta E/E = 12.25\left(\frac{\sigma_g}{v}\right)^4(1-0.8\,p/R_1)(1+M_2/M_1) \quad \text{for} \quad p < R_1$$

$$= 0 \quad \text{for} \quad p > R_1$$

and the mass-loss by

$$\Delta M = 6.1\left(\frac{\sigma_g}{v}\right)^2(1-0.8\,p/R_1)M_2 \quad \text{for} \quad v > v_{crit}(p)$$

$$\Delta M = \tfrac{1}{2}M_2 \quad \text{for} \quad v < v_{crit}(p).$$

This mass-loss can be regarded as an upper limit. Firstly, most collision calculations in Paper I were done for hyperbolic orbits while most mergers in the expanding universe simulations occur between galaxies in bound orbits. Secondly, the "lost" mass was defined as the mass ejected across a spherical surface radius $2R_g$ around the parent galaxy. Most of this mass is still bound to the galaxy but forms an extended low density halo which is dynamically unimportant. Experiments without mass-loss were also made. As described in Paper I, the mass that a galaxy should have lost in collisions is stored until it exceeds half the initial mass of the galaxies. A background particle is then created and sent away from the parent galaxy in a random direction at the escape velocity. In all experiments we have used $\varepsilon \propto R_g \propto M_g$ throughout a run.

Table 1. Results of numerical experiments

	N_f	$R_{i,f}$ Mpc	$H_{i,f}$ 100 km s^{-1} Mpc^{-1}	Ω_f	c	σ_f 100 km s^{-1}	M_g 10^{12} $M_\odot$	R_g Mpc	$v_{crit}(p=0)$ 100 km s^{-1}	ε/R_g	A_f	N_M/N_f
1	341	5.0 –35.5	9.2 –0.5	1	0.8	1.15	6	0.34	7.6	0.35	5	0.13
2	309	2.5	26.1			1.4					10	0.21
3	263	0.7	177.0			1.8					30	0.27
4	247	0.7	177.0			2.4		0.07	16.0	0.43	50	0.29
5	266	0.7	177.0			2.1		0.07	16.0	0.43	50	0.29
6	337	0.7	177.0		0.9	1.1	3	0.07	11.3	0.43	10	0.15
7	258	2.06–35	13.7 –0.5	0.1	0.0	1.5	3	0.17	7.6	0.35	15	0.29
8	166	38.3	17.4	0.15		1.6	6	0.07	16.0	0.57	25	0.33
9	232	48.0	19.41	0.08		1.8	6	0.07	16.0	0.57	30	0.28
10*a	218	48.0	19.41	0.08		1.5	6	0.07	16.0	0.57	30	0.33
b	238	48.0	19.41	0.08		1.4	6	0.34	7.6	0.57	30	0.31
c	260	48.0	19.41	0.08		2.1	6	0.01	44.5	0.57	40	0.24
11	176	1.64–10.5	–0.1	$\sim 10^2$		—					$\simeq 400$	0.41
12	326	2.59–51	–0.6	$\sim 2.10^{-2}$		—					$\simeq$ 4	0.16

Column (1) gives the model number. Underlined numbers are standard models which means that the subsequent models until the next standard model have the same parameters unless specified otherwise. Column (2) gives the final number of galaxies ($N_i = 400$) and column (3) the initial and final radius of the expanding boundary sphere. Columns (4)–(7) give the initial and final Hubble parameter, the final ratio of the density over the critical density Ω_f, the value of c and the final velocity dispersion of the galaxies resp. The galaxy mass and radius are given in columns (8) and (9), the critical velocity for head-on collisions between equal mass galaxies in (10) and ε/R_g in (11). The final amplitude of the covariance function A_f in column (12) was estimated from the typical excess number density of galaxies at a distance $0.8 < r < 1.2$ Mpc from a galaxy. Column (13) gives the final fraction of merged galaxies.

* Note: model 10a was run three times and 10b twice with different random numbers. The given quantities are averages over these runs

2.4. Results

A series of N-body numerical experiments was performed with the aim of testing the sensitivity of the final fraction of merged galaxies and the amplitude of the covariance function to various parameters. The details of each experiment are contained in Table 1.

Models 1–3 have different R_i/R_f, where subscript i denotes the initial state and subscript f the final state. The galaxies in model 4 have smaller size than in models 1–3. Models 5 and 6 were constructed to test the sensitivity to the small scale distribution of the particles. The separation for two neighbouring galaxies with initially zero orbital energy is $s = (2\Omega_i/N_i)^{1/3} R_i$. The galaxies in 5 have initial separation larger than 0.53 s. In model 6 the minimum separation was 0.67 s.

In model 8 and 10 the galaxies were constrained to undergo no mass-loss in collisions.

Models 9 and 10 have a lower final number density.

The initial galactic radius, R_g, was varied in models 10a, b, and c.

Models 11 and 12 are similar to 7 except that in 11 the initial density was $2\varrho_i$ (7), and $0.5\varrho_i$ (7) in model 12, where ϱ_i (7) is the initial density in model 7. These models were run for a same time interval as model 7 and can be regarded as regions of high and low density respectively in the model 7 universe. The excess density around a galaxy in these models was calculated with respect to the mean density in the model 7 universe.

The four principal results from Table 1 are:

(i) The evolution of the amplitude of the covariance function is consistent with standard models (Gott and Rees, 1975).

(ii) The amplitude of the covariance function is approximately proportional to the initial galactic mass.

(iii) From our simulations we estimate galaxy collapse times of $\sim 10^9$ yr for $\Omega < 1$ and $\sim 10^8$ yr for $\Omega = 1$.

(iv) The final fraction of merged galaxies (10–40%) is very insensitive to galactic parameters and to Ω_f. It is weakly correlated with the amplitude of the covariance function.

We shall now discuss the first three results. The final point will be discussed in the following section.

(i) We expect that $A_f \propto (1+z_{ii})^\beta$ where z_{ii} is the initial redshift, with $\beta = 2$ in the linear regime and $\beta = 3+\gamma$ thereafter, when the spatial covariance function $\xi(r)$ is described by a simple power law $\xi(r) = A(r/1\ \mathrm{Mpc})^\gamma$ and $(1+z_{ii}) < \Omega_f^{-1}$ (cf. AGT). In our simulations we find $\gamma \simeq -2$ consistent with previous work. We expect therefore that β varies from ~ 2 at early epochs to ~ 1 at late epochs (see also Sect. 3.1). From A_f and R_i/R_f in models 1, 2, and 4 we find $\langle\beta\rangle \sim 1.2$. A_f in model 3 is lower than expected because the initial separation between neighbours was comparable to ε. The value we find for $\langle\beta\rangle$ is somewhat lower than the value 1.6 used by AGT because merging reduces the number of particles on small scales resulting in a decrease of A_f. For example, AGT find $A_f \simeq 70$ when $z_{ii} = 17$ while we find $A_f \simeq 25$ in model 8 which has comparable parameters (M_g, z_{ii}, and A_f; the index of the initial density fluctuation spectrum does not seem to affect A_f). This indicates that merging has reduced A_f by about a factor 3, and in the absence of merging we would expect $\langle\beta\rangle \simeq 1.2 + \log 3 = 1.7$.

(ii) The mass dependence of A_f is most evident when we compare our results for the $\Omega = 1$ universe with those of AGT. They

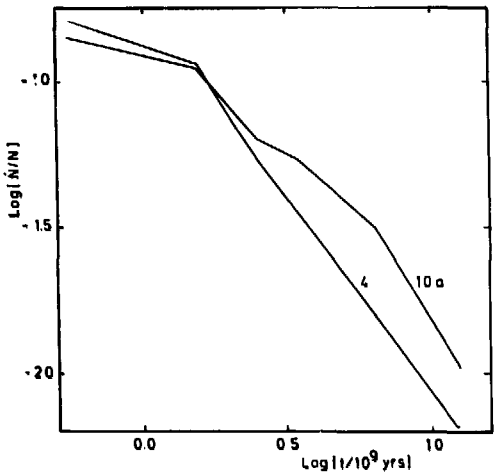


Fig. 1. Evolution of the merging rate in model 4 and 10a. N is the number of mergers per 10^9 yr

give $A_f \simeq 70$ for $z_u = 9.9$ and $M_g = 5\,10^{13}$ $M_\odot$. Assuming $A_f \propto (1+z_u)^{1.7}$ M_g and including a factor 1/3 for the effect of merging, we expect $A_f = 37$ in our model 4, in reasonable agreement with the experimental result. A plausible explanation for the relation between A_f and M_g found here is provided by the infall model of Gunn and Gott (1972). On the basis of this model it is expected that the infall rate onto a galaxy or cluster of galaxies is proportional to its total mass.

(iii) Galaxy collapse times can be estimated for $1+z_u > \Omega_f^{-1}$ by

$$T_c \simeq \pi H_0^{-1} \Omega_f^{-1/2} (1+z_u)^{-1.5}.$$

Using $A_f (1+z_u)^{1.7}$ we find from the results of model 10 that $1+z_u \simeq 40 [M_g/6\,10^{12}\,M_\odot]^{-0.6}$ to obtain $A_f \sim 70$. This yields an estimate of the galaxy collapse time in the case of an open universe with $\Omega_f = 0.1$ of $T_c \simeq 10^9$ yr. In the $\Omega = 1$ universe we find $1+z_u \simeq 60 [M_g/6\,10^{12}\,M_\odot]^{-0.6}$ and $T_c \simeq 10^8$ yr.

3. Properties of Merged Galaxies

3.1. The Total Number of Merged Galaxies and the Merging Rate

The total number of mergers in the simulations is remarkably insensitive to galactic parameters as well as to Ω_f. The reason is that most mergers occur between galaxies in bound pairs (cf. AF). The initial fraction of bound pairs is calculated by assuming that particles are distributed at random with number density n. Then the total fraction of galaxies having their nearest neighbour within radius r is $F(r) = 1 - \exp[-\frac{4}{3}\pi n r^3]$ (Chandrasekhar, 1943). The maximal initial separation for bound pairs, those with zero binding energy, is $s = 0.78\,\Omega_i n^{-1/3}$. Then $F(s) = 1 - \exp[-1.99\,\Omega_i^3]$ which is 86% if $\Omega_i = 1$. Even for an open universe with $\Omega_f \ll 1$, $\Omega_i \lesssim 1$, since

$$\Omega_i = (1 + (\Omega_f^{-1} - 1) R_i/R_f)^{-1}$$

and for $R_i/R_f \sim 23$ and $\Omega_f \sim 0.08$, $\Omega_i = \sim 0.67$ and $F(s) = 45\%$. Therefore, the number of initially bound pairs is very insensitive to Ω_f in our models. The galaxies in a bound pair will eventually merge unless the pair is disrupted in a cluster. From the simulations we see that about 60% of the galaxies merge with their nearest neighbour. The initial velocity dispersion of the galaxies is low and thus the ellipticity of the bound orbits is generally large. Therefore most collisions are almost head-on and the galaxies merge at first impact, forming systems with low characteristic angular momentum (JE, AF). The fraction of merged galaxies can be calculated from the total number of mergers $N_i - N$, where N is the current galaxy number. Assuming that the merging probability is independent of mass, it is not difficult to show that $N_M/N = 1 - N/N_i$, which is typically $\sim 40\%$ at the end of the simulations. The fraction of merged galaxies in the simulations is $\sim 30\%$ indicating that merged galaxies have a higher merging probability than other galaxies.

We now estimate the collision rate assuming that nearest neighbours fall together in the free fall time $(2\,GM_g/R_1^3)^{-1/2}$, where R_1 is the distance between nearest neighbours. This assumption seems justified when nearest neighbours are approaching each other, requiring $\sigma_p > R_1 H$. R_1 can be estimated using the correlation function which can be defined by

$$\xi(r) = \frac{n(r) - \bar{n}}{\bar{n}} = A (r/1\ \mathrm{Mpc})^{\gamma}.$$

Clearly $\xi(r)$ is the typical excess density relative to the mean density $\bar{n}$ at a distance r from a galaxy. The nearest neighbour distance R_1 is then

$$R_1 = \left(\frac{3+\gamma}{4\pi \bar{n} A}\right)^{\frac{3}{3+\gamma}}$$

and the collision rate per galaxy

$$C = \frac{1}{2}\left(\frac{2GM_g}{R_1}\right)^{1/2} \qquad (3)$$

$$= \frac{1}{2}(2GM_g)^{1/2}\left(\frac{4\pi\bar{n}A}{3+\gamma}\right)^{\frac{3}{6+2\gamma}}$$

Substituting the parameters for model 10a at t_f ($\gamma = -2$) we find $R_1 = 5.6$ Mpc and $C = 0.09\ (10^{10}\ \mathrm{yr})^{-1}$. The final collision rate in model 10a averaged over the interval $10^{10} < t < 1.6\,10^{10}$ yr is $0.20 \pm 0.09\,(10^{10}\ \mathrm{yr})^{-1}$. In the next section we will see that the mean mass of colliding galaxies in model 10 is about 2.7 M_g which will increase our estimated C to $0.15\,(10^{10}\ \mathrm{yr})^{-1}$.

In Fig. 1 the evolution of the merging rate in models 10a and 4 is shown. The evolution of the merging rate cannot be described by (3) since $\sigma_p/(R_1 H) \simeq (GM_g/R_1^3)^{1/2} H^{-1} = C H^{-1} \lesssim 1$, at the end of the simulation, but $\sigma_p/(R_1 H) \ll 1$ at earlier epochs. The temporal behaviour of the merging rate is very similar to the increase of mass within a fixed distance D of a galaxy.

$$\frac{1}{M}\frac{dM}{dt} = 4\pi D \frac{d}{dt}[\bar{n} A].$$

Using a proper separation coordinate we have $A = A_f (R/R_f)^{\beta - \gamma}$, where β changes from ~ 2 at early epochs to ~ 1 at late epochs. Assuming $\gamma = -2$ we find

$$\frac{1}{M}\frac{dM}{dt} = 4\pi D \bar{n}_f A_f t^{\frac{2\beta - 5}{3}} \quad \text{if} \quad \Omega = 1$$

$$= 4\pi D \bar{n}_f A_f t^{\beta - 2} \quad \text{if} \quad \Omega < 1.$$

The merging rate at the end of the simulations drops slightly faster than t^{-1} probably because the merging efficiency K, defined by the ratio of the number of mergers over the number of collisions, decreases with t. The evolution of K in some models is given in

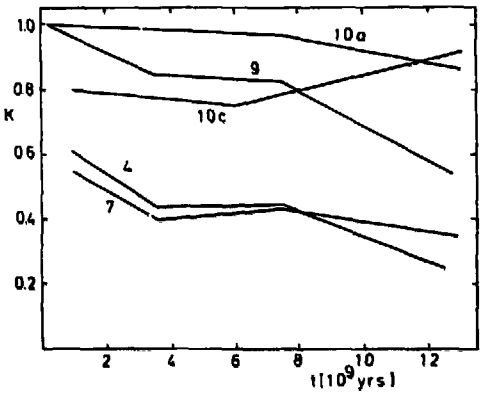


Fig. 2. Ratio of the number of mergers to the number of collisions as a function of time

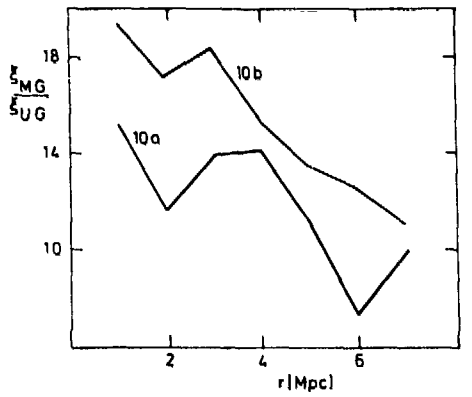


Fig. 3. Ratio of the correlation functions of merged (M) to unmerged (U) galaxies with other galaxies in model 10a and 10b at t_f

Fig. 2. In model 10 $K \lesssim 1$ even at large t, indicating that merging has strongly prevented the formation of virialised clusters with velocity dispersion larger than σ_u. In models 4 and 7, ε/R_u is smaller than in model 10 and it takes a few collisions before the galaxies merge, so that $K \simeq 0.4$. This will increase the covariance function on small scales by about a factor K^{-1} but will not affect the merging rate.

3.2. *Clustering Properties of Merged Galaxies*

The simulations show that on small scales the excess density of galaxies around merged galaxies ξ_{MG} is higher than around unmerged galaxies ξ_{UG} (see Fig. 3). A similar phenomenon was found by Gott et al. (1979, hereafter GTA) who give results for an expanding universe in which two-third of the galaxies were given mass m and one-third mass $2m$. They found that on the scale of a typical interparticle distance R_1 (hereafter called the local scale) the covariance function of the more massive galaxies was about twice that of the covariance function of the less massive ones. On larger scales the two covariance functions become equal. These results are interpreted on the basis of the infall model of Gunn and Gott (1972). In this model the effect of the local density enhancement of a cluster mass m embedded in a homogeneous expanding universe is considered. They show that the infall rate of matter onto the cluster is proportional to m. In our simulations we deal with a discrete distribution of particles of comparable mass. We will show that, as long as virialised groups and clusters of particles do not form, the results of the simulations are consistent with an infall model in which the probability $P(m)$ that a merging pair has total mass m, is proportional to $m\phi(m)$, where $\phi(m)$ is the fraction of all pairs having total mass m. $P(m)$ can be written as

$$P(m) = \frac{\sum_{m_i + m_j = m} (m_i + m_j)}{\sum_{i,j} (m_i + m_j)} \tag{4}$$

where summations are over all possible pairs. Such a probability law is expected if all particles are distributed at random independent of their mass. In this case the probability that a close pair has mass m and separation lying between r and $r+dr$ can be written as $4\pi\bar{n}\phi(m)r^2dr$. The collapse time of a close pair is given by $t_c \simeq (Gm\,r^3)^{-1/2}$ and the probability that a pair has mass m and collapse time in the interval $[t_c, t_c + dt_c]$ is $4\pi\bar{n}Gm\phi(m)t_c dt_c$. For a given collapse time, the probability that a collapsing or merging pair has total mass m is proportional to $m\phi(m)$.

In Fig. 4 we have plotted the function $N_p(m)/\phi(m)$ where $N_p m$ is the number of pairs with total mass m and separation less than R_1. Equation (4) seems to apply throughout the evolution of model 10b, while in 10a the experimental relation deviates from Eq. (4) with increasing time. The particles in model 10a are much smaller and collapsed groups or clusters are less easily destroyed by merging. The merging rate among small galaxies is depressed most strongly and the number of relatively low mass neighbours increases.

The distribution of masses for merging pairs should also obey (4) if merging prevents the formation of clusters. From (4) it follows that the parameter

$$T(m) = \frac{\sum_{m_i + m_j < m} (m_i + m_j) + \frac{1}{2} \sum_{m_i + m_j = m} (m_i + m_j)}{\sum_{i,j} (m_i + m_j)} \tag{5}$$

determined each time a merger occurs between two galaxies of total mass m must be uniformly distributed on $[0, 1]$ with $\langle T \rangle = \frac{1}{2}$. This test was suggested by White (1980, private communication). Figure 5 shows $\langle T \rangle$ for mergers in different time intervals in models 10a, b, and c. The best agreement is found for the models with large galaxy radii. The mean value of all mergers $\langle T \rangle = 0.52$ when $R_u = 340$ kpc, and $\langle T \rangle = 0.58$ when $R_u = 10$ kpc. A further test of Eq. (4) is to count the number of mergers among galaxies of different type during a time interval in which the mass-function is assumed constant. When the merging probability is given by (4) it can be shown that

$$N_{UU} : N_{UM} : N_{MM} = 1 : (\langle m \rangle_M + 1) f_M / f_U : \langle m \rangle_M (f_M / f_U)^2. \tag{6}$$

Here, N_{UM} is the number of mergers among merged and unmerged galaxies, f_M and f_U are the fractions of merged and unmerged galaxies, $\langle m \rangle_M$ is the average mass of the merged galaxies, and unmerged galaxies have unit mass (see Sect. 3.3). In Table 2 the relative numbers of mergers in models 10a, b, c are given with the prediction from (6). The last three columns give the time-

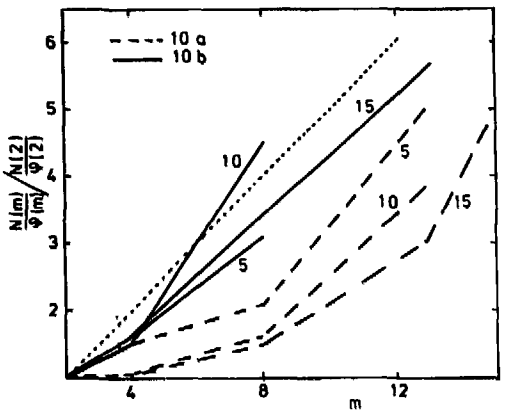


Fig. 4. $N_p(m)$ is the number of pairs with total mass m and separation smaller than the characteristic interparticle distance. The fraction of all pairs having total mass m is $\phi(m)$. The ratio of these quantities normalised to the value for $m=2$ is plotted at different times in models 10a and b. The time is in units of 10^9 yr. The dotted line is the prediction from Eq. (4)

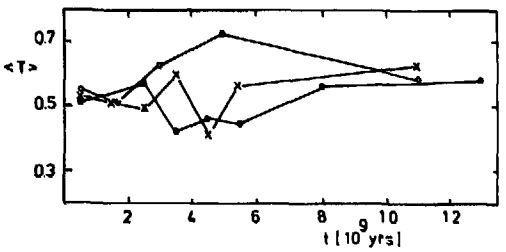


Fig. 5. Evolution of $\langle T\rangle$ [Eq. (5)] in model 10a (crosses), b (dots) and c (open circles). T was determined in a single run of each model

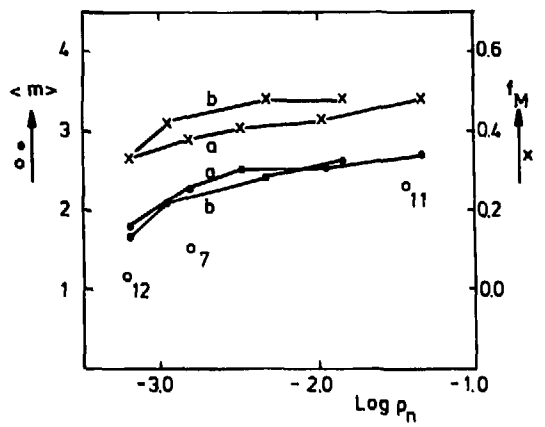


Fig. 6. Mean mass (filled circles) and fraction of merged galaxies (crosses) as a function of neighbour density around galaxies in model 10(a and b). Open circles give the over-all values of $\langle m\rangle$ in models 7, 11, and 12

interval and the average value of f_M and $\langle m\rangle_M$ respectively over this interval. For models 10a, b, and c we find results consistent with (6) at early epochs, but at late epochs N_{VV} is too small in models 10a and c probably due to depressed merging rates for smaller galaxies as described earlier (Fig. 4).

Equation (6) implies that on scales $\sim R_1$, the fraction of massive (or merged) particles is given by

$$f_M' = \tfrac{1}{2}[\langle m\rangle_M/\langle m\rangle + 1] f_M, \tag{7}$$

where f_M is the overall fraction of merged galaxies and $\langle m\rangle$ is the mean mass per galaxy. Inside spheres of chosen radius S, centered on the galaxies, the mean fraction of merged galaxies, mean mass per galaxy and mean number density of galaxies, ϱ_n was calculated. The results for several values of S are plotted in Fig. 6. The open circles give the overall degree of merging measured by $\langle m\rangle = N_i/N_f$ as a function of average density in models 11 and 12 ($S = R_f$), while the filled circles give the degree of merging in model 10 as one goes to smaller scales corresponding to higher densities. On the smallest scales the fraction of merged galaxies is $\simeq 0.45$ in good agreement with (7) since $\langle m\rangle_M/\langle m\rangle \simeq 3.5/1.7 \simeq 2$ and $f_M = 0.3$.

The fraction of merged galaxies and the mean mass of galaxies increases slowly with increasing neighbour density. A crude linear fit to the experimental relation is

$$f_M \simeq 0.2\ \langle m\rangle \simeq 0.1 \log \varrho_n + 0.6. \tag{8}$$

To summarize, the simulations indicate the following clustering properties of merged galaxies.

(i) It is possible to model the merging probability of a pair of galaxies by assuming that it is proportional to the sum of the masses of the two merging galaxies. This dependence of the merging probability on mass implies that the covariance function of merged galaxies is amplified and slightly steepened on small scales.

(ii) The fraction of merged galaxies and the mean galaxy mass increase slowly with the neighbour density. On a local scale the merged galaxy fraction is well described by (7).

3.3. Mass Functions

In Fig. 7 some final mass-functions are given. The mass functions in models 10a, b, c are very similar and can be approximated by a simple powerlaw $n(m)dm \propto m^{\alpha}dm$ with $\alpha \simeq -2.2$ as also found by JE. This value of α is found to be independent of galaxy radius in contrast to the work of AF who find a slope of $\alpha \simeq -3$ for small galaxy radius. However, when the loss of orbital energy in non-merger collisions is zero we also obtain a steeper final mass function approaching $\alpha = -3$.

The final mass function in model 11 illustrates the effect of mass-loss. The mean mass of unmerged particles does not decrease significantly during a run with mass-loss included, since most collisions lead to mergers. The final fraction of the total mass in background particles is about 1.5 (or $\sim \frac{1}{4} N_i/N_f$) when $K \simeq 1$ and somewhat lower when $K \simeq 0.5$. As noted in Sect. 2.3, our prescription for mass-loss should be regarded as an upper limit.

The mass function evolution was simulated with a Monte Carlo code under the assumption that the merging probability is prescribed by (4). All galaxies had equal mass initially ($N_i = 3000$). Results are given in Fig. 8. In model 10 we found $N_f/N_i \simeq 0.6$. The mass function obtained in the simulation when $N/N_i = 0.6$ has slope $\alpha = -2.2$ in good agreement with the final mass function in 10. We conclude that the general evolution of the mass function can be explained by (4).

Table 2

Model	N_{CC} :	N_{CM} :	N_{MM}	t [10^9 yr]	f_M	$\langle m\rangle_M$
10a exp	1 :	0.64±0.14 :	0.13±0.05	$1<t<3$	0.19	2.36
theor.	1 :	0.79 :	0.13			
10a exp.	1 :	3.5 ±0.8 :	0.9 ±0.3	$3<t<16$	0.29	3.02
theor.	1 :	1.65 :	0.51			
b exp	1 :	1.2 ±0.1 :	0.3 ±0.1	$3<t<16$	0.23	2.9
theor.	1 :	1.15 :	0.25			
c exp	1 :	2.4 ±0.4 :	0.5 ±0.2	$3<t<16$	0.20	2.92
theor.	1 :	0.98 :	0.19			

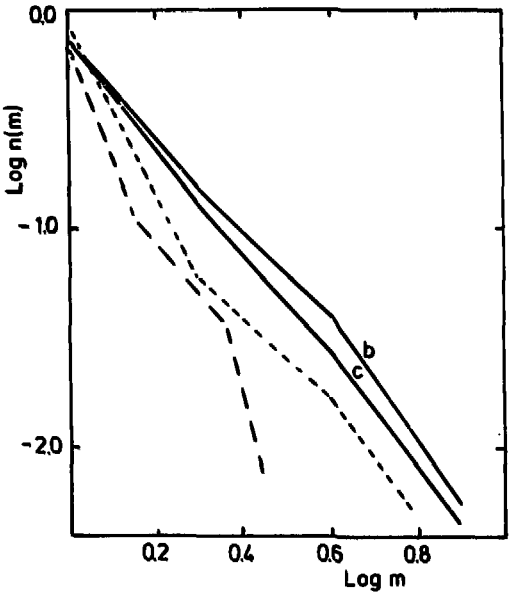


Fig. 7. Fraction of the total number of particles per unit interval of mass at t_f. Full lines are the results for model 10a and b. The dashed line is the final mass function in model 11, showing the effect of mass-loss. The dotted line shows the final mass function in model 10c when loss of energy in collisions not resulting in mergers is zero

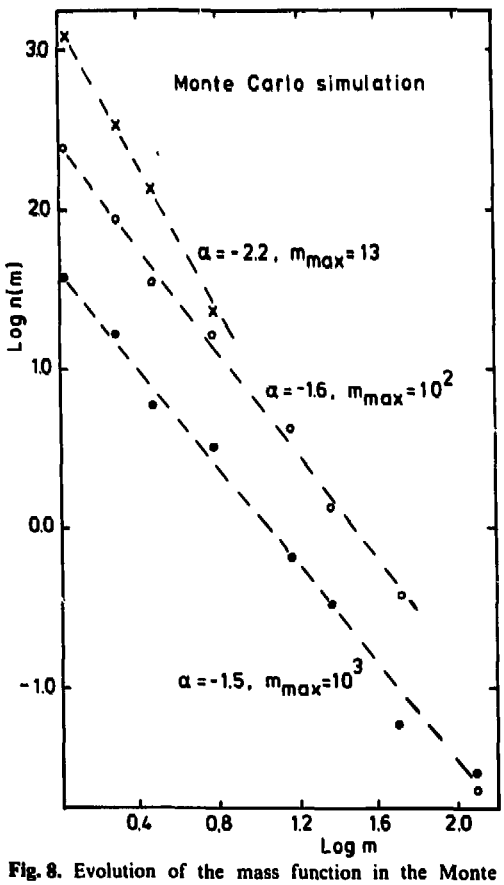


Fig. 8. Evolution of the mass function in the Monte Carlo simulation with 3000 equal mass particles initially. The dashed lines are powerlaw approximations to the mass functions for $N_i/N = 1.6$ (crosses), 5.5 (open circles) and 33 (filled circles)

When the simulations are continued the mass function reaches a self-similar form with $\alpha = -1.5$. This limiting form was also found in numerical experiments by Nakano (1966) and is in agreement with the self-similar solution of the coagulation equation when the cross-section of the particles is proportional to their mass (Trubnikov, 1971; Silk and White, 1978). Trubnikov's solution: $n(m)dm \propto m^{-1.5} e^{-m/m^*} dm$ is identical to the Press and Schechter (1971) mass function in the Poisson case. Therefore, Eq. (4) yields the proper limiting form of the mass function and we expect that the galaxy mass function evolves according to (4) as long as all galaxies in collapsed systems merge and no clusters form. It will describe the evolution of the cluster mass function or multiplicity function when merging ceases.

4. *Comparison with Observations*

The question we now turn to is whether one can identify systems that have undergone significant merging. The simulations give

Table 3

				$\bar{A}_{SS}$:	$\bar{A}_{SG}$:	$\bar{A}_{GG}$	k	λ
DG		Sample I	E	1	: 1.7 :	6.5	3.1	1.5
			S0	1	: 1.5 :	5.8	3.0	1.5
		Sample II	E	1	: 1.5 :	7.7	3.8	1.9
			S0	1	: 1.3 :	4.1	2.5	1.6
	f_G	f_S	G	N_{SS}:	N_{SG} :	N_{GG}		
KK	0.17	0.83	E+S0	1	: 0.48:	0.22	3.2	2.0
NI	0.10	0.55	E	1	: 0.5 :	0.16	3.5	2.5
NII	0.17	0.55	S0	1	: 0.47:	0.18	1.9	1.8

two clues from analysis of the fraction of merged galaxies and the clustering properties of merged galaxies. In fact we show that these properties support the hypothesis that elliptical and lenticular galaxies arise from merging spirals. We shall discuss later additional evidence from the intrinsic properties of early type galaxies in support of such a galaxy formation model.

4.1. *The Fraction of Early Type Galaxies*

The fraction of merged galaxies in the simulations is somewhat higher than the fraction of elliptical galaxies. However, the agreement is surprisingly good considering the uncertainties in our model and in the specific interpretation of the observations.

The observed fraction of elliptical galaxies among the bright galaxies in magnitude limited samples is about 0.1–0.15. The total fraction of elliptical and lenticular galaxies is about 0.2–0.3. These fractions depend on the luminosity function of galaxies of different types. There is observational evidence that the fraction of ellipticals among bright galaxies increases with luminosity, the mean absolute magnitude of ellipticals being brighter than spirals by about half a magnitude (Van den Bergh and McClure, 1979). This would make the real fraction of elliptical galaxies 0.05–0.08. The uncertainties in the simulations arise from: (i) the unrealistic initial mass function. (ii) the choice of the initial distribution of the particles. (iii) the merging process strongly inhibiting the formation of clusters in our simulations. This is most pronounced in model 10b although the galaxies in 10b have the realistic circular rotation velocity of ~ 250 km s^{-1}. It is possible that small visible cores of halos do not merge as easily as the halos they are embedded in. They might not merge at all when the galaxies are part of a collapsing cluster as in the White and Rees (1978) model.

4.2. *Clustering Properties of Early Type Galaxies*

We first note that the fraction of merged galaxies in the simulations is a very weak function of neighbour density (Fig. 4). Dressler (1979) has investigated the observed galaxy composition as a function of density in 55 clusters of galaxies. He finds that the fraction of E and SO galaxies varies slowly over a range of about three orders of magnitude in projected density. He also finds that at low densities the fraction of SO galaxies rises faster with density than the fraction of elliptical galaxies. In Sect. 5 we will see that these results fit naturally in the merger picture. The amount of clustering among galaxies in a magnitude limited survey depends on the depth of the catalogue, which is determined by the characteristic luminosity of the galaxies. The measured amount of clustering is sensitive to the bright end of the luminosity function (Peebles and Hauser, 1974). The luminosity functions of E, SO and spiral galaxies are very similar. However there is evidence that the average luminosity of ellipticals brighter than $M_v = -19$ is higher than disk galaxies by about a factor $\lambda \simeq 2$ (Van den Bergh and McClure, 1979). Therefore, in interpreting the observed correlation functions, we use the following crude rules: (1) The real clustering scales with the depth D of the catalogue as $A \propto D^{\gamma+2} \propto \lambda^{\gamma+2}$ or $A \propto \lambda$ if we use $\gamma = -2$. (2) The observations underestimate the cross-correlation between elliptical or SO and spiral galaxies by a factor $\lambda^{3/2}$ (or $\lambda^{-3/2}$ if $\lambda < 1$).

We first compare the simulation results with those of Davis and Geller (1976, DG) who study galaxy correlations as a function of morphological type. Specifically they use two samples (I and II) of galaxies from the Uppsala Catalogue to study the covariance functions of E, SO and spiral galaxies. Sample II is identical to sample I except that the galaxies in the Virgo and Coma clusters are excluded. Their results show that the correlation functions for E and SO galaxies are indeed higher and have a steeper slope as described in Sect. 3.2. The number of spiral neighbours within a distance r of an elliptical galaxy can be calculated from the amplitude A_{FS} and slope γ_{FS} of the cross correlation function for ellipticals and spirals by

$$n_{FS}(<r) = 4\pi \bar{n}_S A_{FS} (\gamma_{FS}+3)^{-1} r^{\gamma_{FS}+3}$$

where $\bar{n}_S$ is the mean number density of spiral galaxies. The number of galaxies on a local scale, say ~ 1 Mpc, is $n_{FS} = 4\pi \bar{n}_S \bar{A}_{FS}$, where $\bar{A}_{FS} = A_{FS}(\gamma_{FS}+3)^{-1}$. The relative numbers of close pairs of different types can be estimated from

$$N_{SS} : N_{FS} : N_{FF} = n_{SS} : n_{FS}\frac{\bar{n}_F}{\bar{n}_S} : n_{FF}\frac{\bar{n}_F}{\bar{n}_S} = \bar{A}_{SS} : 2\frac{\bar{n}_F}{\bar{n}_S}\bar{A}_{FS} : \left(\frac{\bar{n}_F}{\bar{n}_S}\right)^2 A_{FF}.$$

In our simulations Eq. (6) was valid as long as virialised clusters did not form. Galaxies have formed virialised clusters and we will assume that during the collapse of these clusters (6) was valid and a fraction of ellipticals given by (7) was formed in high density regions. In these virialised regions galaxies meet randomly. Thus, we assume that

$$N_{SS} : N_{FS} : N_{FF} = 1 : 2f_F'/f_S' : (f_F'/f_S')^2.$$

where

$$f_F'/f_S' = k f_F/f_S$$

and

$$k = \tfrac{1}{2}(\langle m\rangle_F/\langle m\rangle_S + 1). \qquad (9)$$

Therefore

$$\bar{A}_{SS} : \bar{A}_{FS} : \bar{A}_{FF} = 1 : k : k^2.$$

Correcting for a difference in mean luminosity between E and S galaxies as discussed previously, we find for the observed ratios ($\lambda > 1$)

$$\bar{A}_{SS} : \bar{A}_{FS} : \bar{A}_{FF} = 1 : k/\lambda^{3/2} : k^2/\lambda. \qquad (10)$$

In Table 3 the ratios for sample I and II of *DG* are given together with the values for k and λ derived from (10). The mass of E and SO galaxies derived in this way gives $\langle m\rangle_{E+SO}/\langle m\rangle_S \simeq 5$ which is slightly larger than the mass of merged galaxies in the simulations. The results in Table 3 further imply that E and SO galaxies have a larger mean brightness and a larger mass-to-light ratio than spiral galaxies.

We can also make a comparison with the work of Van Albada (1979) who has studied the small scale clustering of galaxies in the Uppsala Catalogue. The galaxies were divided into 10 intervals of 0.5 apparent magnitude. For each galaxy the 10 nearest neighbours were listed (in all 10 mag intervals separately). A local background density is estimated from the angular distance r to the 10th nearest neighbour which gives the expected value of the distance to the nearest neighbour, $\langle r_1 \rangle$, assuming a random distribution of galaxies. The observed distribution of $x = r/\langle r_1 \rangle$ can be compared with the Poisson distribution after normalisation of the two distributions for $x > 1$. In this way the excess of the galaxy density over the *local* density at $x < 1$ is found. For $x < 1$ and for an apparent magnitude difference between a galaxy and its neighbour less than 2.5 mag the excess densities P_E and P_{SO} of galaxies around E and SO galaxies are larger than around spiral galaxies. Van Albada's results yield

$$P_S : P_{SO} : P_E = 1 : 1.6 : 3.2$$

indicating that the clustering behaviour and hence probably also the mass of SO galaxies is intermediate between that of elliptical and spiral galaxies.

The fraction of ellipticals in groups with small crossing times (Field and Saslaw, 1971) is larger than the overall fraction. This also holds for galaxies in binaries. Karachentsev and Karachentseva (1974, KK) classify 17% of the galaxies in their catalogue as ellipticals (taking ellipticals and lenticulars together), while their sample of 1206 double galaxies contains 27% ellipticals. In Turner's sample of binaries (Turner, 1976a) there are 94 pairs containing only galaxies classified as pure E or pure S (Noerdlinger, 1979a, NI). The classification is taken from the Uppsala catalogue containing 10% ellipticals and 55% spirals. The subsample contains 21% ellipticals. Following the arguments leading to (10) the ratios of the numbers of SS, ES, and EE pairs are

$$N_{SS} : N_{SE} : N_{EE} = 1 : \frac{2k}{\lambda^{3/2}} f_E/f_S : \frac{k^2}{\lambda} (f_E/f_S)^2.$$

The observed ratios in the KK and NI sample and in a sample of binaries containing only SO galaxies (Noerdlinger, 1979b, NII) are given in Table 3. The second and third column contain the overall fraction of early-type and spiral galaxies. The values of k and λ derived from the observed ratios agree with those found from the covariance functions. We conclude that the covariance function data and the data on binaries are consistent with the interpretation of early-type galaxies as merger products.

Although the evidence presented here cannot be regarded as particular convincing proof, the data do suggest that the number of neighbours around a mass-concentration is indeed correlated with mass and luminosity. We shall pause briefly here to discuss some additional evidence.

As noted by GTA (1979) the forms of the observed 3- and 4-point correlation functions are consistent with the infall model. They also note that the observed ratio of the excess density of galaxies around Abell clusters over the excess density around galaxies (Seldner and Peebles, 1977) can be explained in this model.

Longair and Seldner (1979) have measured the cross-correlation function between the positions of 3CR-radiogalaxies (R) having redshifts $z < 0.1$ and positions of the Lick galaxy counts (G). They find $A_{RG} = 4$ à 5 A_{GG}. This is expected if the number of neighbours on a local scale is proportional to mass, since strong radiosources have $\langle M_V \rangle = -2.3$ which is about 1.5 mag brighter than normal galaxies in a magnitude limited survey (Sandage, 1972a; Auriemma et al., 1977).

When binaries are formed according to (4) they will preferentially contain massive galaxies. We expect that their mass function has a higher characteristic mass and flatter slope than the field mass function. Recently White and Valdès (1980) have studied the luminosity function of close binary galaxies. They find that the luminosity function of these galaxies has a higher characteristic luminosity but does not show a flatter slope at the faint end relative to Felten's field luminosity function (1977) as would be expected. In another recent study Heiligman and Turner (1980) find that the luminosity function of galaxies in compact groups is flatter than the field luminosity function. It seems that both studies agree qualitatively that in binaries as well as in compact groups massive galaxies are overrepresented relative to the field.

5. Evolution of Galaxy Mass-functions and the Origin of Morphological Types

Let us now investigate the evolution of a more realistic mass function due to merging or infall during the epoch of cluster formation. We note here that not only ellipticals, but also bulges of disk galaxies might be formed by merging. Consider a small galaxy falling in a larger one. It will spiral towards its centre losing orbital energy and angular momentum through dynamical friction. The orbital energy will be transformed partly into internal energy of the disk. The disk will be heated but may survive if the infallen mass is smaller than the disk mass, and the bulge will grow. We then expect that galaxies with large bulge-to-disk ratio have stellar disks that are hotter than galaxies with low B/D. Recent observational studies, indicate that SO galaxies have bulge-to-disk ratios (B/D) which are systematically larger than those of spiral galaxies (Dressler, 1979; Burstein, 1979a). These galaxies appear to have "thick disks" as well (Tsikoudi, 1977; Burstein, 1979b). The colours and gas content of SO's indicate that they form a class intermediate between spirals and ellipticals. Their clustering properties, B/D-ratio and disk structure now suggest that they are the results of mergers between galaxies of unequal mass. We attempt to discuss this quantitatively in the following model.

5.1. The Model

Let us speculate that all bright ($M_V \lesssim -16$) galaxies were initially formed as disk galaxies resembling late-type spirals. It is proposed that these galaxies evolve along the sequence (Sc–Sb–Sa)–SO–E as they merge with other (smaller) galaxies in the subsequent clustering process. A key parameter determining the galaxy's morphological type in such a model will be Δ, the ratio of the accreted mass to the original mass.

We have chosen an initial mass function which simulates the observed galaxy luminosity function (Schechter, 1976)

$$\begin{aligned} n(m/m^*)d(m/m^*) &= (m/m^*)^{-1.25} && 0.015 < m/m^* < 1 \\ &= (m/m^*)^{-2.75} && 1 < m/m^* \lesssim 3 \\ &= 0 && m/m^* > 3 \end{aligned}$$

where $n(m)$ is the fraction of galaxies in the interval $(m/m^*, m/m^* + d(m/m^*))$ and m^* is the mass of a galaxy with absolute magnitude $M_V = -21.5$. The evolution of this mass function was numerically calculated using the Monte-Carlo code discussed in Sect. 3.3. Since the merging probability is proportional to the total mass of the merging galaxies, the most massive galaxies will swallow

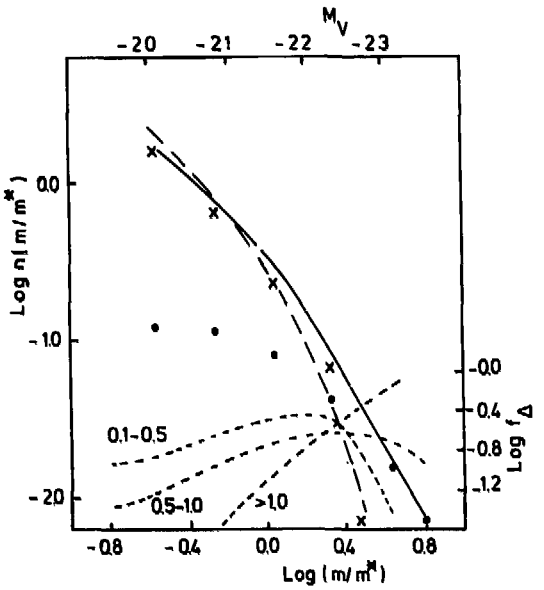


Fig. 9. Mass function for all particles (full line), for the particles with $\Delta < 0.5$ (crosses) and $\Delta > 0.5$ (dots) in the Monte Carlo simulation, when $\langle m\rangle/\langle m\rangle_i = 1.6$. The dashed line is the TYS luminosity function (see Sect. 5.3). The dotted lines give f_Δ, the fraction of galaxies having Δ within the range indicated near the curves

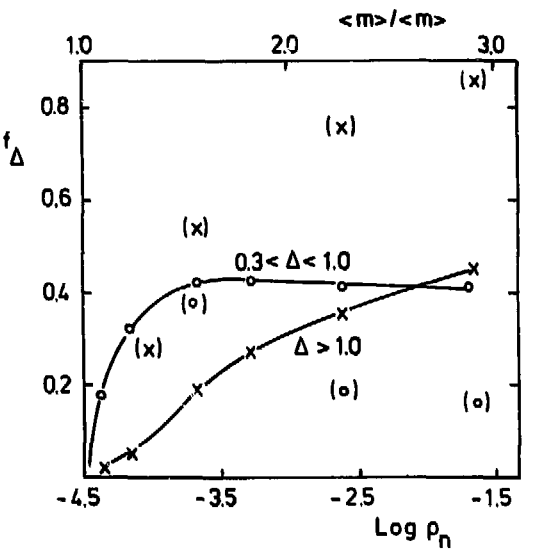


Fig. 10. Fraction of galaxies with $\Delta > 1$ (crosses) and $0.3 < \Delta < 1.0$ (open circles) plotted against $\langle m\rangle/\langle m\rangle_i$ and neighbourdensity ϱ_n. The fractions were calculated for $m > 0.4\ m^*$ (full lines) and for $m > m^*$ (between brackets)

Table 4

$\langle m\rangle/\langle m\rangle_i$ [$\langle m\rangle_i = 0.86\ m^*$]		$\Delta < 0.1$	$0.1 < \Delta < 0.5$	$0.5 < \Delta < 1.0$	$\Delta > 1.0$
1.1	f_Δ	0.81	0.11	0.06	0.02
	$\langle m\rangle_\Delta/m^*$	0.85	1.2	1.2	2.7
1.6	f_Δ	0.39	0.26 (0.18)	0.16 (0.20)	0.19 (0.54)
	$\langle m\rangle_\Delta/m^*$	0.75	1.13	1.2	2.9
2.0	f_Δ	0.28	0.24	0.18	0.3
	$\langle m\rangle_\Delta/m^*$	0.65	1.05	1.14	3.6

most mass. We shall study the mass function where the low-mass cut-off will have negligible effect, and we assume that at the present epoch $\langle m\rangle/\langle m\rangle_i \simeq 1.6$, as found in the N-body experiments. Here $\langle m\rangle_i$ is the mean initial mass for $m > 0.4\ m^*$.

5.2. Results

Results of the Monte Carlo simulations are given in Figs. 9 and 10 and in Table 4. Figure 9 gives the mass functions, when $\langle m\rangle/\langle m\rangle_i = 1.6$, for: (1) all galaxies, (2) galaxies having $\Delta > 0.5$, and (3) the galaxies having $\Delta < 0.5$. The mass-function for galaxies having $\Delta > 0.5$ is flatter than for other galaxies. In the same figure the fraction of galaxies f_Δ within a certain range of Δ, is plotted logarithmically against mass. Note that the slope of the mass function decreases with increasing Δ. Table 4 gives the fractions f_Δ and mean mass $\langle m\rangle_\Delta/m^*$ of the galaxies at three stages in the evolution of the mass function. All quantities were calculated for $m > 0.4\ m^*$. The fractions given in parentheses were calculated for $m > m^*$. It appears that f_Δ is quite sensitive to the choice of this lower limit. In Fig. 10 the relation between f_Δ and $\langle m\rangle_\Delta/\langle m\rangle_i$ is plotted for $0.3 < \Delta < 1.0$ and for $\Delta > 1.0$. We have used Eq. (8) to relate $\langle m\rangle/\langle m\rangle_i$ to the number density of galaxies ζ_n.

5.3. Comparison with Observations

Bulge-to-disk ratios are correlated with Hubble type (Roberts et al. 1975; Wakamatsu, 1976), but the range of bulge-to-disk masses that should be assigned to a Hubble type is uncertain.

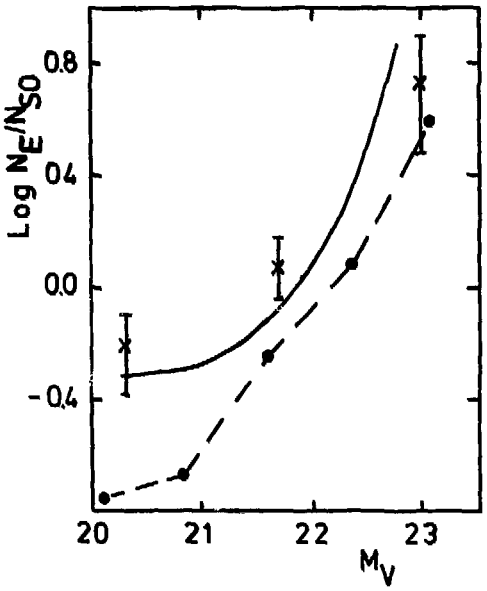


Fig. 11. Ratio of the number of ellipticals to the number of S0 galaxies from Van den Bergh and McClure (1979, crosses, the error bars are 1σ) and TYS (full line). The dashed line gives the result of the Monte Carlo simulation

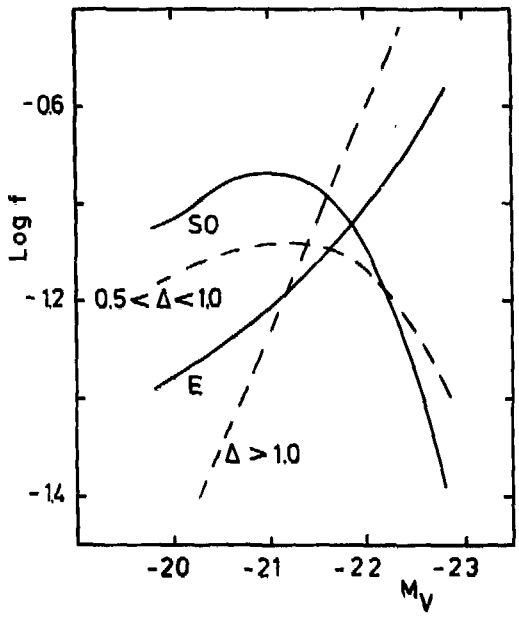


Fig. 12. The fraction of elliptical and S0 galaxies from TYS (full lines) and the results of the Monte Carlo simulation (dashed lines)

Burstein (1979a) has measured B/D values for three Sc and twelve SO galaxies. The B/D values of the spirals ranged from <0.07 to 0.2. Ten out of twelve SO galaxies had $0.5<B/D<1.5$. Quantitative estimates of B/D for different galaxy types are also given by Dressler (1979). These are 0.2, 0.2–0.5, and >0.5 for Sc, Sb, and Sa galaxies respectively. In regions of low galaxy density SO galaxies tend to have low B/D. There is typically a difference of about 1 mag between total and bulge magnitude for SO's in these regions suggesting that SO galaxies have $B/D\simeq 0.5$. The high B/D values for early type galaxies may partly be due to fading of the disk (Kent, 1980). Henceforth, we will assume that ellipticals have $\Delta>1.0$ and SO galaxies $0.5<\Delta<1.0$.

The mass function in Fig. 9 can be compared directly with luminosity functions of galaxies if M/L is taken to be constant (upper scale), the dots should be shifted to the left by $\simeq 0.75$ mag if $(M/L)_{E,SO}\simeq 2\,(M/L)_S$. The luminosity function for all galaxies (without correction for internal absorption) given by Tammann et al. (1979, TYS) normalised to our luminosity function is shown by the dashed line in Fig. 9. We have converted their magnitude scale to M_V using $M_V=M_B-0.8$. A shift in luminosity by ~ 0.75 mag for galaxies with $\Delta>0.5$ seems required to bring the functions in agreement. A more important result from Fig. 9 is that we expect the slope of the mass functions for different types to decrease with increasing B/D. We expect the fraction of ellipticals to increase almost linearly with mass.

Shapiro's (1971) data indicated that the luminosity function of ellipticals is indeed flatter than for other galaxies. The luminosity functions for spiral and SO galaxies presented by Dressler (1979) have a slope of about -2.5 for $M_V<-20.5$, while for ellipticals it is about -1.5. Van den Bergh and McClure (1979) have determined the luminosity function of E and SO galaxies from a sample studied by Sandage and Visvanathan (1978). The ratios of the number of E to the number of SO galaxies in this sample are given in Fig. 11. The ratios calculated from the TYS luminosity functions and the results of the Monte Carlo simulations are also shown. The slopes of the functions in the three cases are similar. Again, dimming of galaxies with $\Delta>0.5$ by about 0.75 mag would improve the fit.

In Fig. 12 the fractions of E and SO galaxies as a function of M_V from the TYS data are given. The dashed lines are the results from the Monte-Carlo simulation for $\Delta>1.0$ and $0.5<\Delta<1.0$. These fractions were calculated assuming $(M/L)_{\Delta<0.5}=2(M/L)_{\Delta<0.5}$. As expected a population of galaxies with small Δ has to be present before the population of galaxies with larger Δ can grow (Fig. 10). A similar behaviour for the frequencies of SO and E galaxies as a function of density was found by Dressler. His results apply to galaxies in rich clusters and are therefore not directly comparable to ours.

This preliminary exploration of a scenario in which bright galaxies are formed as late-type spirals which evolve along the sequence (Sc–Sa)–SO–E due to merging gives predictions that are already in qualitative agreement with observations in spite of all uncertainties discussed above.

6. Discussion

The general context in which we have placed the merging process in this paper is clearly different from the cannibalism process

described by Hausman and Ostriker (1978). If the galaxies are distributed at random initially, peaks in the density distribution, such as massive galaxies or proto-groups and clusters, will act as nuclei for the clustering process on local scales, the infall rates being proportional to mass. High-density peaks may exist in this initial distribution. Merging among the galaxies in these peaks may produce *cD*-galaxies which tend to gather clusters of galaxies around them afterwards via infall. In these clusters the massive galaxies lose orbital energy through dynamical friction and spiral towards the centre of the cluster where they can be cannibalised by a giant elliptical galaxy.

We have stressed in this paper how merging or infall might have played an important role in establishing the differences between the morphological types of galaxies. The model presented in Sect. 5 fits naturally in a scenario in which both galaxies and clusters form in a hierarchical clustering process such as the one discussed by White and Rees (1978). In their model, dark halos grow by gravitational clustering of objects formed shortly after recombination. Substructure is destroyed during the collapse of bound units. Any residual gas will, if ionised, be locked to the comoving frame by Compton drag on the microwave background. This would prevent it from falling into the potential wells of the halos until $z \simeq 50$ (Hogan, 1979). At some epoch ($z \gtrsim 10$) the gas will fall into the cores of the halos, cooling and forming rotating disks. The characteristic mass and radius of bright galaxies might have been determined by the criterium that cooling and collapse timescales are comparable (Silk, 1977; Rees and Ostriker, 1977). Note that at low luminosities ($M_v \sim -18$) most galaxies are gas-rich spiral (Sc) and irregular galaxies. Below $M_v = -16$ we find only irregular galaxies and dwarf ellipticals. Many observed properties of galaxies such as the color-magnitude relation, constant central surface brightness and metallicity gradients may have been established during the epoch when dissipation played an important role (Tinsley and Larson, 1978; Norman and Silk, 1980). The observation that the present abundances of morphological types and the luminosity function of galaxies is at most weakly dependent on the density of the surrounding medium, however, suggests that the formation of predominantly stellar disk galaxies was completed before clusters of galaxies started to form. We have assumed that the presently observed bright galaxies descend from this population of galaxies. Clusters of galaxies will now be able to form since the relatively small luminous disks do not merge as easily as their halos. Nevertheless a considerable fraction of them will have merged with other galaxies by the present time. It seems plausible that this leads to the formation of a bulge, heating and even destruction of the stellar disks and depletion of its neutral gas content. These processes will occur preferentially at the high mass-end of the mass function and in regions of high galaxy density.

The product of a merger between two equal mass spiral galaxies is more massive but not necessarily more luminous if star formation ceases after the merging process is completed. Young stars in the arms of spiral galaxies contribute roughly 25% to the total light (in the *U*-band) of a spiral galaxy (Schweizer, 1976; Bash, F. private communication). This contribution disappears within $\sim 10^8$ yr. The subsequent decrease in luminosity is of order 0.5 mag in 10^{10} yr. (Sandage, 1972b; Tinsley, 1972). So after $\sim 10^{10}$ yr the luminosity of such a merged galaxy might be comparable to the luminosity of one spiral while its mass is twice as large. According to Turner (1976b) and to the results of this paper (Sects. 4.2 and 5) the mass-to-light ratio of ellipticals is indeed about twice as large as for spirals. Although no calculations of mergers of rotating stellar disks have been performed, there are indications that merger products of galaxies of comparable mass tend to have a Hubble density profile (White, 1979; Villumsen, 1979) and rotate slowly (JE, AF) as do elliptical galaxies. Recent observations indicate that bulges rotate faster than ellipticals (Illingworth et al., 1980; Kormendy and Illingworth, 1980). It will be interesting to see whether this is also consistent with the origin of bulges as proposed here. At the moment it seems that there is much evidence speaking in favor of a model in which merging is responsible of an evolution of bright galaxies along the sequence (Sc–Sb–Sa) –SO–E.

Conclusions

The numerical *N*-body simulations given here model the situation where clusteres of galaxies form from small random fluctuations in the initial distribution of galaxies corresponding to low initial velocity dispersion. The simulation particles, which all have equal mass initially are supposed to represent the halos of typical bright galaxies with $L = L^*$, and total mass-to-light ratio of about 150. The principal conclusions are as follows.

1. If the starting time of the simulations is chosen to produce about the observed amplitude of the covariance function, the final fraction of merged galaxies is typically 30%, varying from $\simeq 15\%$ in low density regions to $\simeq 50\%$ in high density regions. Since most mergers occur between glaxies in bound pairs, this fraction is very insensitive to galactic parameters and not particularly sensitive to either the final density or the final amplitude of the covariance function.

2. The galaxy collapse time estimated from the initial redshift in our simulations required to produce the observed amplitude of the covariance function is $\simeq 10^9$ yr for $\Omega \simeq 1$ and $\simeq 10^8$ yr for $\Omega = 1$.

3. The clustering properties and merging probability of the particles in the simulations are consistent with Eq. (4) as long as merging prevents the formation of clusters. When clusters form, the merging rate is depressed most strongly for the smallest particles.

4. The mass function is well described by a power-law approaching a self-similar form identical to the Press and Schechter mass function for an initially Poissonian distribution.

5. Observational data on the clustering properties of early-type galaxies are in good agreement with their interpretation as merged galaxies. Adopting this interpretation we find that the characteristic mass of ellipticals is several times larger than spirals while their luminosity is about a factor two larger. The clustering properties of SO galaxies suggest that their mass and luminosity is intermediate between that of spirals and ellipticals.

6. The form of the three and four-point correlation functions, the clustering of 3C-radiogalaxies and the luminosity functions of galaxies in binaries and compact groups indicate that galaxy mass is indeed correlated with local galaxy density in our universe.

7. It is suggested that galaxies were formed as mainly stellar disks at $z \gtrsim 10$ and their luminous spherical components form by merging or infall during the subsequent epoch of cluster formation resulting in an evolution of galaxies along the sequence (Sc–Sb–Sa)–SO–E. This model predicts that the mass function of galaxies flattens from late to early-type galaxies. At the bright end of the luminosity function we expect to find mainly early-type galaxies. It is possible to explain qualitatively the observed frequencies of morphological types as a function of mass and neighbour density in this type of model.

Acknowledgements. I am very grateful to Sverre Aarseth for making his code available to me and for his friendly and helpful advise on how to incorporate collision effects. I wish to thank also Colin Norman for many useful discussions and thorough reading of the manuscript. Michael Fall and Simon White are acknowledged for many critical and useful comments on an earlier version of the manuscript. Tjeerd van Albada kindly provided his analysis of galaxy clustering prior to publication.

References

Aarseth, S.J., Gott, J.R., Turner, E.L.: 1979, *Astrophys. J.* **228**, 664
Aarseth, S.J., Fall, S.M.: 1979, *Astrophys. J.* **236**, 43
Auriemma, C., Perola, G.C., Ekers, R., Fanti, R., Lari, C., Jaffe, W., Ulrich, M.H.: 1977, *Astron. Astrophys.* **57**, 41
Burstein, D.: 1979a, *Astrophys. J.* **234**, 435
Burstein, D.: 1979b, *Astrophys. J.* **234**, 829
Chandrasekhar, S.: 1943, *Rev. Model. Phys.* **15**, 2
Davis, M., Geller, M.J.: 1976, *Astrophys. J.* **208**, 13
Dressler, A.: 1980, *Astrophys. J.* **236**, 351
Felten, J.E.: 1977, *Astron. J.* **82**, 861
Field, G.B., Saslaw, W.C.: 1971, *Astrophys. J.* **170**, 199
Gott, J.R., Rees, M.J.: 1975, *Astron. Astrophys.* **45**, 365
Gott, J.R., Turner, E.L.: 1976, *Astrophys. J.* **209**, 1
Gott, J.R., Turner, E.L., Aarseth, S.J.: 1980, *Astrophys. J.* **234**, 13
Gunn, J.E., Gott, J.R.: 1972, *Astrophys. J.* **176**, 1
Illingworth, G., Schechter, P.L., Gunn, J.E.: 1980 (in preparation)
Hausmann, M., Ostriker, J.P.: 1978, *Astrophys. J.* **224**, 320
Heiligman, G.M., Turner, E.L.: 1980, *Astrophys. J.* **236**, 745
Hogan, C.J.: 1979, *Monthly Notices Roy. Astron. Soc.* **188**, 781
Jones, B.J.T., Efstathiou, G.: 1979, *Monthly Notices Roy. Astron. Soc.* **189**, 27
Karachentsev, D., Karachentseva, V.E.: 1974, *Astron. Zh.* **51**, 724 (English transl. in *Soviet Astron. Astrophys. J.* **18**, 428
Kormendy, J., Illingworth, G.: 1980 (in preparation)
Longair, M.S., Seldner, M.: 1979, *Monthly Notices Roy. Astron. Soc.* **189**, 433
Miyoshi, K., Kihara, T.: 1975, *Publ. Astron. Soc. Japan* **27**, 333
Nakano, T.: 1966, *Progr. Theor. Phys.* **36**, 515
Noerdlinger, P.: 1979, *Astrophys. J.* **229**, 470
Noerdlinger, P.: 1979, *Astrophys. J.* **229**, 877
Peebles, P.J.B., Hauser, M.G.: 1974, *Astrophys. J. Suppl.* No. 253, **28**, 19
Press, W., Schechter, P.: 1974, *Astrophys. J.* **187**, 425
Rees, M.J., Ostriker, J.P.: 1977, *Monthly Notices Roy. Astron. Soc.* **179**, 541
Roberts, W.W., Roberts, M.S., Shu, F.H.: 1975, *Astrophys. J.* **196**, 381
Roos, N., Norman, C.A.: 1979, *Astron. Astrophys.* **76**, 75
Sandage, A.: 1972a, *Astrophys. J.* **178**, 1
Sandage, A.: 1972b, *Astrophys. J.* **178**, 25
Sandage, A., Visvanathan, N.: 1978, *Astrophys. J.* **228**, 742
Schechter, P.: 1976, *Astrophys. J.* **203**, 297
Seldner, M., Peebles, P.J.B.: 1977, *Astrophys. J.* **215**, 703
Shapiro, S.L.: 1971, *Astron. J.* **76**, 291
Silk, J.: 1977, *Astrophys. J.* **211**, 638
Silk, J., White, S.D.: 1978, *Astrophys. J.* **223**, L 59
Tammann, G.A., Yahil, A., Sandage, A.: 1979, *Astrophys. J.* **234**, 775
Tinsley, B.M.: 1972, *Astrophys. J.* **178**, 319
Tinsley, B.M., Larson, R.B.: 1979, *Monthly Notices Roy. Astron. Soc.* **186**, 503
Toomre, A., Toomre, J.: 1972, *Astrophys. J.* **178**, 662
Toomre, A.: 1977, in *The Evolution of Galaxies and Stellar Populations*, eds. B. M. Tinsley and R. B. Larson, New Haven: Yale University Observatory, p. 401
Trubnikov, B.A.: 1971, *Soviet. Phys. Doklay* **16**, 124
Tsikoudi, V.: 1979, *Astrophys. J.* **234**, 842
Turner, E.L.: 1976a, *Astrophys. J.* **208**, 20
Turner, E.L.: 1976b, *Astrophys. J.* **208**, 304
Turner, E.L., Gott, J.R.: 1976, *Astrophys. J.* **209**, 6
Turner, E.L., Aarseth, S.J., Gott, J.R., Blanchard, N.T., Matthieu, R.: 1979, *Astrophys. J.* **228**, 684
Van Albada, T.S.: 1980 (in preparation)
Van Albada, T.S., Van Gorkom, J.N.: 1977, *Astron. Astrophys.* **54**, 121
Van den Berg, S., McClure, R.D.: 1979, *Astrophys. J.* **231**, 671
Villumsen, J.V.: 1979 (preprint)
Wakamatsu, K.: 1976, *Publ. Astron. Soc. Japan* **28**, 397
White, S.D.M., Rees, M.J.: 1978, *Monthly Notices Roy. Astron. Soc.* **183**, 341
White, S.D.M.: 1978, *Monthly Notices Roy. Astron. Soc.* **184**, 185
White, S.D.M.: 1979, *Monthly Notices Roy. Astron. Soc.* **189**, 831
White, S.D.M., Valdès, F.: 1980, *Monthly Notices Roy. Astron. Soc.* **190**, 55

CHAPTER IV

EVOLUTION OF RICH CLUSTERS OF GALAXIES

Abstract.

The evolution of a rich cluster of galaxies having a realistic initial distribution of masses is numerically simulated. The sizes and masses of the simulation particles are such that they represent halos presumed to be attached to galaxies initially. The ratio of the amount of dark to luminous material is assumed constant. The upper end of the mass spectrum is most strongly effected by the merging process occurring predominantly during the expansion and subsequent collapse of the cluster. The amount of mass segregation in the collapsed cluster is shown to be consistent with the observed luminosity segregation in the Coma cluster. In a previous paper it was proposed to interprete the increase of Δ, the ratio of the accreted over the initial mass of galaxies, as an evolution of galaxies along the sequence (Sc-Sb-Sa)-SO-E. The results presented here show that such a model is indeed consistent with recent observations such as the over-all abundance of different morphological types, the local density-morphological type relation and the mass function of elliptical galaxies. The amount of mass loss from galaxies in high velocity encounters after the collapse of the cluster is estimated. The results are consistent with the observational evidence that elliptical galaxies in spiral-poor and cD clusters have smaller sizes than in spiral-rich clusters.

1. Introduction.

The scenario in which the clustering of galaxies is the result of a hierarchical clustering process during the expansion of the universe is consistent with current observations (e.g. Peebles 1980). In such a picture galaxies are formed at a redshift $z \gtrsim 10$. Clusters form later from the density fluctuations in the distribution of galaxies. During the expansion and subsequent collapse of the protocluster, smaller subclusters, which generally have higher initial amplitudes, will already

form inside the protocluster (White, 1976) and the formation of the cluster may be regarded as a hierarchical merging process of subclusters. Violent relaxation will bring the cluster into virial equilibrium within about one dynamical time. These later dynamical stages are best described in numerical N-body simulations (Peebles, 1970; Aarseth and Hills, 1972 and White, 1976) which can incorporate physical processes such as subclustering and violent relaxation. The density profiles of the simulated clusters show good agreement with observed profiles, although strong mass segregation was found in White's model in which the particles were assigned masses according to a Schechter function. He concluded from the absence of luminosity segregation observed in Coma that only a small fraction ($\sim$ 20%) of the total cluster mass could have been attached to the galaxies initially. The observational situation regarding the luminosity segregation in Coma has altered due to the work of Quintana (1979) and of Capelato et al. (1980), who show that there is a definite segregation of luminosity in this cluster. This seems in agreement with two other recent observational results: (i) the existence of a clear correlation between the local density of galaxies and the abundance of morphological types (Dressler, 1980; Gisler, 1980) and (ii) a difference between the luminosity function of ellipticals and the luminosity function of other galaxies in the sense that the fraction of ellipticals increases with luminosity (Van den Bergh and McClure, 1979; Dressler, 1980; Tammann, Yahil and Sandage, 1979; Thompson and Gregory, 1980). It seems that the observed segregations of mass and of morphological type are readily explained in a model in which early-type galaxies are identified with the products of mergers between galaxies of comparable mass (Toomre and Toomre, 1972; Toomre, 1977; Roos and Norman, 1979, Paper I). This was confirmed in subsequent cosmological simulations (Jones and Efstathiou, 1979; Aarseth and Fall, 1980; Roos, 1981, Paper II). As discussed in Paper II these simulations strongly indicate that the observed correlation of density with mass and morphological type is primarily the result of infall of small galaxies onto larger ones during the early stages of the formation of clusters. The infall rate, and thus the number density of neighbours around a galaxy was found to be roughly proportional to galaxy mass, in agreement with the infall model by Gunn and Gott (1972).

In previous N-body simulations in which inelasticity effects are

included, particles were given equal masses initially. Furthermore, the number of particles in the simulations was too low (400-1000) to allow rich clusters ($N \gtrsim 100$) to form. It seemed therefore worthwhile to simulate the evolution of a rich cluster of galaxies starting with a realistic initial mass spectrum and to compare the results with the most recent available observational data. As a key parameter determining the morphological type of a galaxy we will adopt Δ, the ratio of the infallen mass in the form of smaller galaxies over the original mass (see Paper II).

In section 2 the initial conditions and results of the numerical experiments are presented. Also presented are some comparison experiments, such as an experiment with a slightly larger initial expansion velocity to simulate the expanding universe in which the cluster may form. In section 3 the numerical results are compared with observations and in section 4 the evolution of a rich cluster after its collapse and virialisation is discussed. The results are summarized in section 5.

2. Numerical experiments.

2.1 Initial conditions.

(i) Mass function.

The luminosity function of galaxies in and outside clusters is well described by a function of the form (Schechter, 1976)

$$n(L)\, dL = n^* \, (L/L^*)^{\alpha} \, e^{-L/L^*} \, dL \qquad (1)$$

where $n(L)\, dL$ is the number of particles having luminosity in the interval $[L, L+dL]$ and n^* is determined by the total number of galaxies per unit volume. The characteristic luminosity L^* is about $4\ 10^{10}\ L_\odot$ (in the visual) and $\alpha \simeq -1.25$. The form of this function is identical to the mass function derived by Press and Schechter (1974) for self similar clustering (see also Efstathiou, Fall and Hogan, 1979). In the case of a Poissonian initial fluctuation spectrum they find $\alpha = -3/2$. In Paper II it was found empirically that, for Poissonian initial conditions, the merging probability for a pair of galaxies was proportional to their combined mass . It was pointed out that this would lead to exactly the

same mass function with $\alpha = -3/2$. Therefore we have adopted the following initial mass function

$$n(m)\ dm = \begin{cases} n^* \ (m/m^*)^{-1.5}\ e^{-m/m^*}\ dm, & m > m_{min} \\ 0 \qquad\qquad\qquad\qquad\quad , & m < m_{min} \end{cases} .$$

For the characteristic mass of the particles m^* we have chosen $m^* = 6 \times 10^{12}\ M_\odot$ yielding a mass-to-(visual) light ratio of $150 M_\odot/L_\odot$ and the simulation particles are to be identified with galactic halos. The minimum mass of the particles was $m_{min} = 0.1\ m^*$ yielding a mean initial mass per particle of $\langle m \rangle_i = 0.35\ m^*$. The total number of particles required to simulate a rich cluster having typical mass $10^{15}\ M_\odot$ is then about 500. Without the cut-off of the mass spectrum the mass below m_{min} would account for 28% of the total mass. The contribution of this mass to the mass accreted by galaxies having a mass m is about a fraction $M(< m_{min})/M(< m)$ of the total mass accreted, where $M(< m)$ is the total mass of galaxies having mass smaller than m. This is about 60% at $0.4\ m^*$, 40% at m^* and 30% at $2.5\ m^*$. Although the adopted cut-off will thus have a significant effect on the evolution of the mass spectrum (Section 2.2.2) we have not lowered it in order to keep the number of particles in the simulation and hence the computation time within reasonable limits.

(ii) Collision cross-sections.

The radius of the particles having mass m^* is 150 kpc yielding a (one-dimensional) internal velocity dispersion of these simulated halo galaxies of about 200 km sec^{-1}. The radius of the particles scales with mass as $R_g \propto m^{\frac{1}{2}}$. Assuming constant mass-to-light ratio this is consistent with the observation that the (central) surface brightness of (elliptical) galaxies is independent of luminosity (Faber and Jackson, 1976). The inelasticity of the encounters among the particles depends on their assumed internal velocity dispersion, which should scale as $m^{1/4}$. We have ignored this and adopt for all particles an internal velocity dispersion of 200 km sec^{-1}. The loss of orbital energy in collisions and the merging criteria were the same as in Paper II. Mass loss was ignored. The gravitational potential for the mutual interaction of two galaxies was calculated from

$$\phi_{ij} = -G\ m_i m_j\ (r_{ij}^2 + \varepsilon^2)^{-\frac{1}{2}},$$

where G is the gravitational constant, r_{ij} is the distance between the galaxies and ε, the usual softness parameter, is equal to 0.6 times the radius of the largest galaxy (see Paper II).

(iii) Cluster parameters.

Initially the particles are distributed at random in a spherical volume of radius R_i (subindices i and f denote initial and final respectively). From the amount of clustering produced in expanding universe simulations it was found that galaxies are formed within $\sim 10^9$ yr after the Big Bang. This will be approximately the starting time of the cluster simulations. The general Hubble expansion velocity is then about $H_i \approx 10^3$ km sec^{-1}Mpc^{-1}. The collapse time of the Coma cluster which has a radius $R_{cl} \approx 2.8$ Mpc and mass $M_{cl} \approx 2 \times 10^{15}\ M_\odot$ is

$$T_c \approx 5\ 10^9 \left[\frac{R_{cl}}{2.8\ \text{Mpc}}\right]^{3/2} \left[\frac{M_{cl}}{2\times10^{15}\ M_\odot}\right]^{-3/2} \text{yr}$$

(Gunn and Gott, 1972). In terms of the initial parameters of the protocluster it is

$$T_c \approx \frac{\pi}{H_i}\ (\Omega_i - 1)^{-3/2}$$

for $\Omega_i \approx 1$, where Ω_i is the initial ratio of the density over the critical density. It follows that $\Omega_i \approx 1.7$ to obtain a collapse time of 5×10^9 yr. The initial radius of the expanding sphere is then about 2 Mpc for a cluster of mass $1.5\times10^{15}\ M_\odot$ ($N_i = 700$). This yields a reasonable value of 2.3 R_g for the mean interparticle distance in the initial distribution. The radius of the sphere containing all the particles at maximum expansion can be estimated from

$$R_{max}/R_i = (1 - \Omega_i^{-1})^{-1}$$

and the radius of the virialised cluster will be $\sim\frac{1}{2}R_{max} = 2.5$ Mpc. The models were allowed to evolve for a total time of 1.5×10^{10} yr.

(iv) Computational technique.

The numerical code used to integrate the equations of motion was identical to the version of the Aarseth-Ahmad-Cohen code used in Paper I.

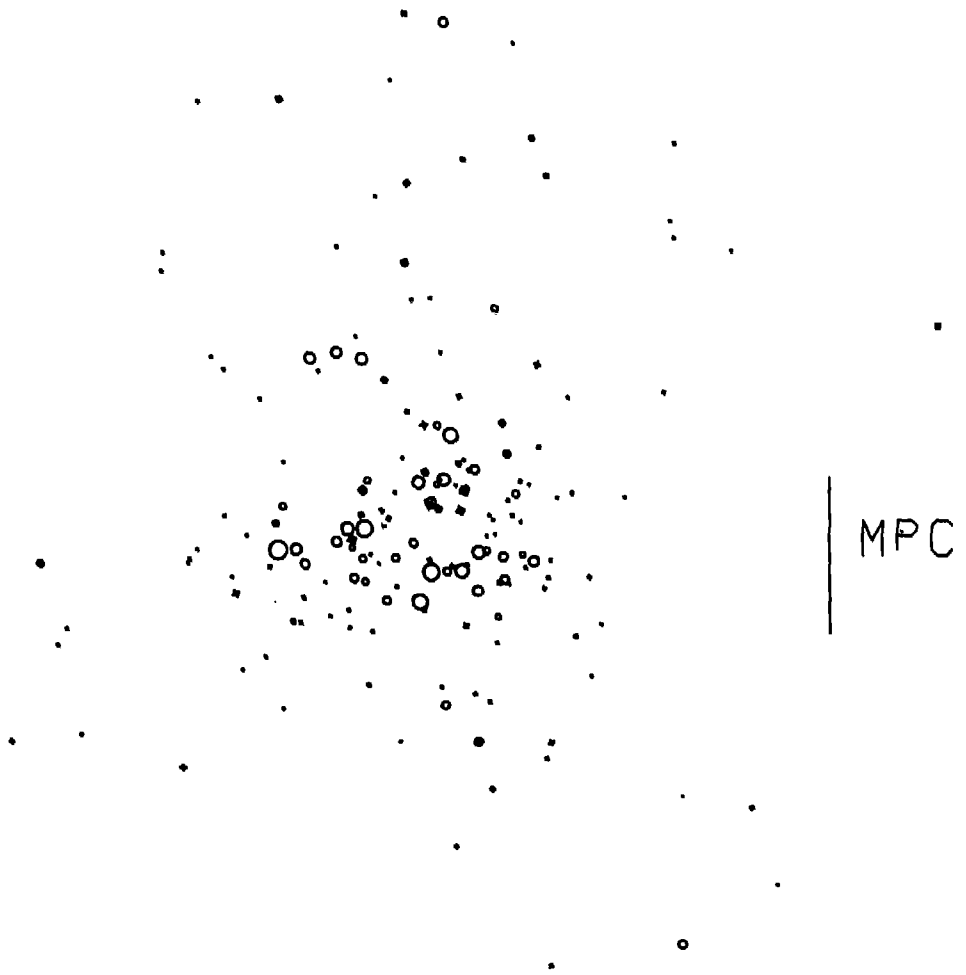


Figure 1. Projected distribution of particles at the end of the simulation in Model 1. The sizes of the particles as shown are smaller by a factor 7 than in the Model. Crosses correspond to the initial mass of the particles and circles to the infallen mass. The crosses are omitted for $\Delta > 1$.

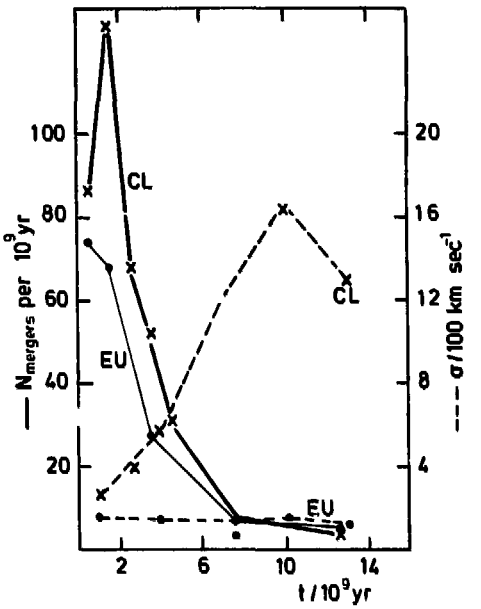


Figure 2. Number of mergers per 10^9 yr and dispersion in random velocities in Models 1 (Cl) and 5 (EU) as a function of time. Error estimates can be made from Table 2.

Table 1. Model Parameters.

		N_i	H_i [100 km sec^{-1} Mpc^{-1}]	Ω_i	N_f	σ_f [100 km sec^{-1}]
Clusters	1	700	12.0	1.75	240	13
	2	350	42.	1.05	132	2.7
	3	350	36.2	1.4	98	9
	4	350	18.2	5.6	78	11
Exp. Un.	5	700	19.4	0.66	410	1.2

The computations were carried out on an Amdahl V7B computer.

2.2 Results.

Besides the standard cluster model (Model 1) we have calculated some other models. We also did an expanding universe calculation (Model 5) with $\Omega_f = 0.07$ and $H_f = 60$ km sec^{-1} Mpc^{-1}. At the starting time this

Table 2. Merging Probabilities for Different Mass Groups.

Model 1 (C1)

$t/10^{10}$ yr	P_{m_1}	P_{m_2}	P_{m_3}
0.0-0.2	$(1.4\pm0.2)10^{-1}$	$(3.6\pm0.4)10^{-1}$	$(9.7\pm1)\ 10^{-1}$
0.2-0.5	$(4.2\pm1)\ 10^{-2}$	$(2.5\pm0.3)10^{-1}$	$(3.8\pm0.4)10^{-1}$
0.5-1.0	$(2\ \pm2)\ 10^{-3}$	$(4\ \pm1.5)10^{-2}$	$(8.5\pm1.5)10^{-2}$
1.0-1.5	$(2.5\pm2)\ 10^{-3}$	$(1\ \pm1)\ 10^{-2}$	$(4.5\pm1)\ 10^{-2}$

Model 5 (EU)

$t/10^{10}$ yr	P_{m_1}	P_{m_2}	P_{m_3}
0.0-0.2	$(0.9\pm0.1)10^{-1}$	$(2.1\pm0.3)10^{-1}$	$(5.1\pm0.1)10^{-1}$
0.2-0.5	$(1.5\pm0.5)10^{-2}$	$(1.2\pm0.2)10^{-1}$	$(1.9\pm0.3)10^{-1}$
0.5-1.0	$(3.2\pm1.5)10^{-3}$	$(3.2\pm0.1)10^{-2}$	$(5.8\pm1.4)10^{-2}$
1.0-1.5	$(2.4\pm2)\ 10^{-3}$	$(1.9\pm1)\ 10^{-2}$	$(4.4\pm1)\ 10^{-2}$

model differs only from the cluster model in the initial expansion velocity. Parameters for the models are given in Table 1. Column 1 gives the model number. The number of particles, expansion velocity and density parameter at the starting time are given in columns (2), (3) and (4) respectively. The final number of particles and (3-dimensional) dispersion in random velocities are given in columns (5) and (6). In Figure 1 the projected distribution of particles in the standard cluster at the end of the simulation is given. The final radial distribution of mass in cluster simulations has been discussed by Peebles (1970) and by White (1976). The distribution in our final model 1 is similar to that found by these authors.

(i) Merging rate.

In Figure 2 the total number of mergers per 10^9 yr as a function of time is given for the standard cluster model (C1) and for the expanding universe (EU) model. In the same figure the dispersion in random velocities in both models is given. The velocity dispersion in the cluster model is steadily increasing during the expansion and subsequent

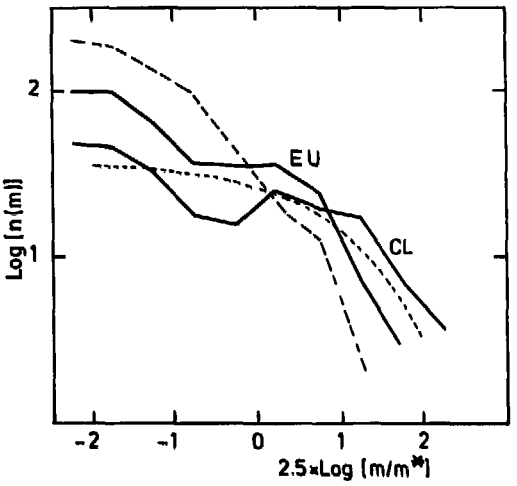


Figure 3. Final mass functions in Model 1 and 5. Here n(m) is the number of particles per interval Δ log m = 0.2. The dashed line is the initial mass function. The curve of short dashes is the mass function of the form given by equation (1) with α = -1 and a break at 2.5 m , normalized to yield the same total number of particles as in the final Model 1.

collapse due to the development of subclustering (cf. White, 1976). Although the collision rate in the cluster increases during that stage, the merging rate decreases because the velocity dispersion becomes larger than σ_g. As a result the final merging rates in the two models in Figure 2, which differ widely in mean density, are remarkably similar.

Merging probabilities as a function of time for galaxies in three mass groups are given in Table 2. The probabilities P_{m_1}, P_{m_2} and P_{m_3} are defined as the total number of mergers per 10^9 yr among galaxies, with the mass of the most massive one lying in the range $0.16 < m_1/m^* < 0.4$, $0.4 < m_2/m^* < 1$ and $m_3/m^* > 1$ respectively, divided by the total number of galaxies in that mass range. The errors in Table 2 represent standard deviations from the mean. For a random combination of pairs the ratio $P_{m_1} : P_{m_2} : P_{m_3}$ would be equal to $N(< m_1) : N(< m_2) : N(< m_3)$, where $N(< m_1)$ is the number of particles having mass $< m_1$. The latter ratio is about 0.7 : 0.9 : 1 which is clearly different from the ratios given in Table 2. The probability is about proportional with mass. It was found in paper II that this is not due to the larger size of the more massive galaxies and that the infall model provided a good

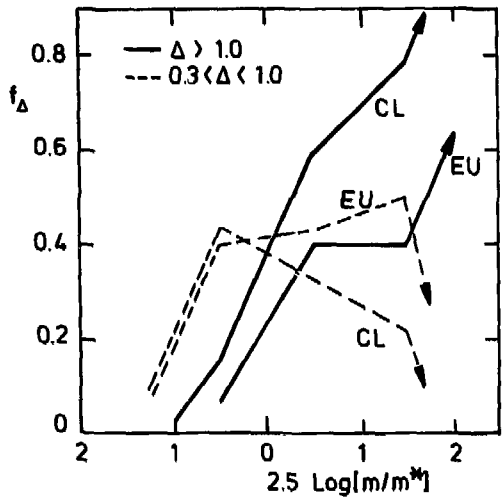


Figure 4. Fraction of particles having Δ in the indicated range as a function of mass in the final states of Models 1 and 5. Error estimates can be made from Table 3 and Figure 3.

explanation of this behaviour. It is surprising that this relation seems to hold even at late epochs in the expanding universe model as well as in the cluster model.

(ii) Mass functions.

In Figure 3 the initial mass function and the final mass functions for Models 1 and 5 are shown. The flatness of the final mass functions at the lower mass end of the spectrum and the depression at $m/m^* \approx 0.6$ are probably due to the low mass cut-off. Figure 4 shows the fractions of galaxies for different ranges of Δ, the ratio of the amount of accreted mass over the original mass of the particles. As can be seen from Figure 4 the depression in the mass functions in Figure 3 lies just in the region where a considerable fraction of the galaxies has accreted an amount of mass comparable to their own mass thereby moving to higher total mass. Although at small masses the amount of accreted mass is underestimated due to the cut-off of the mass spectrum, the increase in the fraction of particles having large with mass is mainly due to the infall rate being proportional to mass.

In Figure 5 the final mass functions of some clusters having differ-

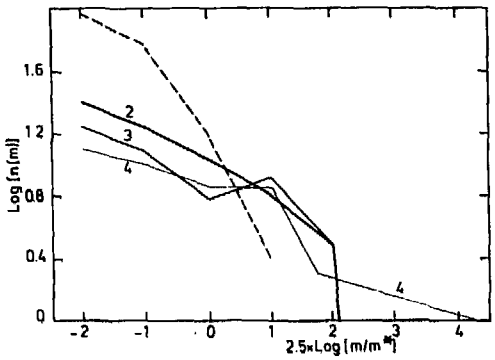


Figure 5. Final mass functions in Models 2, 3 and 4.

Table 3. Mass Segregation at the end of the computation for Model 1.

$\langle m \rangle / m^*$	N	$\langle BE \rangle / \langle BE \rangle_{all}$	$\langle R \rangle^{\dagger}$	$\langle R_{proj} \rangle$	$\sigma_{3d}^{\dagger}$	σ_{1d}
0.4	96	0.86	3.9	3.2	13.5	7.3
1.0	52	0.85	3.8	2.9	14.1	10.2
2.8	41	1.05	2.9	2.0	13.5	8.8
6.2	39	1.2	2.9	2.2	12.8	8.8
17.0	12	1.9	1.2	0.9	10.8	6.8

† Radii are in Mpc's, velocities in 100 km sec^{-1}.

ent collapse times are given (see Table 1). After the collapse of Model 4, which has a very short collapse time, a central massive galaxy grows by merging with other galaxies. The mass of the cannibal is 40 m^* at the end of the simulation.

(iii) Mass segregation.

The amount of mass segregation in the final cluster is quite strong, a result also found by White (1976). In Table 3 some quantities for the different mass groups in the final cluster model are given. The mean mass and total number of particles per mass group are given in columns (1) and (2) respectively. The ratio of the binding energy per unit mass in these mass groups over the mean binding energy of all particles is given in column 3. Columns 4, 5, 6 and 7 contain the mean radius, mean projected radius, 3 dimensional velocity dispersion and line-of-sight

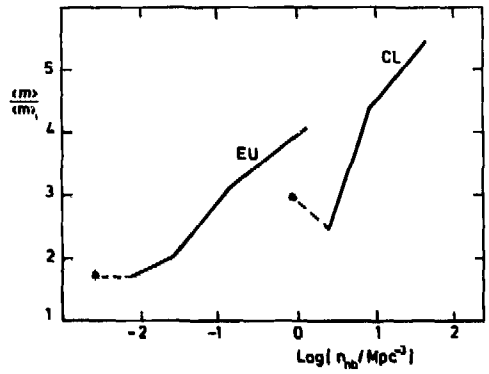


Figure 6. Ratio of final over initial mean mass of the particles in Model 1 and 5 as a function of number density of neighbours as defined in the text. Stars (*) give the over-all mean mass and number density of the particles in the Models.

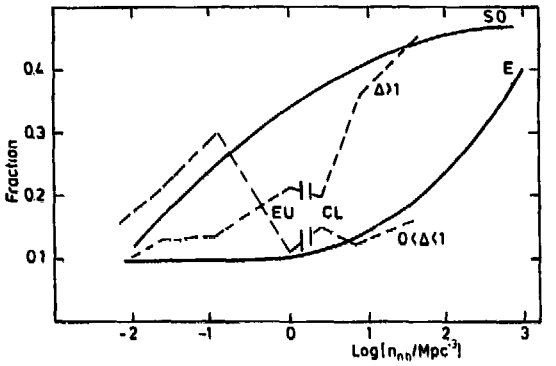
velocity dispersion respectively. The most massive particles are found closer to the cluster centre and have a higher mean binding energy per unit mass than the smaller ones. The difference in binding energy between the most massive and the lighter particles seems somewhat smaller than in White's model although the final mass function in our model extends to larger masses. From his Figure 4 we find that the ratio of the mean binding energy per unit mass of the 14% most massive particles having masses in the range $0.51 < m/m^* < 4$ to the 57% light particles with $m < 0.18\ m^*$ is about 2.2. In our model the ratio of the mean binding energies of the 14% most massive particles over the 57% lightest particles is about 1.8. The difference may be due to the depletion of massive galaxies in the central region of the cluster due to merging.

We have also determined the relation between the mean mass of particles and the (local) density of neighbouring galaxies similar to the relation between the frequency of morphological types and local (projected) density determined observationally by Dressler (1980). For each particle the radius of the sphere, centered on the particle, containing the two nearest neighbours was determined. The relation between the mean mass and neighbour density of the particles in these spheres is given in Figure 6. The mass-local density relation in the cluster is in reasonable agreement with the extension of the mass-local density

relation in the expanding universe simulation indicating that a similar mechanism is causing the mass segregation in both models. White (1976) ascribes the steady increase of the binding energy of the most massive particles during the expansion and collapse of the cluster to the effect of two-body relaxation. However, the two-body relaxation formalism is not strictly applicable under the conditions existing during these early stages of the model. In Paper I it was found that the infall model of Gunn and Gott (1972) provided a good explanation of the clustering process during the early stages of the expansion. If the particles are distributed at random initially, the most massive particles will act as condensation nuclei for the hierarchical clustering process, the infall rate being about proportional to mass. This leads to a steady increase of the binding energy of the most massive particles (see White's Figure 4) which tend to be found in the central regions of the subclusters forming during the expansion and subsequent collapse of the cluster.. Note that merging tends to decrease the partical density around massive particles, which leads to lower binding energies and a steeper mass-local density relation.

The velocity distribution of the particles is effected during the formation and merging of subclusters by the time-dependent changes in the mean potential at the position of the particles (Lynden-Bell, 1967). This process, usually called violent relaxation, is independent of the mass of the particles and tends to destroy mass segregation in subclusters as they merge to form larger systems. However, the numerical experiments indicate that violent relaxation is much less effective in destroying mass segregation than one might expect. We can understand this by considering the merging process of subclusters in more detail. When two merging subsystems start to overlap the particles in both systems feel a rapidly changing field which will redistribute the kinetic energy in orbital motion over random motions of the particles in the subsystems. The increase of internal energy in the subsystems is mainly due to the tidal field which the subsystems exert on each other when they overlap. The tidal force affects the outer parts of the subsystems most strongly and the central parts of the merging systems sink towards each other, transferring their orbital energy and angular momentum to the outer parts. Much of the mass segregation in the subclusters may thus survive during the merging process due to the

(7a)



(7b)

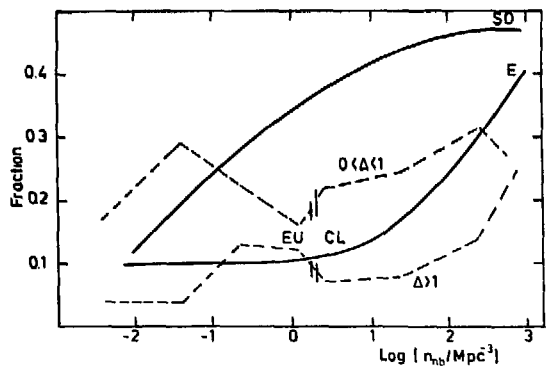


Figure 7. a. The dashed lines give the fraction of particles having Δ in the range indicated as a function of neighbour density (as in Figure 6) at the end of the simulation of Model 1 (to the right of the break) and of Model 5 (to the left of the break). Full lines give the fractions of SO and E galaxies from Dressler (1980).

b. Results of calculations similar to those in Figure 7a, in which the particles have effective sizes smaller by a factor 5.

different behaviour of the core and halo particles in the merging subclusters.

(iv) Δ Segregation.

Since the fraction of particles with large values of Δ increases with mass, it is not surprising that mass segregation is accompanied by a segregation between particles with large and small values of Δ. In Figure 7a the fraction of particles having $\Delta > 1$ and the fraction

having $0 < \Delta < 1$ are given as a function of local neighbour density (as defined above). Note that at high densities the fraction of particles having $\Delta > 1$ increases at the expense of the particles with $0 < \Delta < 1$ due to merging in the central regions of collapsed (sub)clusters. In Figure 7b the results of simulations in which the particles were given effective collison radii of $1/5\ R_g$ are shown. The merging rate *after* the collapse of clusters is now considerably depressed resulting in a higher fraction of particles with $0 < \Delta < 1$.

3. Comparison with Observations.

The calculations discussed in the previous section model the evolution of a cluster in a very crude and schematic way, and it may therefore be useful to mention some of the difficulties that hamper a comparison of the results of the simulations with observations.

(a) Although the initial mass function is consistent with a Poisson distribution of particles, the initial small-scale distribution of the particles may be different from a distribution resulting from hierarchical clustering prior to the starting time of the model.

(b) The particles represent dark galactic halos, formed in the hierarchical clustering process. The epoch of cluster formation is assumed to set in when small luminous disk galaxies are formed in the cores of these halos. These galaxies may then survive during the subsequent clustering process due to their small cross sections (White and Rees, 1978). The destruction of halos via tidal interactions is not accounted for in the model. The merging rate among the simulation particles in the collapsed cluster is therefore probably considerably higher than the merging rate among galaxies in rich clusters. This will predominantly affect the high mass part of the mass spectrum because the massive particles tend to be in the central regions of the cluster.

(c) The mass swallowed by galaxies having small masses is underestimated due to the low mass cut-off in the mass spectrum.

(d) The assumption of constant mass-to-light ratio may not be valid. In Paper II it was found that the clustering properties of early-type galaxies were consistent with a higher mass and higher mass-to-light ratio for these galaxies than for late-type galaxies.

(e) Although it seems plausible to explain the presently observed morphological galaxy types as the result of an evolution along the sequence (Sc-Sb-Sa)-SO-E due to an increase in the parameter Δ (Paper II), we do not know exactly what range in Δ to assign to a particular type.

Keeping these points in mind we will now compare the properties of the simulation particles discussed in the previous section with observed properties of galaxies.

(i) Luminosity functions.

Luminosity functions of galaxies in rich clusters decrease exponentially (or steeper) with increasing luminosity at the bright end if one excludes the brightest cluster members (Schechter, 1976; Dressler, 1978a). Sandage and Hardy (1973) found that these first ranked galaxies tend to be brighter in rich dense clusters than in poor clusters while the next brightest galaxies are less luminous, suggesting that the first ranked cluster member has grown at the expense of the other bright galaxies in the cluster. This effect tends to steepen the bright side of the luminosity function and to shift L^* towards lower luminosities (Hausman and Ostriker, 1978). The final mass function in the cluster simulation is not as steep as the observed luminosity functions. The final mass function in Model 1 can be described reasonably well by a function of the form $m^{-1}\ e^{-m/m_1}$ with $m_1 = 2.5\ m^*$. The flatness of the mass function may partly be due to the low mass cut off and to the high merging rate among the massive particles in the central regions of collapsed (sub)clusters (see b and c above). Note that the run-away of a first ranked galaxy in the centre of a cluster, accompanied by a steepening of the bright side of the luminosity function is seen in the high density cluster of Model 4. A similar effect is also found in other simulations (Aarseth and Fall, 1980).

A number of observational investigations has shown that the fraction of elliptical galaxies increases towards the bright end of the galaxy luminosity function (Van den Bergh and McClure, 1979; Dressler, 1980; Tamman, Yahil and Sandage, 1979; Thompson and Gregory, 1980). In Paper I it was suggested that this could be explained in the infall model if ellipticals are identified with particles that have accreted more than their own mass during the epoch of cluster formation. This result is confirmed here (see Figure 4). Although the amount of accreted

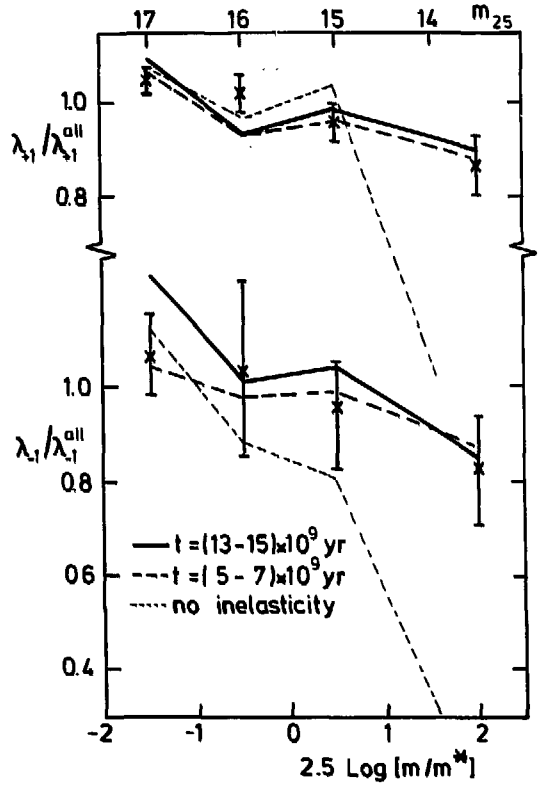


Figure 8. Mean separation distance (λ_{+1}) and harmonic mean separation distance (λ_{-1}) (averaged over the results for different R_{max}; see text) for particles in different mass groups in Model 1 as a function of mass. The over-all values, λ_{+1}^{all} and λ_{-1}^{all} for different R_{max}, are given in Table 4. The dotted line is for a model similar to Model 1 in which the particle encounters were purely elastic. Crosses give the mean separation distances for galaxies in different groups of apparant (visual) magnitude in the Coma cluster (Capelato et al.1980). The upper and lower scales are matched for a mass-to-light ratio of 500 $M_\odot/L_\odot$.

matter is underestimated at small masses, the increase in the fraction of particles with large Δ at $m > m^*$ must be quite realistic. The mass spectrum of particles with large Δ extends to larger masses than the particles with $\Delta = 0$. However, the luminosity function of ellipticals (first ranked galaxies excluded) does not seem to extend to larger luminosities than those of spirals (see Dressler, 1980). This can be explained if both the mass and mass-to-light ratio of ellipticals are

Table 4†

	Model I			Coma
R_{max}	6	3	1.5	1.3
λ_{+1}^{all}	3.0	1.7	1.0	0.95
λ_{-1}^{all}	1.4	0.9	0.6	0.6

† Distances in Mpc's.

about a factor two higher than those of spiral galaxies (see point (d)). Note that this would also yield a steeper luminosity function than the mass function given in Figure 3.

(ii) Luminosity segregation.

The frequency of early-type galaxies in rich clusters increases with the number density of neighbours (Dressler, 1980). In our model this implies the presence of mass segregation. Luminosity segregation will be much weaker than mass segregation if the mass-to-light ratio of these galaxies is larger than the mass-to-light ratio of the less massive spiral galaxies. Luminosity segregation will also be weakened by tidal stripping of galaxies in the central regions of rich clusters (section 4.1). Therefore, it should perhaps not be surprising that little evidence for luminosity segregation in rich clusters has been found (Oemler, 1974; Dressler, 1980; see also Bahcall, 1977 and references therein). Nevertheless, recent studies of the luminosity distribution in the Coma cluster by Quintana (1979) and by Capelato et al. (1980) have revealed a definite luminosity segregation in this cluster. Quintana found that the core radius of the radial distribution of galaxies in Coma increases with decreasing galaxy luminosity. Capelato et al. have shown that the mean separation and mean harmonic separation of galaxies in different luminosity classes depend systematically on luminosity. The mean separation and harmonic mean separation of the galaxies in a luminosity group are defined by

$$\lambda_{+1} = \frac{1}{N_p} \sum_{i>j} r_{ij},$$

and by

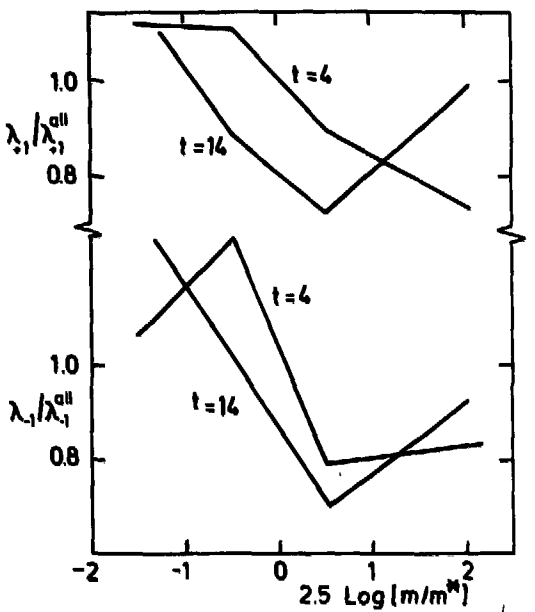


Figure 9. Mean separation distance and harmonic mean separation distance as a function of mass in Model 4 (see Figure 8).

$$\lambda_{-1} = \left[\frac{1}{N_p} \sum_{i>j} \frac{1}{r_{ij}} \right]^{-1}$$

respectively, where N_p is the number of all pairs in that luminosity group and r_{ij} are the pair separations. In Figure 8 the results obtained by Capelato et al. are compared with the results from Model 1 (averaged over the situation at 13, 14 and 15×10^9 yr after the start of the simulation). The mean separations for all particles, λ_{+1}^{all} and λ_{-1}^{all}, depend on the size of the region considered. In Table 4 the values for λ_{+1}^{all} and λ_{-1}^{all} in the cluster model are given for different sizes of the spherical region centered on the centre of mass of the cluster. Capelato et al. have studied a region of radius 1.3 Mpc (H_o = 50 km sec^{-1} Mpc^{-1}) around the centre of Coma. The mean radii in the cluster model given in Figure 8 are obtained by averaging the results for R_{max} = 1.5, 3 and 6 Mpc to decrease the statistical noise due to small numbers. This probably does not weaken the comparison since no systematic change in the correlation between the mean separations with luminosity was found when varying R_{max} in the cluster model. Comparing the final situation in

Model 1 with: (1) the results of Capelato et al., (2) the situation in Model 1 directly after its collapse ($t \approx 6\ 10^9$ yr), and (3) the final situation in a model, similar to model 1, in which encounters among particles were fully elastic, we reach the following conclusions.

(1) The mass segregation in the numerical cluster model is consistent with the observed segregation of luminosity in Coma.

(2) Mass segregation comes about mainly during the initial expansion and subsequent collapse of the cluster.

(3) The mean separation distance of the particles in the highest mass group is significantly increased due to merging among these particles. The increase in mean separation distance among the brighter cluster members in model 4 is illustrated in Figure 9. This confirms Dressler's interpretation of the observed lack of luminosity segregation in cD clusters (Dressler, 1978b) as the result of merging of bright galaxies with the central giant.

(iii) Type segregation.

A well defined relation between the frequencies of E and SO galaxies and local density was found by Dressler (1980). In Figure 7a his results are compared with the relation between local density and the fractions of particles having $\Delta > 1$ and $0 < \Delta \leq 1$ respectively. The total fraction of early-type (E + SO) galaxies agrees reasonably with the fraction of particles having $\Delta > 0$. At low densities there is reasonable agreement between the fractions of elliptical and SO galaxies and the fractions of particles with $\Delta > 1$ and $0 < \Delta \leq 1$ respectively. At higher densities however, the fraction of ellipticals is lower than the fraction of particles with $\Delta > 1$ and the fraction of SO's is higher than the fraction of particles with $0 < \Delta < 1$. The rise of the fraction having $\Delta > 1$ at the expense of the fraction having $0 < \Delta < 1$ must be due to merging among particles in the central regions of collapsed clusters (see point b above). This is supported by the results shown in Figure 7b of simulations in which the effective radius of the particles is smaller by a factor 5.

4. Evolution after Collapse of the Cluster.

4.1 Tidal stripping.

As the velocity dispersion of the galaxies in the protocluster rises, the merging rate drops, but the collision rate steadily increases. The internal energy of two colliding galaxies increases due to the tidal force which they exert on each other. This will lead to mass loss from the outer parts of the galaxies. Galaxies in rich clusters may have lost a considerable fraction of their halos by this mechanism (Gallagher and Ostriker, 1972; Richstone, 1974). A realistic numerical simulation of the tidal stripping proces requires that the particles representing the stripped halo mass are less massive than the galactic particles. This implies that the number of particles in the simulation rises inversely proportional to the mass of the galactic particles. Therefore, we have not allowed mass loss to occur in the simulations. Instead, we will give here a simple analytic estimate of the mass loss occurring after the collapse of the cluster using the cross sections obtained in Paper I.

The fractional change of internal energy of a galaxy colliding head-on with another galaxy of equal mass is approximately given by

$$\Delta U/U = 6\ (\sigma_g/V_{coll})^2$$

(see Paper I), where V_{coll} is the relative velocity of the colliding galaxies at smallest separation distance. The relative velocity at infinity is related to V_{coll} by

$$V^2 = V_{coll}^2 - 8\ \sigma_g^2$$

where σ_g is (3-dimensional) internal velocity dispersion of the galaxies. For high relative velocities ($V >> \sigma_g$) the fractional change in internal energy can be approximated by

$$\Delta U/U = 3.6\ (\sigma_g/V)^2 \qquad (2a)$$

which is in reasonable agreement with the impulsive approximation

(Spitzer, 1958). The decrease of $\Delta U/U$ with increasing impact parameter p can be approximated for $p < R_g$ by

$$\Delta U/U \propto (1 - 0.8\ p/R_g) \tag{2b}$$

where R_g is twice the radius containing half the total mass M_g of the galaxy. For $p > R_g$ the impulsive approximation suggests $\Delta U/U \propto (p/R_g)^{-4}$. For a homogeneous distribution of identical galaxies having number density n_g, the rate of change in internal energy of a galaxy is given by

$$\frac{1}{U}\frac{dU}{dt} = n_g\ \pi\ R_g^2 \int_0^\infty \int_{\sigma_g}^\infty \psi(p)\ \phi(V)\ 3.6\ \left(\frac{\sigma_g}{V}\right)^2\ V\ dV\ dp$$

where

$$\psi(p) = \begin{cases} (1-0.8\ p/R_g)\ p/R_g^2\ , & p < R_g \\ (0.2\ (p/R_g)^{-4}\ p/R_g^2\ , & p > R_g \end{cases} \tag{3}$$

The velocity distribution of the galaxies in a collapsed and virialised cluster is approximately Maxwellian. The distributions of relative velocities is then given by

$$\phi(V)\ dV = \frac{3^{3/2}}{2\ \sqrt{\pi}\ \sigma_{cl}^3}\ e^{-3V^2/4\sigma_{cl}^2}\ \ V^2\ dV, \tag{4}$$

where σ_{cl} is the (3-dimensional) velocity dispersion of galaxies in the cluster. The lower limit of the integration over V is determined by σ_g since collisions with $V < \sigma_g$ result in merging of the colliding galaxies. Performing the integrations in (2) yields

$$\frac{1}{U}\frac{dU}{dt} \simeq 2.5\ n_g\ \pi R_g^2\ \sigma_g\ (\sigma_g/\sigma_{cl})\ e^{-3\ \sigma_g^2/4\sigma_{cl}^2}.$$

Note that the fractional change of energy found here is more than an order of magnitude larger than Richstone's result (1976). The behaviour of the energy change with collision velocity in his simulations is in conflict with the impulsive approximation. Presumably, the discrepancy is due to deficiencies of his model. Substituting parameter values appropriate for a galaxy moving in the core of typical rich clusters like Coma (Abell 2256) we find

$$\frac{1}{U}\frac{dU}{dt} = K_c \left[\frac{R_g}{100\ \text{kpc}}\right]^2 \tag{6}$$

where

$$K_c = 3.2 \left[\frac{\sigma_{cl}}{1000\ \text{km sec}^{-1}}\right]^{-1} \left[\frac{n_c}{1000\ \text{Mpc}^{-3}}\right] \left[\frac{\sigma_g}{200\ \text{km sec}^{-1}}\right]^2 (10^9\text{yr})^{-1} \tag{7}$$

We do not know how the density distribution of a galaxy is affected when its internal energy increases due to one or several collisions. However, the impulsive approximation and the results of numerical simulations suggest that the outer parts will be preferentially stripped leaving the central regions intact. We will assume that galaxies are approximately isothermal within some tidal radius R_g and that the velocity dispersion in the central parts of the galaxy does not change in collisions, implying

$$-dU/U = dM_g/M_g = dR_g/R_g. \tag{8}$$

It may be worthwhile to mention that the stripped mass is not necessarily lost from the galaxy in the sense that it is not bound to the galaxy anymore. A large fraction of the stripped mass may form an extended, loosely bound halo around the galaxy which is easily stripped in subsequent collisions. Integration of (6) using (8) yields

$$R_g(t) = R_g(0)\ (1 + 2\ K_c \left[\frac{R_g(0)}{100\ \text{kpc}}\right]^2\ t_9)^{-\frac{1}{2}} \tag{9}$$

where t_9 is the time elapsed since the collapse of the cluster in 10^9 yr. This result indicates that galaxies (or galaxy halos) having an initial size of 100 kpc, will be stripped to a radius of $\sim$ 18 kpc within 5×10^9 yr when they are moving in the core of Coma. Outside the core region the ratio $R_g(t)/R(0)$ decreases as $\langle n_g\rangle^{\frac{1}{2}}$ where $\langle n_g\rangle$ is the mean density along the orbit of a galaxy. The density profiles in rich clusters are well described by bounded isothermal models in which the density decreases outside a core radius R_c as

$$n(r) = n_c(r/R_c)^{\gamma} \quad \text{for} \quad R_c < r < R_{cl}$$

where the core radius R_c is typically about 300 kpc, $R_{cl} \approx 10\ R_c$ and $\gamma \approx -2.6$ (Bahcall, 1977, see also Seldner and Peebles, 1977). The mean density along the orbit of galaxies moving at a distance r from the cluster centre varies as $n_g \propto r^{-2.6}$ for galaxies on circular orbits. Galaxies moving on radial orbits through the cluster core spend about a fraction r/R_c of an orbital period in the core indicating that $\langle n_g \rangle \approx n_c(r/R_c)^{-1}$. The size of galaxies at a distance r from the centre of a cluster can now be estimated from (9) if K_c is replaced by

$$K(r) = K_c\ (r/R_c)^{-(1 \text{ to } 2.6)}. \qquad (10)$$

Observational evidence that galaxies on the central regions of rich clusters are being tidally stripped has been presented by Strom and Strom (1978, see also Kormendy, 1977). They show that the characteristic size of galaxies in rich clusters decreases with increasing galaxy density as $\sim n_g^{-0.3}$. In the central regions of cD clusters, the effective radius of ellipticals, defined as the radius containing half the total luminosity, is about a factor 4 smaller than in Spiral-rich clusters while their luminosity is about ½ magnitude smaller. The effective radius of typical bright ellipticals ($M_V = -21.5$) in Spiral-rich clusters is about $10^{0.8}$ kpc. The density profiles of elliptical galaxies are well represented by a Hubble law (Hubble, 1930) for which L(r), the total luminosity within a distance r of the centre, increases as ln r in the outer regions of the galaxy. The radius containing all the light is therefore about $10^{0.8} \times e^2 \approx 50$ kpc. From (9) we see that the radius of a galaxy of this size, moving in the core of a rich cluster will indeed decrease by about a factor 4 in 10^{10} yr. For an elliptical galaxy this implies a decrease in total luminosity of about 40%. The mean distance to the cluster centre of the galaxies within 1 Mpc ($\sim 3R_c$) of the cluster centre is about $2R_c$. Using (9) we find that these galaxies will decrease by about a factor 2.5 in size. For the galaxies between $3R_c$ and $10R_c$ this is about 1.5. These estimates are in agreement with the findings of Strom and Strom and support their suggestion that the dependence of galaxy size on cluster density is due to tidal stripping of galaxies in collisions.

The above analysis suggests that a significant fraction of the total light from the central region (< 1 Mpc) of a dense rich cluster

originates from a diffuse background component of material stripped from the cluster galaxies. Such a component does not contribute more than 20% to the total light emitted by the Coma cluster (Melnick et al., 1978, see also Thuan and Kormendy, 1977). In some cD clusters the central giant has a huge envelope extending over $\sim$ 1 Mpc. There are two observational indications that these envelopes consist of material stripped from the cluster galaxies: (i) the stellar velocity dispersion in the envelope of the cD in the Abell cluster A 2029 increases with radius (Dressler, 1979) and (ii), cD's in poor groups, in which high velocity encounters are clearly unimportant, have total luminosities comparable to that of cD's in rich clusters, but do not have extended envelopes (Schild and Davis, 1979; Thuan and Romanishin, 1981).

4.2 Evolution of the merging rate.

The merging probability per galaxy in a homogeneous distribution of identical galaxies having number density n_g can be calculated from

$$P_{merger} = n_g \pi R_g^2 \int_0^{P_{max}} (p/R_g)\, d(p/R_g) \int_0^{V_{max}} \phi(V)\, V^3\, dV$$

where

$$P_{max} = 2 R_g$$

and

$$V_{max} = 1.1 \sigma_g (1 - p/4 R_g)$$

(Paper I). The velocity integral yields

$$\frac{4}{\sqrt{3\pi}} \sigma_{cl} \left[1 - e^{-3V_{max}^2/4\sigma_{cl}^2} (1 + 3V_{max}^2/4 \sigma_{cl}^2)\right] =$$

$$\propto \frac{2}{\sqrt{3\pi}} \sigma_{cl} (3 V_{max}^2/4\sigma_{cl}^2)^2 \text{ for } V_{max} << \sigma_{cl},$$

and we find

$$P_{merger} \simeq 1.7\ (10^9\ yr)^{-1}\ (\sigma_g/\sigma_{cl})^3 \left[\frac{n_g}{1000\ Mpc^{-3}}\right] \times$$
$$\times \left[\frac{R_g}{100\ Kpc}\right]^2 \left[\frac{\sigma_g}{200\ km\ sec^{-1}}\right] \quad \text{for } \sigma_g << \sigma_{cl}. \tag{11}$$

Applying this result to our Model 1 cluster after collapse and virialization using $\langle n_g \rangle \simeq 100\ Mpc^{-3}$ and $\sigma_{cl} \simeq 1300\ km\ sec^{-1}$ we find a merging probability per galaxy of about $10^{-2}\ (10^9\ yr)^{-1}$ in good agreement with the merging probability given in Table 2. The result given above can be generalized using (9) and (10) to obtain the merging probability as a function of distance to the cluster centre of time passed since the collapse of the cluster. We find

$$P_{merger}\ (r,t) = 0.11\ (10^9\ yr)^{-1}\ (\sigma_g/\sigma_{cl})^2\ K(r)\ R_{100}^2 \times$$
$$\times\ (1 + 2\ K(r)\ R_{100}^2\ t_9)^{-1} \tag{12}$$

where R_{100} is the radius of the galaxies in 100 Kpc at t = 0. For $t_9 >> (2\ K(r)\ R_{100}^2)^{-1}$ this yields

$$P_{merger} \simeq 0.05\ (10^9\ yr)^{-1}\ (\sigma_g/\sigma_{cl})^2\ t_9^{-1} \tag{13}$$

and the merging probability is independent of the distance to the cluster centre.

The mean merging probability of the galaxies in numerical cluster models drops by about a factor 10 during the collapse of the cluster due to the increase in velocity dispersion in the cluster. From the above analysis we conclude that after collapse the merging probability decreases further due to tidal stripping of galaxies, approaching the density independent value given by (13). For a typical rich cluster for which $t_9 \sim 5$ this is about a factor $(\sigma_g/\sigma_{cl})^2$ smaller than the mean merging rate among bright galaxies in the universe (Paper II).

4.3 Dynamical friction and cannibalism.

Loss of orbital energy of galaxies in clusters due to inelastic collisions does not play an important role if the orbital energy of the

galaxies is much larger than their internal energy. However, the most massive galaxies in a cluster tend to lose orbital energy and spiral towards the cluster centre due to interaction with the lighter galaxies (two-body relaxation) or with a smooth background of stripped halo material (dynamical friction). In both cases the time scale for loss of orbital energy of a galaxy of mass M_g moving in a region of density ρ is approximately given by

$$t_{df} = 10^{10}\ \mathrm{yr} \left[\frac{\sigma_{cl}}{1000\ \mathrm{km\ sec^{-1}}}\right] \Big/ \left[\frac{M_g}{10^{12}\ M_\odot}\right] \left[\frac{\rho}{10^{15}\ M_\odot\ \mathrm{Mpc^{-3}}}\right] \qquad (14)$$

The most massive galaxies in the core of a rich cluster might thus form a central giant (cD) within a Hubble time if they do not lose too much mass in collisions. The growth of a central giant in the centre of rich clusters has been studied by Ostriker and Tremaine (1975, see also Hausman and Ostriker, 1978). It may be possible to estimate the present growth rate of first ranked galaxies (FRG's) in clusters from the fraction that is observed to contain multiple nuclei (Rood and Leir, 1979; Hoessel, 1980). According to Rood and Leir about 25% of the FRG's are double with projected separation smaller than $\sim$ 30 kpc (H_o = 50 km sec^{-1} Mpc^{-1}). They give a magnitude difference between the two components of the doubles of $\Delta M \leq 1$. The secondary component would then be one of the $\sim$ 5 brightest galaxies in the cluster. This makes chance projection very unlikely and, at the same time, yields a present mass doubling time for FRG's which is comparable to the Hubble time (Tremaine, 1981). However, both conclusions depend critically on the luminosity of the secondary component, which may not be easy to determine. The total luminosity of the double FRG's in Hoessel's sample is only about 10% higher than the mean luminosity of the other FRG's which may indicate that the two components differ by more than two magnitudes. This would mean that (i) a considerable fraction of the double nuclei may be chance projections, and (ii) the estimated growth rate, which is proportional to the mass of the secondary component squared, is much smaller, even if all the doubles are bound systems. The investigation of this paper suggests that cD's are formed at an early stage in the evolution of clusters. They may still be growing in collapsed rich clusters at the present epoch but probably at a much lower rate due to stripping of cluster galaxies in high velocity collisions.

5. Discussion and conclusions.

We have numerically simulated the evolution of a rich cluster of galaxies in the context of the hierarchical clustering model. The relatively small number of particles whose equations of motion can be integrated within reasonable limits of computation time impedes a realistic simulation of the formation of rich clusters of galaxies in several ways: (i) both the small and large-scale initial distributions of the particles should be the result of the clustering process prior to the starting time of the model. Our choice of initial conditions may be a poor representation of such a situation. (ii) The cut-off in the mass-function at small masses has some notable effects on the mass function evolution under merging (see section 2.2.2). (iii) Mass loss from galaxies in encounters is not accounted for in the model. (iv) No distinction was made between dark and luminous matter.

Despite these deficiencies it is possible to draw some important conclusions.

1. The merging probability of the particles in the simulation is found to be about proportional to their mass in agreement with the results of Paper I.

2. Mass segregation in the final cluster is strong (see also White, 1976). Mass segregation is already present in the subcluster that form before the collapse of the cluster and can be interpreted in the infall model. It is not destroyed when subclusters merge to form the final cluster. The amount of mass segregation in terms of the change with mass of the mean separation among the particles in different mass groups is in agreement with observed mass segregation in the Coma cluster.

3. As a result of 1 and 2 the fraction of particles having large Δ, the ratio of accreted mass over initial mass, increases with mass. This is consistent with recent observations showing that the fraction of ellipticals increases with luminosity.

4. The mean separation among the most massive particles increases due to merging among these particles supporting Dressler's (1978b) interpretation of the apparent absence of luminosity segregation in cD clusters.

5. The relation between local density and fraction $f(\Delta)$ (or f_Δ) of galaxies within some range of Δ is in reasonable agreement with the morphological type-local density relation found by Dressler if we

identify the particles having $\Delta > 1$ with ellipticals and the particles having $0 < \Delta < 1$ with early-type disk galaxies. The steep increase in $f(\Delta > 1)$ at the expense of $f\ (0 < \Delta < 1)$ in the simulation can probably be attributed to the fact that the galaxies are not allowed to lose mass which leads to merging rates for these particles which are too high.

6. An estimate of the mass loss rate of galaxies in rich clusters due to collisions with other galaxies indicates that galaxies in the core of a rich cluster will be tidally stripped to a radius of about 20 kpc within $\sim 5\ 10^9$ yr. The estimated decrease in size of galaxies in rich clusters seems in good agreement with the observational results of Strom and Strom (1978).

7. The merging probability of galaxies in rich clusters decreases due to tidal stripping of galaxies and approaches a density independent value which is about a factor $(\sigma_g/\sigma_{cl})^2$ lower than the mean merging probability of galaxies outside rich clusters, where σ_g and σ_{cl} are the velocity dispersion in the galaxies and in the clusters respectively.

The situation in the core of a rich cluster may be somewhat more complicated since massive galaxies tend to be dragged towards the cluster centre where they may form, or be swallowed by, a central cannibal. The merging rate for this central giant probably has been decreasing after the collapse of the cluster, but it may still be higher than for galaxies in the field.

The observed decrease with z ($z < 1$) in the fraction of Spiral-like galaxies in rich clusters (Butcher and Oemler, 1978) may be related to the stripping of galaxies in the initial epoch shortly after collapse of clusters. The dependence of merging probability on environment may be important for models in which activity in galactic nuclei is caused by mergers (Roos, 1981). The steep evolution of the merging rate by a factor of order 100 during and after the collapse of a rich cluster, which occurs at a relatively recent epoch ($z \leq 1$) warrants a comparison with the cosmological evolution of active galaxies.

References.

Aarseth, S.J. and Hills, J.G., 1972, Astron. and Astrophys. 21, 255.
Aarseth, S.J. and Fall, S.M., 1980, Ap.J. 236, 43.

Bahcall, N., 1977, Ann. Rev. of Astron. and Astrophys. 15, 505.
Butcher, N. and Oemler, A., 1978, Ap.J. 226, 120.
Capelato, H.V., Gerbal, D., Mathez, G., Mazure, A., Salvador-Solé, E. and Sol, H., 1980, Ap.J. 241, 521.
Carter, D., 1977, M.N.R.A.S. 178, 137.
Dressler, A., 1978a, Ap.J. 223, 765.
Dressler, A., 1978b, Ap.J. 226, 55.
Dressler, A., 1978, Ap.J. 222, 23.
Dressler, A., 1979, Ap.J. 231, 659.
Dressler, A., 1980, Ap.J. 236, 351.
Efstathiou, G., Fall, S.M. and Hogan, C., 1979, M.N.R.A.S. 189, 203.
Faber, S.M and Jackson, R.E., 1976, Ap.J. 204, 668.
Gallagher, J.S. and Ostriker, J.P., 1972, Ap.J. 77, 288.
Gisler, G.R., 1980, Astron. J. 85, 623.
Gunn, J.E. and Gott, J.R., 1972, Ap.J. 176, 1.
Hausman, M.A. and Ostriker, J.P. 1978, Ap.J. 224, 320.
Hoessel, J.G., 1980, Ap.J. 241, 493.
Jones, B.J.T. and Efstathiou, G., 1979, M.N.R.A.S. 189, 27.
Kormendy, J., 1977, Ap.J. 218, 333.
Melnick, J., White, S.D.M. and Hoessel, J., 1978, IAU Symp. no. 79.
Oemler, A., 1974, Ap.J. 194, 1.
Ostriker, J.P. and Tremaine, S.D., 1975, Ap.J. Lett. 202, 113.
Peebles, P.J.E., 1970, Astron. J. 75, 13.
Peebles, P.J.E., 1980, "The Large-Scale Structure of the Universe", Princeton University Press, Princeton, New Jersey.
Press, W.H. and Schechter, P., 1974, Ap.J. 187, 425.
Quintana, H., 1979, Astron. J. 84, 15.
Richstone, D.O., 1976, Ap.J. 204, 642.
Rood, H.J. and Leir, A.A., 1973, Ap.J. Lett. 231, 3.
Roos, N. and Norman, C.A., 1979, Astron. and Astrophys. 76, 75, Paper I.
Roos, N., 1981, Astron. and Astrophys., 95, 349, Paper II.
Sandage, A. and Hardy, E., 1973, Ap.J. 183, 743.
Schechter, P. 1976, Ap.J. 203, 297.
Schild, R. and Davis, M., 1979, Astron. J. 84, 311.
Seldner, M., and Peebles, P.J.E., 1977, Ap.J. 215, 703.
Spitzer, L., 1958, Ap.J. 127, 17.
Strom, S.E. and Strom, K.M., 1978, Ap.J. Lett. 225, 93

Tammann, G.A., Yahil, A. and Sandage, A., 1979, Ap.J. 234, 775.

Thompson, L.A. and Gregory, S.A., 1980, Ap.J. 242, 1.

Thuan, T.X. and Kormendy, J., 1977, Publ. Astron. Soc. Pacific, Vol. 89, 466.

Thuan, T.X., and Romanishin, W., 1981, preprint.

Toomre, A. and Toomre, J., 1972, Ap.J. 178, 662.

Toomre, A., 1977, in "The Evolution of Galaxies and Stellar Populations" eds. B.M. Tinsley and R.B. Larson (New Have, Yale University Observatory).

Tremaine, S.D., 1980, preprint.

Van den Bergh, S. and McClure, R.D., 1979, Ap.J. 231, 671.

White, S.D.M., 1976, M.N.R.A.S. 177, 717.

White, S.D.M., 1977, M.N.R.A.S. 179, 33.

White, S.D.M. and Rees, M.J., 1978, M.N.R.A.S. 183, 341.

CHAPTER V

GALAXY MERGERS AND ACTIVE NUCLEI

Abstract.

The tidal disruption rate of stars near a central black hole in a galactic nucleus may be considerably enhanced during a merger of a galaxy with a smaller companion due to scattering of stars into loss-cone orbits by the perturbing gravitational field of the intruding galaxy. For conditions in galactic nuclei similar to those in the nucleus of the Milky Way, a maximum luminosity of $\sim 10^{46}$ erg sec^{-1} and a total energy production of about 10^{59} $[M_{hole}/10^7\ M_\odot]$ erg in $\sim 10^7$ yr may be obtained in such an event. Using a merging rate consistent with observational data and with cosmological simulations the fraction of galaxies which are active at some level of total luminosity is estimated. The agreement between the model prediction and observed luminosity functions of active galactic nuclei is encouraging. The cosmological evolution of active galaxies is discussed briefly in the context of the merger model.

If both merging galaxies contain a central black hole, a binary black hole system will be formed at the centre of the combined galaxy. Numerical three-body experiments were performed to investigate the evolution of the binary due to three-body interactions with galactic stars. It is concluded that the minimum separation distance that the binary may reach via interactions with these stars is probably sufficiently small so that gravitational damping can take over and ensure merging of the holes within a Hubble time.

1. Introduction.

Massive black holes are frequently invoked as the primary component of the powerful energy sources in QSO's and active galactic nuclei. Such objects may have grown by accretion of gas produced by tidal disruption of stars in the vicinity of the hole or by star-star collisions in the nuclei of galaxies (Hills, 1975; Young, 1977; Frank, 1978).

Tidal disruption of stars near the black hole is a particular attractive fuelling mechanism because it occurs very close to the hole. The debris of the tidally disrupted stars will be strongly bound to the hole and is likely to be swallowed by it. In galactic nuclei of typical densities 10^7 $M_\odot$ pc^{-3} and core radii of about one pc this may yield luminosities of $\sim 10^{44}$ erg sec^{-1}. The disruption rate of stars near the hole is limited by the existence of a loss-cone within some critical radius of the hole. Within this critical radius stars that are disrupted near the hole cannot be replaced by other stars via two-body relaxation in one stellar orbital time (Frank and Rees, 1976; Lightman and Shapiro, 1977). In order to explain the high luminosities of $\gtrsim 10^{46}$ erg sec^{-1}, observed in active galaxies and QSO's, higher central densities are required. In such nuclei, however, most gas is produced in star-star collisions occurring at a much larger distance from the hole (Frank, 1978) and the fraction of gas consumed by the hole is again uncertain.

In previous investigations the galaxies were considered as isolated systems, undisturbed by other galaxies. However, interactions among galaxies are probably quite common. Most galaxies are observed to have close companions (Holmberg, 1969) and the appearance of peculiar galaxies can often be interpreted as the result of tidal interactions with a nearby galaxy (Toomre and Toomre, 1972). Such interactions, which are likely to result in rapid merging of the two stellar systems within a few galactic rotation times, may have played an important role in the evolution of galaxies. It seems therefore worthwhile to investigate whether there might be a link between merging and activity in galactic nuclei. There are several reasons for suspecting such a relation. It is well known that strong radio sources are mostly associated with elliptical galaxies, and it has been argued that ellipticals may well be merger products (Toomre and Toomre, 1972; Jones and Efstathiou, 1979; Aarseth and Fall, 1980; Roos, 1981). Furthermore some nearby active elliptical galaxies show signs of recent merging, e.g. Fornax A (Schweizer, 1980) and Centaurus A (Tubbs, 1980). Finally, precession of radio beams, as inferred from rotational symmetry observed in many extended radio sources, may be explained by the existence of a binary black hole in the nucleus produced by merging (Begelman, Blandford and Rees, 1980).

In this paper we will investigate a simple mechanism to enhance

the tidal disruption of stars near a black hole in a galactic nucleus during a merger: repopulation of loss-cone orbits via scattering of stars by the intruding galaxy. The tidal disruption of stars near a black hole in a galactic nucleus may be enhanced considerably over the (quasi-) stationary disruption rate in this way, and luminosities of $\sim 10^{46}$ erg sec^{-1} may be obtained for nuclear densities of $\sim 10^7\ M_\odot\ pc^{-3}$.

We know very little about the conditions in galactic nuclei, but for our galaxy we have information on the mass distribution to within 1 pc of its centre (Oort, 1977). The density keeps rising with decreasing distance to the centre as $r^{-1.8}$ to within r = 1 pc, where it reaches a value of about $10^6\ M_\odot\ pc^{-3}$. If our galaxy contains a massive central object of a few times $10^6\ M_\odot$ as suggested by infrared observations (Wollman et al., 1977), this is about the distance where the stellar distribution starts to obey the cusp solution with density rising as $r^{-7/4}$ and velocity dispersion as $r^{-\frac{1}{2}}$ (Bahcall and Wolf, 1976). There is no evidence for the existence of a core region where the density profile is flat. A similar situation may exist in other galaxies. Observed surface brightness profiles of elliptical galaxies are often fitted to a Hubble or a King law. However, Schweizer (1979) has pointed out that if the effect of seeing is properly taken into account there is little evidence that galaxies have flat central density profiles. This has important implications for fuelling models of central black holes. The high central density in our galaxy is favourable for growth of a black hole via accretion of gas produced by tidal disruption of stars or in star-star collisions. These processes clearly cannot be very important if massive black holes are embedded in extended low-density cores. In this paper we will generally take the conditions in our galaxy as fairly representative for conditions in other galaxies and assume that galaxies do not have flat central density profiles.

In section 2 we will first discuss the dynamical evolution of a merging pair of galaxies one or both of which containing a central massive black hole. If both stellar systems contain a black hole a binary black hole will be formed at the centre of the combined galaxy due to dynamical friction. The next stage of evolution of the binary depends on the amount of energy that it loses to stars coming within its orbital radius. Numerical three-body experiments were performed to explore this stage of the binary evolution. The results of these simula-

tions are presented in section 2.2 and their implications discussed in section 2.3. In section 3.1 we briefly review the (quasi-)stationary fuelling rate of a massive black hole in a galaxy which is not disturbed by intruders. In section 3.2 the tidal disruption rate during a merger is estimated and discussed for some particular cases. Some other observable phenomena which may accompany the activity of a galaxy in the model presented here, such as precession of radio beams and activity of the secondary component, are also discussed. In section 4 our knowledge of the merging rate among galaxies and of the dynamics of the merging process is combined with the results of section 3 to estimate the fraction of galaxies that are active at a given level at the present epoch. The total luminosity function of active galaxies is then estimated from available radio, optical and X-ray luminosity functions and compared with our predictions. The cosmological evolution of the space density of active galaxies is also discussed briefly. Results are discussed and summarized in section 5.

2. Dynamics of the Merging Process.

2.1 Merging of galaxies and the formation of binary black holes.

The merging process of a small galaxy having mass m_g can be described in terms of dynamical friction. The smaller galaxy, moving with velocity V at a distance r from the centre of the larger galaxy, loses orbital energy E according to

$$\frac{dE}{dt} = -4\pi\, G^2\, \frac{m_g^2\ \rho(r)}{V(r)}\, \ln \Lambda, \qquad (1)$$

if $V(r) \gtrsim \sigma(r)$, the one-dimensional velocity dispersion of the stars in the larger galaxy (Spitzer, 1962). G is the gravitational constant. The standard Coulomb logarithm $\ln \Lambda$ will be of order unity if we take Λ equal to the ratio of the sizes of the two galaxies. The density of the larger galaxy at r is given by $\rho(r)$. We assume that this density is given by a power law

$$\rho(r) = \rho_o \, (r/r_o)^{\gamma} \qquad (2)$$

where ρ_o is the density at some reference distance r_o. It will be generally assumed in this paper that the stellar distribution is isothermal, implying $\gamma = -2$ ($\sigma(r) = \sigma$). Cosmological simulations indicate that most mergers occur between galaxies in bound orbits (e.g. Aarseth and Fall, 1980). The orbital velocity V of the two galaxies will then be comparable to σ, and the dynamical friction time at r, defined as $t_{df} = (\frac{1}{E}\frac{dE}{dt})^{-1}$ is well approximated by

$$t_{df}\,(r) = t_{dyn}\,(r)\; M_g(r)/m_g. \qquad (3)$$

The dynamical time $t_{dyn}(r)$ is defined as r/σ and $M_g(r)$ is the large galaxy's mass within r. As the smaller galaxy spirals inward, it will be tidally stripped and its density distribution will be truncated at some radius r_t where the densities of the two systems become comparable. If both galaxies have a density distribution of the form (2) with indices γ and γ', the mass ratio of the two galaxies varies as $M_g(r)/m_g(r_t) \propto r^{3(\gamma-\gamma')/\gamma'}$. For $\gamma = \gamma'$ this is a constant and the smaller galaxy spirals inward with radial velocity $\sigma\, m_g/M_g$. At some distance the galactic density profiles must change: either there is a core or a central black hole. If the smaller galaxy does not contain a black hole it will lose its identity when the density of the large galaxy becomes equal to the core density of the smaller one. More interesting is the case in which both galaxies contain a massive central black hole. At some separation distance where $M_g(r)$ is comparable to M, the mass of the black hole at the centre of the more massive galaxy, the black holes form a binary. Dynamical friction tends to circularize the orbit of a pair of merging galaxies and the eccentricity of the black hole binary orbit is therefore probably very small. As the orbit shrinks dynamical friction becomes less effective since stars can only interact with the binary when the encounter lasts longer than one orbital period (Heggie, 1975). When the binary separation becomes less than the cusp radius, defined by

$$r_h = G\,M/\sigma^2 \qquad (4)$$

the orbital velocity rises above σ (the binary becomes hard). In this

regime the binary loses energy by the slingshot mechanism: stars that come within the orbit of the binary are ejected with velocities comparable to the orbital velocity of the binary. The evolution via three-body interactions is halted at a radius r_{lc} when the number of stars on orbits passing through r_{lc} becomes too small to carry away the binding energy of the binary. Before we can discuss the final evolutionary stages of the binary we have to know how efficient the slingshot mechanism is in removing energy from a massive binary, and since the three-body problem is a classic unresolved problem we must resort to numerical experiments to obtain the necessary information.

2.2 Numerical calculations of slingshot efficiency.

Large regions in the phase space of the three-body problem have been explored numerically by Saslaw et al. (1974), Valtonen (1975) and by Hills (1975a). More recently Hills and Fullerton (1980) have performed a numerical study of the energy transfer as a function of m_*, the mass of the incoming particle. Their experiments, however, are restricted to the case of an equal mass binary on a circular orbit.

Let us define the slingshot efficiency η_{sl} by

$$\langle \Delta E \rangle = \eta_{sl} \tfrac{1}{2} m_* V_{orb}^2 \qquad (5)$$

where ΔE is the fractional energy loss of the binary per interaction and V_{orb} is the orbital velocity of the binary, having components of mass m and M. The binary is assumed to be surrounded by a spherically symmetric distribution of particles with an isotropic Maxwellian velocity distribution with dispersion smaller than the orbital velocity of the binary. The average in (5) is taken over interactions with particles that come within a specified distance R_{max} of the binary's centre of mass. The slingshot efficiency may depend on R_{max}, m, M and e, the eccentricity of the binary orbit.

A standard variable-order variable-step Adams integration routine was used to integrate the set of differential equations describing the three-body system (Hall and Watt, 1976). The fractional change in orbital energy due to numerical effects depends on e. It is less than 10^{-6} per orbital period for e = 0 and rises to 10^{-4} for e = 0.99 (cf. Van

Albada, 1968). The values of η_{sl} were determined in runs in which 500 particles, having mass $m_* << m$, were allowed to interact one by one with the binary. The particles start from a shell, centered on the centre of mass of the binary, with a radius four times the semi-major axis of the binary. When a particle passes the shell again on its way outward its change in energy and angular momentum are measured and the next particle is sent off.

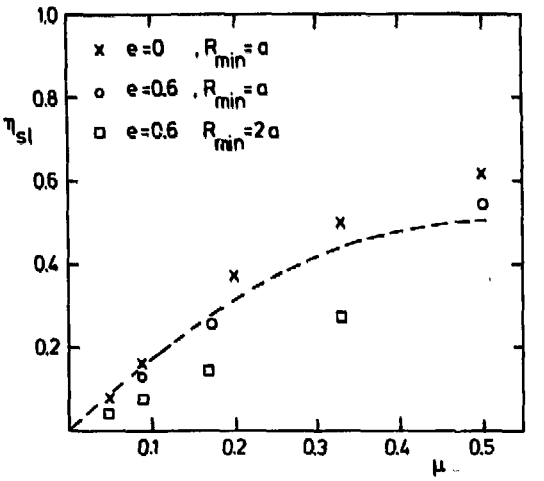


Figure 1. Slingshot efficiency, defined by equation (5) as a function of reduced mass of massive binary. The dashed line is the approximation given by (6).

In Figure 1 the slingshot efficiency is shown as a function of μ, the reduced mass of the binary. For $\mu = \frac{1}{2}$ ($e = 0$) we find $\eta_{sl} \approx 0.6$ in good agreement with the value ~ 0.5 found by Hills and Fullerton. Note that the experimental results are reasonably approximated by the relation

$$\eta_{sl} = 2 \frac{\mu}{m+M} . \qquad (6)$$

We can understand this result assuming that the test particles interact either with m, with probability P_m, or with M with probability P_M, and that they fly off to infinity with the centre of mass (c.o.m.) velocity of m or M respectively, times some constant c. We then find

$$\langle \Delta E \rangle = \left[P_m \left(\frac{M}{m+M}\right)^2 + P_M \left(\frac{m}{m+M}\right)^2 \right] c \, \tfrac{1}{2} m_* V_{orb}^2 .$$

Assuming further that P_m is proportional to m^2 (as in dynamical friction) and to the c.o.m. velocity of m, we find $\eta_{sl} = c\,\mu/(m+M)$. The numerical result found here implies that the binary evolution is independent of the mass-ratio m/M for m << M. Note that this differs from the relation $\Delta E/E = m_* m/M^2$ adopted by Begelman et al. (1980) to describe the evolution of a massive hard binary.

We have also measured the exchange of angular momentum between the binary and the test particles. The angular momentum of the binary is given by

$$L^2 = \mu \frac{GmM}{2E}\,(1-e^2).$$

In the simulations with e = 0 we find $\left\langle \frac{\Delta L/L}{\Delta E/E} \right\rangle = 0.5$, which implies that the eccentricity will not change. For e = 0.6, however, we find a value of 0.7 ± 0.2 and the binary eccentricity may increase. More accurate numerical simulations are required to confirm this.

2.3 Further evolution of the binary.

We have found that the dynamical evolution of the binary inside the cusp radius r_h is independent of the mass ratio m/M for m << M, in contrast to the region where the dynamical frictior formula applies. The binary then interacts only with stars passing through its orbital radius r_b. The total mass of these 'loss-cone' stars is given by

$$M_{lc}(r_b) = 4\pi \int_{r_b}^{\infty} \rho(r)\, r^2\, \theta^2\, dr \qquad (7)$$

where θ^2 is the fraction of such stars at a distance r from the binary. In the case of an isotropic velocity distribution θ is given by

$$\theta = \begin{cases} (r_b r_h)^{\frac{1}{2}}/r & r > r_h \\ (r_b/r)^{\frac{1}{2}} & r < r_h \end{cases} \qquad (8)$$

Frank and Rees, (1976). The domain of integration may consist of three parts: a cusp region for $r < r_h$ with $\gamma_1 = -7/4$ (Bahcall and Wolf, 1976; Lightman and Shapiro, 1977; Cohn and Kulsrud, 1978), a core for $r_h < r < r_c$ with $\gamma = 0$ and an outer region with $-3 < \gamma_2 < -2$. We find

$$M_{lc}\ (r_b) = (\gamma_1 + 3)\ M_{cusp}\ \frac{r_b}{r_h}\left[\frac{1}{\gamma_1+2} + \frac{r_c}{r_h}\ (1 - \frac{1}{\gamma_2+1})\right]. \qquad (9)$$

M_{cusp}, the total mass of the stars within r_h, is given by

$$M_{cusp} \simeq \frac{4\pi}{\gamma_1+3}\ \rho_h\ r_h^3$$

where ρ_h is the density at r_h. Using $\rho_h = \rho_c$ and $r_c = \frac{GM_c}{3\sigma^2}$ we can write

$$M_{cusp} = \frac{81}{\gamma_1+3}\ M^3/M_c^2 \qquad (10)$$

where M_c is the core mass.

The dynamical evolution time of the binary for $r_{lc} < r_b < r_h$ is approximately given by

$$t(r_b) = t_{dyn}\ (r_h)\left[M_{lc}(r_b)\eta_{sl}/\mu\right]^{-1} \simeq t_{dyn}(r_h)r_b/r_h. \qquad (11)$$

From the definition of η_{sl} we see that the binary orbit stops shrinking at a binary separation $r_b = r_{lc}$, defined by

$$M_{lc}\ (r_{lc})\ \eta_{sl}/\mu = 1. \qquad (12)$$

This yields with (9) and (6)

$$\frac{r_{lc}}{r_h} = \frac{1}{81}\ \frac{\mu M_c^2}{\eta_{sl} M^3}\left[\frac{1}{\gamma_1+2} + \frac{1}{3}\ \frac{M_c}{M}\ (1 - \frac{1}{\gamma_2+1})\right]^{-1}. \qquad (13)$$

Clearly it is difficult to bring the two holes close together via three-body scattering if they are embedded in a large galactic core. However, if this is not the case we may use $M_{cusp} = M_c$. From (10) we see that $M_{cusp} \simeq 4\ M$ and equation (13) then yields, with $\gamma_1 = -7/4$, $\gamma_2 = -2$ and $\eta_{sl} = 2\mu/M$

$$r_{lc}/r_h \simeq 0.015. \qquad (14)$$

The minimum separation that the binary may reach via three-body interactions with stars passing through the binary orbit may be smaller than

$\sim r_{lc}$ if only a fraction f of these stars interacts with the smaller binary component, acquires a velocity kick $\Delta V \sim V_{orb}$ and becomes unbound (to the binary), while the binding energy of the other stars is hardly affected. These stars may pass several times through the binary orbit until they are also flung out of the stellar system bound to the binary with velocity $\sim V_{orb}$. The minimum binary separation will then be smaller than r_{lc} by about a factor f. The argument below equation (6) suggests that $f \sim m/M$. In that case the minimum separation of the binary will be

$$\frac{m}{M} r_{lc} \lesssim r_{b,min} \lesssim r_{lc}.$$

Gravitational radiation ensures merging of the black holes within a time Δt when they reach a separation

$$\frac{r_{gr}}{r_h} = 0.01 \left[\frac{M}{10^7\ M_\odot}\right]^{-1/4} \left(\frac{m}{M}\right)^{1/4} \left[\frac{\sigma}{200\ \text{km sec}^{-1}}\right]^2 \left[\frac{\Delta t}{10^{10}\ \text{yr}}\right]^{1/4} f(e), \quad (15)$$

where f(e) is a function of the eccentricity of the orbit, which is of order unity for $e \simeq 0$ (Wagoner, 1975). Comparing (15) with (13) we conclude that merging of galaxies is likely to produce long-lived binary black holes if these holes are embedded in large galactic cores. If the density profile is not flat outside the cusp radius r_h, the holes may reach a separation where gravitational radiation will ensure merging within a Hubble time.

3. Fuelling the Black Hole.

3.1 Stationary fuelling of the black hole.

We will first discuss briefly the fuelling of a black hole in the nucleus of a galaxy via tidal disruption and via star-star collisions, when the central region is undisturbed by other galaxies. We will use our galaxy as an example.

The stellar density at distance of $\sim$ 1 pc from the centre of our galaxy is $\sim 10^6\ M_\odot pc^{-3}$ and the velocity dispersion of the stars $\sim$ 120 km sec^{-1} (Oort, 1977). Infrared observations of the galactic nucleus indicate

the presence of gas at $r \sim 0.4$ pc with velocity dispersion $\sim 200\,\mathrm{km\,sec^{-1}}$ (Wollman et al., 1977). This is consistent with the presence of a black hole of mass $5.10^6\ M_\odot$.

Tidal disruption of solar-type stars occurs at

$$r_T/r_h \simeq 10^{-5}\, M^{-2/3}\, \sigma^2. \qquad (16)$$

The gas produced at r_T is strongly bound to the black hole and a large fraction may be swallowed by the hole. Note that the Schwartzschild radius of the hole is $r_S/r_h = 5.10^{-7}\,\sigma^2$ and that stars may be swallowed whole when $M \gtrsim 10^9\ M_\odot$ (Hills, 1975). At a large distance r from the hole the flux of stars moving in orbits that will bring them within r_T is

$$F(r) = 2.10^3\, \rho(r)\, r^2\, \sigma(r)\, \pi\, \theta_{lc}^2\,(r) \quad M_\odot\ yr^{-1}$$

where θ_{lc} is defined by (8) with r_b replaced by r_T. Using

$$\sigma(r) = \begin{cases} \sigma, & r > r_h \\ \sigma(r/r_h)^{-\frac{1}{2}}, & r < r_h \end{cases}$$

we find

$$\frac{F(r)}{M_\odot\ yr^{-1}} = 0.06\, \rho_h\, \sigma^{-1}\, M^{4/3} \begin{cases} (r/r_h)^{\gamma_2}, & r > r_h \\ (r/r_h)^{\gamma_1+\frac{1}{2}}, & r < r_h \end{cases} \qquad (17)$$

where $\rho_h = \rho(r_h)$. Within a critical radius r_{crit}, stars cannot be scattered in or out of loss-cone orbits by interactions with other stars in one dynamical time scale and these orbits become depleted. Equation (17) is valid only outside r_{crit} where the velocity distribution is isotropic. The largest contribution to the tidal disruption rate comes from the region near r_{crit} and the tidal disruption rate is about $F(r_{crit})/7$ (Frank, 1978). Following the discussion by Frank and Rees (1976) and

---

✲ Note

Here and throughout the rest of the paper radii, masses, densities and velocities will usually be expressed in 1 pc, $10^7\ M_\odot$, $10^7\ M_\odot\ pc^{-3}$ and 200 km sec^{-1} respectively.

using

$$\rho_h \approx M^{-2}\ \sigma^6 \tag{18}$$

which is valid for $r_c = r_h$, we find

$$\left.\begin{array}{l} \text{for } r_{crit} > r_h,\ (r_{crit}/r_h)^{\frac{3+\gamma_2}{2}} \\ \text{for } r_{crit} < r_h,\ (r_{crit}/r_h)^{\frac{4+\gamma_1}{2}} \end{array}\right\} = 2.3\ \rho_h^{-1/12}\ \sigma^{3/2}.$$

In our galaxy $r_{crit}/r_h \approx 1.5$ and the tidal disruption rate is about $5.10^{-4}\ M_\odot\ yr^{-1}$.

The inner edge of the cusp around black holes is set by star-star collisions. Within a radius r_{coll}, defined by

$$r_{coll}/r_h \approx 0.2\ \sigma^2, \tag{19}$$

solar-type stars cannot be deflected over large angles without colliding. The cusp only extends over the region $r_{coll} < r < r_h$. Most collisions occur at r_{coll} which is much larger than r_T. The fate of the gas liberated in collisions with other stars may therefore be different from that of stars disrupted by the tidal force of the black hole. The collision rate in the cusp is about (cf. Van Bueren, 1978)

$$C \approx 5.10^{-3}\ M^3\ \rho_h^2\ \ \sigma^{-7}\ yr^{-1}. \tag{20}$$

In our galaxy this is about $10^{-3}\ yr^{-1}$, a value comparable with the tidal disruption rate. A fuelling rate F of a black hole yields a luminosity

$$L \approx 6.10^{45}\ \text{erg sec}^{-1}\left[\frac{F}{M_\odot\ yr^{-1}}\right]\left[\frac{\eta}{0.1}\right]$$

where η is the fraction of the rest mass energy that is radiated during accretion of the mass by the black hole. This efficiency factor may vary from 0.057 to 0.42 depending on the angular momentum of the black hole (Bardeen, 1970). A tidal disruption rate of $5.10^{-4}\ M_\odot\ yr^{-1}$ yields a

luminosity of 3.10^{42} $(\eta/0.1)$ erg sec^{-1}. The limits to the present energy output of the galactic nucleus are much lower, but there is evidence for recurrent explosions from the nucleus yielding an average energy output of $\sim 10^{42}$ erg sec^{-1} (van Bueren, 1978).

It is interesting to note that a stationary fuelling rate of 5.10^{-4} $M_\odot$ yr^{-1} is consistent with the growth of a 5.10^6 $M_\odot$ black hole in one Hubble time. If the growth time of black holes in galactic nuclei is about one Hubble time we can derive a relation between the central velocity dispersion of stars and the mass of a black hole in the nucleus of a galaxy.

Using (18) we find from the tidal disruption rate

$$M = \begin{cases} 3.5\ \sigma^{2.1}, & r_{crit} > r_h \\ \sigma^{3/7}, & r_{crit} < r_h \end{cases}$$

and from the collision rate

$$M \simeq 5\ \sigma^{5/2}.$$

Collisons may dominate the growth for $\sigma > 1/2$, just before r_{crit} becomes smaller than r_h. The mass given by the relations $M = 3.5\ \sigma^{2.1}$ for $\sigma < 1/2$ and $M = 5\ \sigma^{2.5}$ for $\sigma > 1/2$ may be regarded as an upper limit to the mass of a black hole in a galactic nucleus if tidal disruption of stars and star-star collisions are the dominant fuelling mechanisms.

3.2 Enhanced tidal disruption rate during mergers.

Stars within a distance r_{crit} of a black hole in a galactic nucleus may be scattered into loss-cone orbits by the tidal force of an intruding galaxy, leading to an increase of the tidal disruption rate. The enhanced disruption rate will then be determined by the effective critical radius q, at which stars can be scattered into loss-cone orbits within one (stellar) orbital period. The perturbation of a stellar orbit by the intruder can be calculated analytically if the stellar orbital period is much shorter than the orbital period of the merging galaxies. (cf. Heggie, 1975.) A mean value of q ($<<$ r, the distance between the centres of the merging galaxies) could then be obtained by integrating

over all stellar orbits. However, if the orbital periods are comparable we have to deal with the full complexity of the three-body problem for which there are no general solutions. We will give here a very simple, but crude estimate for q, ignoring its dependence on the orientation of the stellar orbits with respect to the direction of the perturber and neglecting the motion of the perturber. Our basic assumption is that the scattering of stars by the perturbing mass can be treated as a random process analogous to two-body relaxation.

In order to scatter stars into loss-cone orbits we need a mean velocity change in one dynamical time of

$$\Delta V = h_T/q$$

where h_T, the angular momentum per unit mass of stars passing the black hole near r_T, is given by

$$h_T = \sigma r_h^{\frac{1}{2}} r_T^{\frac{1}{2}}.$$

The dynamical time of the stars at q is given by

$$t_{dyn} \simeq \begin{cases} q/\sigma & , \quad q > r_h \\ (q/r_h)^{\frac{1}{2}}\, q/\sigma, & q < r_h \end{cases} \tag{21}$$

and the required acceleration is

$$\Delta V/t_{dyn} = \begin{cases} \sigma^2\, r_h^{\frac{1}{2}}\, r_T^{\frac{1}{2}}\, q^{-2} & , \quad q > r_h \\ \sigma^2\, r_h\, r_T^{\frac{1}{2}}\, q^{-5/2}, & q < r_h \end{cases}. \tag{22}$$

The tidal acceleration of a star at q due to a mass m_g at r is

$$\frac{\Delta V}{\Delta t}_{tidal} \simeq 2\, G\, m_g\, q/r^3.$$

Using $G\, M_g(r) = r\, \sigma^2$ we find

$$\frac{\Delta V}{\Delta t}_{tidal} = \begin{cases} \sigma^2\, f\, q/r^2 & , \quad r > r_h \\ \sigma^2\, f\, r_h/r^3 & , \quad r < r_h \end{cases} \tag{23}$$

where f is $m_g/M_g(r)$. The tidal acceleration equals the required acceleration for

$$q = \begin{cases} 0.12\ r^{2/3}\ M^{2/9}\ \sigma^{-1/3}\ f^{-1/3} & q,r > r_h \\ 0.16\ r^{4/7}\ M^{1/3}\ \sigma^{-4/7}\ f^{-2/7} & q < r_h < r \\ 0.16\ r^{6/7}\ M^{1/21}\ f^{-2/7} & q,r < r_h \end{cases} \qquad (24)$$

where M is again the mass of the black hole at the centre of the most massive galaxy. Calculating F(q(r)) from (17) and (24) and substituting $\gamma_1 = -7/4$ we find for the tidal disruption rate as a function of the distance of the intruding galaxy from the centre of the larger galaxy

$$\frac{F(r)}{M_\odot\ yr^{-1}} = \begin{cases} 6.10^{-2}\ (0.12)^\gamma\ r^{2.\gamma/3}\ \rho_h\ M^{(12-7\gamma)/9} \times \\ \qquad \times \sigma^{(5\gamma-3)/3}\ f^{-\gamma/3}, & q,r > r_h \\ 0.6\ r^{-5/7}\ \rho_h\ M^{13/6}\ \sigma^{-78/28}\ f^{5/14}, & q < r_h < r \\ 0.6\ r^{-15/14}\ \rho_h\ M^{53/21}\ \sigma^{-7/2}\ f^{5/14}, & q,r < r_h . \end{cases} \qquad (25)$$

This relation is valid only for $q > r_{coll}$. The precise behaviour of the density profile at $r < r_{coll}$ is not known. However, it seems plausible that the collision time does not increase with decreasing r within r_{coll} implying $\gamma \lesssim -1/2$ in that region. The fuelling rate then reaches a maximum value determined by the conditions at this inner edge of the cusp

$$F_{max} = 0.45\ \rho_h\ \sigma^{-7/2}\ M^{4/3} \quad M_\odot\ yr^{-1}.$$

Note that this value is not very sensitive to M if (18) and the relations at the end of section 3.1 are valid.

We will now discuss this result for some particular cases. We will assume that both merging galaxies have a power law density distribution with $\gamma_2 = -2$, implying that f is a constant.

Case 1. A merger of a small galaxy with our galaxy assumed to contain a central black hole of $5.10^6\ M_\odot$. We use the same galactic parameters as in the previous section ($\rho_h = 0.25, \sigma = 0.6$). In Figure 2 the tidal disruption rate is given for a mass ratio $f = 10^{-2}$. At a separation distance of the merging galaxies of $360\ f^{\frac{1}{2}}$ pc the tidal disruption rate overtakes the stationary rate of $\sim 5.10^{-4}\ M_\odot\ yr^{-1}$, calculated in the

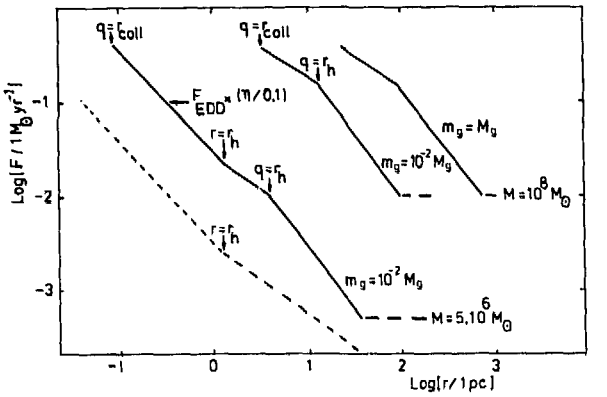


Figure 2. Tidal disruption rate of a central black hole during a merger of a galaxy with a smaller companion. The mass-ratio of the two galaxies is assumed constant. The stellar density and velocity dispersion at the cusp radius r_h are $2.5\times10^6 M_\odot pc^{-3}$ and 120 km sec^{-1} respectively for $M = 5.10^6 M_\odot$ (case 1) and $1.6\times10^6 M_\odot$ pc^{-3} and 300 km sec^{-1} for $M = 10^8 M_\odot$ (case 2). The dashed line is the tidal disruption rate near a secondary black hole of mass $5.10^5 M_\odot$ assumed present in the smaller galaxy in case 1 (see section 3.3).

previous section. The dynamical friction time, given by

$$t_{df} = 10^4 \; r \; \sigma^{-1} \; f^{-1} \; yr \qquad (26)$$

is then about $6.10^6 \; f^{-\frac{1}{2}}$ yr. At $r = 38 \; f^{\frac{1}{2}}$ pc ($F = 10^{-2} \; M_\odot \; yr^{-1}$, $t_{df} = 6.10^5 \; f^{-\frac{1}{2}}$ yr) the value of q reaches that of the cusp radius and the disruption rate subsequently increases less rapidly with decreasing r. The total number of stars disrupted by the black hole in this first stage is $\sim 10^4 \; f^{-\frac{1}{2}}$, only about a factor three higher than in the unperturbed case. F(r) changes its slope again at $r = r_h$ ($F(r_h) = 0.11 \; f^{5/14} \; M_\odot \; yr^{-1}$ $t_{df} = 2.10^4 \; f^{-1} yr$). In the present case this occurs when $q/r_h = 0.15 \; f^{-2/7}$ before the tidal disruption rate reaches its maximum value at $q = r_{coll} = 0.07 \; r_h$. During this stage about $10^4 \; f^{-9/14}$ stars are disrupted. The binary evolution time is given by (11) for $r < r_h$. The disruption rate rises above the Eddington rate

$$F_{EDD} = 0.2 \; M \; (0.1/\eta) \; M_\odot \; yr^{-1}$$

and reaches the value 0.4 $M_\odot$ yr^{-1} when $q = r_{coll}$ at $r = 0.4\ f^{1/3}$. The binary evolution time is then $3.10^4\ f^{-1/3}$ yr. The stellar debris produced in this stage can keep the black hole radiating at the Eddington luminosity for $10^5\ (\eta/0.1)\ f^{-1/3}$ yr. If the fuelling rate stays at a level of 0.4 $M_\odot$ yr^{-1} until $r = r_{coll}$ this is about 10^6 yr. For this period the merger may enhance the fuelling rate by more than two orders of magnitude.

Case 2. Merging of two galaxies, one of which containing a massive black hole of mass $10^8\ M_\odot$. We have chosen a velocity dispersion at r_h of 300 km sec^{-1} and $\rho_h = 1.6\ 10^6\ M_\odot\ pc^{-3}$ (see equation 18) yielding a black hole growth time via accretion of gas produced in star-star collisions of $\sim 10^{10}$ yr. Now q becomes equal to r_{coll} at $r = 34\ f^{\frac{1}{2}}$ pc which is larger than $r_h = 4.4$ pc for $f > 0.02$. The total number of stars disrupted during the whole period of enhanced activity, which lasts $\sim 10^7\ f^{-\frac{1}{2}}$ yr, is $10^6\ f^{-\frac{1}{2}}$, about 10 times more than in the unperturbed case. Most of these stars are disrupted in the second evolutionary stage ($r_{coll} < q < r_h$), where a maximum luminosity of $\sim 3.10^{45}$ erg sec^{-1} is obtained during $\sim 10^6\ f^{-\frac{1}{2}}$ yr.

We conclude that under conditions similar to those in the nucleus of our galaxy, merging of galaxies may yield luminosities of $\sim 3.10^{45} \times (\eta/0.1)$ erg sec^{-1}. Small galaxies ($f \simeq 10^{-2}$) are probably most effective in causing activity in galaxies (see section 4). The total energy production during a merger with such a small companion is of the order $10^{59} \left[\frac{M}{10^7\ M_\odot}\right] (\eta/0.1)$ erg in $\sim 10^7$ yr.

3.3 The secondary component.

If the smaller infalling galaxy also contains a massive black hole the tidal disruption rate of stars near this secondary black hole may be enhanced as well. The tidal disruption rate is then also given by equation (25) with $f \simeq 1$, if the parameters of the larger galaxy are replaced by those of the smaller one. The dotted line in Figure 2 gives the tidal disruption rate of the secondary black hole in case 1 of the preceding section. The velocity dispersion in the smaller galaxy is 40 km sec^{-1}, the black hole mass $5.10^5\ M_\odot$. The stellar density at its cusp

radius is $3.5\ 10^5\ M_\odot\ pc^{-3}$. Note that the mass-ratio of the two galaxies is not a constant anymore. At a separation distance $r \sim r_h$ the mass-ratio changes from 10^{-2} to 10^{-1}, the mass-ratio of the black holes. This will probably not effect the dynamical evolution since inside the cusp radius the binary evolution is independent of the mass ratio. The only difference may be that the primary black hole reaches F_{max} at a slightly larger separation. The important conclusion we can draw from the dashed line in Figure 2 is that the fuelling mechanism discussed in this paper probably implies the existence of double sources of energy with small separations.

3.4 Precession of beams.

The observed large scale symmetry in many extended radio sources may be explained as due to precession of radio beams or jets through which energy is being transported from the nucleus to the outer lobes of the source (e.g. Ekers et al., 1978). Such beams will be aligned with the rotation axis of a hole, presumed to be in the galactic nucleus, even when the angular momentum of the matter feeding the hole has a different orientation (Rees, 1978). If a binary black hole is formed during a merger the spin axes of the holes will generally precess around their orbital angular momentum vector, the precession period being given by

$$t_{prec} \simeq 3.10^{10}\ r^{5/2}\ (M/m)\ M^{-3/2}\ yr \qquad (27)$$

(Begelman et al., 1980). From the discussion of the binary evolution we conclude that the probability to observe a binary with some separation r has a minimum at $r = r_h$, increases with $t_{ev} \propto r_h/r$ for $r < r_h$ (equation 11), and goes rapidly to zero at $r \lesssim 10^{-3}$ pc when energy loss due to gravitational radiation becomes important. For $r > r_h$ the precession period is larger than a Hubble time. The most probable binary separations for $r < r_h$ are $10^{-3} \lesssim r/r_h \lesssim 10^{-2}$ yielding precession periods of $\sim 10^3$ to $\sim 10^7$ yr. The time spent at these separation distances is probably larger than the evolution time given by equation 11 due to depletion of stars on orbits through the binary orbit. The maximum probability at $r < r_h$ may be estimated from $t_{ev} \simeq t_{gr}$, where $t_{ev} \simeq 10^4\ M^2\ \sigma^{-5}\ r^{-1}$ yr and t_{gr}, the time scale for loss of orbital energy due

to gravitational radiation is given by

$$t_{gr} = 2.4 \times 10^{18} \, (M/m) \, M^{-3} \, r^{4} \text{ yr.} \qquad (28)$$

This occurs at $r = 1.3\times10^{-3} \, (m/M)^{1/5} \, M \, \sigma^{-1}$ pc, yielding a most probable precession period of $2.10^{3} \, (M/m)^{\frac{1}{2}} \, M \, \sigma^{-5/2}$ yr. Begelman et al. (1980) derive a precession period of $\sim$ 500 yr from the observed curvature of the jet in 3C 273 on a scale of 0.01". This implies a rather short gravitational radiation time scale of a few million years. The probability that the presence of a secondary black hole in this nearby QSO is accidental seems very small. If the explanation of the jet curvature in 3C 273 proposed by Begelman et al. is right, then the short lifetime of the binary suggests that its occurrence is causally related to the QSO phenomenon.

4. Total Luminosity Function of Active Galaxies.

The maximum total luminosity of active galactic nuclei at the present epoch is about 10^{46} erg sec^{-1}. Radio observations of active galaxies indicate that the total energy production in active nuclei can be $\sim 10^{60}$ erg in 10^{6} to ^{8} yr. These values are compatible with the activity produced during a merger of a galaxy with a smaller companion (section 3.2) and it seems therefore worthwhile to investigate what fraction of galaxies is expected to be active in the merger model, and to make a comparison with observations.

4.1 The model.

The merging rate of galaxies can be estimated from observations as well as from numerical experiments. Most galaxies have a smaller companion ($m_g/M_g \sim 0.1$) within a few galactic radii (Holmberg, 1969). Assuming that galaxies have dark halos extending to $\sim$ 100 kpc, the dynamical friction time for these companions is about 10^{10} yr (see also Tremaine, 1976) yielding a growth rate for the larger galaxies of about $M_g \, (10^{11} \text{ yr})^{-1}$. The number of neighbours (of comparable size) within a distance r of a galaxy can also be calculated from the two-point

correlation function (Peebles, 1980, Gott and Turner, 1979). From this function we find that the mean number of neighbours within a distance r ($10\ \mathrm{kpc} \lesssim r \lesssim 10\ \mathrm{Mpc}$) is about $4\pi\bar{n}\ 70\ (r/1\ \mathrm{Mpc})^{1.2}$ (we adopt H_o = 50 km sec^{-1} Mpc^{-1}) where $\bar{n}$, the mean number density of galaxies brighter than L^*, the break in the luminosity function, is about $2.10^{-3}\ \mathrm{Mpc}^{-3}$. The relative velocities of close neighbours is a few hundred km sec^{-1}. Most of these binaries will therefore experience strong dynamical friction and will merge within a few galactic rotation periods if galaxies have extended halos. This also yields a growth rate for bright galaxies of about $M_g\ (10^{11}\ \mathrm{yr})^{-1}$. A similar merging rate was deduced by Toomre (1977) from the fraction of galaxies that are observed to be tidally distorted due to interaction with a nearby neighbour. The merging rate at the present epoch derived from observations is in good agreement with the value found in cosmological simulations (Roos, 1981, Paper I).

We will now estimate the fraction of galaxies $\phi(L)$ having bolometric luminosities L (equal to $6.10^{45}\ (\eta/0.1)$ erg sec^{-1} times the fuelling rate) in the interval $[L,\ 2.5\times L]$. We will do this for galaxies of mass M_g containing a central black hole of mass M. We assume that the galaxies have power law density profiles with $\gamma = -2$ and stellar velocity dispersion σ. Assuming that the infall rate is proportional to mass (paper I and II) and that the mass function of infalling galaxies is about proportional to m_g^{-1} for small m_g (Schechter, 1976) we find for the mean time between subsequent mergers of a galaxy of mass M_g with smaller galaxies of mass m_g $\Delta t = 10^{11}\ m_g/M_g$ yr. The fraction $\phi'(r)$ of galaxies having a companion at a distance in the interval $[r,\ r+dr]$ is

$$\phi'(r) = \frac{1}{\Delta t}\frac{dt}{dr} \tag{29}$$

for r smaller than some distance r_{sat}, defined by $\int_0^{r_{sat}} \phi'(r)dr = 1$, where "saturation" of $\phi'(r)$ occurs (see next section). As discussed in section 2 the smaller galaxy will spiral inward with constant velocity ($r > r_h$)

$$\frac{dr}{dt} = 10^{-4}\ f\ \sigma\ \mathrm{pc\ yr}^{-1},$$

if the two galaxies have similar density profiles. Knowing $\phi'(r)$ we can calculate the fraction of galaxies having a fuelling rate F in the

interval $[F, 2.5 \times F]$ from

$$\phi(F) = \phi'(r) \frac{F}{2.5} \frac{dr}{dF}.$$

This yields with (25) and (18)

$$\phi(F) \simeq \begin{cases} 10^{-7} F^{-3/4} M^{2/3} \sigma^{1/4} f^{-3/2}, & F < F^* \\ 0.2 \times 10^{-7} F^{-7/5} M^{5/42} \sigma^{3.5} f^{-3/2}, & F > F^* \end{cases} \quad (30)$$

The break in this function occurs at $F^* = 0.06\ \rho_h\ M^{4/3}\ \sigma^{-1}\ M_\odot\ yr^{-1}$ when $q = r_h$. As dynamical friction is taken over by three-body scattering at $r \lesssim r_h$ the dynamical evolution speeds up (if $f < 1$) and $\phi(F)$ probably decreases strongly upon extrapolation of (30) to larger F. Note that the part of $\phi(F)$ with an index $-7/5$ does not extend to large F. The maximum F for this part of $\phi(F)$ is determined by the condition $r = r_h$ (as in case 1 of the previous section) yielding $F = 0.59\ \rho_h\ M^{1.45}\ \sigma^{-1.36}\ f^{5/14}\ M_\odot\ yr^{-1}$, or $q = r_{coll}$, yielding

$$F = 0.45\ \rho_h\ M^{4/3}\ \sigma^{-7/2}\ M_\odot\ yr^{-1}.$$

The largest contribution to $\phi(F)$ comes from mergers with small companions. Firstly, because the dynamical friction time at some fuelling level is longer, and secondly because these galaxies are more abundant. However, the galaxies that are being swallowed may not be smaller than $\sim 10^{-2}\ M_g$ since the dynamical friction time then exceeds the Hubble time. Converting fuelling rates F to luminosities L and taking $f = 10^{-2}$, the fraction of galaxies having luminosity within $[L, 2.5 \times L]$ due to infall of small galaxies is given by

$$\phi(L) = \begin{cases} 2.10^{-3}\ L_{44}^{-3/4}\ (\eta/0.1)^{3/4}\ M^{2/3}\ \sigma^{1/4}, & L < L^* \\ 6.10^{-3}\ L_{44}^{-7/5}\ (\eta/0.1)^{7/5}\ M^{5/42}\ \sigma^{3.5}, & L > L^* \end{cases} \quad (31)$$

where $L_{44} = L/10^{44}$ erg sec^{-1} and $L_{44}^* = 5\ M^{-2/3}\ \sigma^5\ (\eta/0.1)$. As in (30) the region with $\phi(L) \propto L^{-7/5}$ extends over a small range in L to $L_{44} = 30(\eta/0.1)\ M^{-0.55}\ \sigma^{4.64}\ f^{5/14}$ for $r = r_h$ ($q > r_{coll}$), or to $L_{44} = 30(\eta/0.1)\ M^{-2/3}\ \sigma^{2.5}$ for $q = r_{coll}$ ($r > r_h$). Note that the choice of f introduces a considerable uncertainty in the absolute value of $\phi(L)$.

The luminosity function of active galaxies could in principle be obtained from (31) by performing an integration over the product of $\phi(L)$

and the mass function of central black holes in galactic nuclei. This requires information about the relation between galactic luminosity (or central velocity dispersion) and central black hole mass. For the moment we are merely interested to see whether equation (31) is consistent with luminosity functions of active galactic nuclei derived from observations. We will use as a first approximation for the theoretical function of active nuclei having total luminosity L in the interval $[L, 2.5 \times L]$

$$\rho(L) = n^* \, \phi(L, M^*, \sigma^*) \qquad (32)$$

where n^* is the number density of bright galaxies above the break of the galaxy luminosity function, and M^* and σ^* are the mass of the central black hole and the velocity dispersion in these galaxies. In Figure 3 $\rho(L)$ is drawn for $n^* = 2.10^6$ Gpc^{-3} (Felten, 1977). $M^* = 10^7 \, M_\odot$ and $\sigma^* = 150$ km sec^{-1}. Note that a proper calculation of $\rho(L)$ will yield a luminosity function extending over a larger range in luminosity.

4.2 The observations.

Our next task is to compare this result with available observational data on the space density of active galaxies at the present epoch. Luminosity functions for active galaxies have been determined for radio, optical and X-ray wavebands. However, the determination of the bolometric (total) luminosity function requires information on the spectral distribution of the radiation emitted by active galactic nuclei (and/or QSO's). For some individual sources like 3C 273 the spectrum is reasonably well known from 10^9 - 10^{20} Hz (Ulrich, 1981), but this is still an exception. Nevertheless, the presently available observational data indicate that the luminosities at different wavebands are correlated, and that the global spectral distribution of the radiation emitted in the range 10^{13} - 10^{18} Hz can be characterized by a power law with a spectral index which is similar for different objects. The energy emitted at optical wavelengths seems a good indicator of the total energy emitted by Seyferts and QSO's. The mean radio-to-optical spectral index α_{RO} of radio-loud QSO's is about 0.6, implying $L_{opt}/L_{radio} = 10^{2.4}$ and the optical-to-X-ray index α_{OX} is about 1.25 (Ku et al. 1980; Zomorani et al., 1980) implying $L_{opt}/L_X = 10^{3/4}$, where L_{radio},

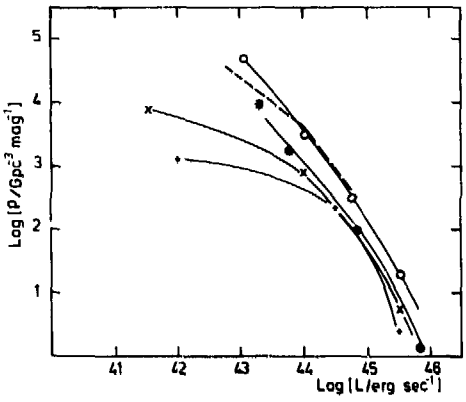


Figure 3. Luminosity functions of active galaxies. Luminosities are converted to total luminosities as described in the text (section 4.2).
o X-ray luminosity function of Seyfert galaxies.
● Local, optical luminosity function of QSO's.
** Space density of Seyfert nuclei (optical).*
× Radio luminosity function of E+SO galaxies (1.4 GHz).
+ Radio luminosity function of E+SO cores (5 GHz).
The dashed line gives the estimated luminosity function of active galaxies derived from (31) using the parameters given in the text (section 4.1).

L_{opt} and L_X are the luminosities integrated over one decade in frequency at 10^9, 10^{15} and 10^{18} Hz respectively. The scatter in the values of α is large and in some cases most energy is emitted in the infrared (Rieke and Lebofsky, 1979), or in the gamma-ray domain as in 3C 273 (Swanenburg et al., 1978), but in these cases too the optical luminosity may give a good order of magnitude estimate of the bolometric luminosity. Radio galaxies seem to behave differently. While the energy emitted at radio wavelengths is comparable to that of radio-loud QSO's, the optical luminosity of the nuclei of radio galaxies is in most cases much less than that of QSO's. Still there is some evidence that the energy production rate in these objects may also be larger by about two orders of magnitude than their radio luminosity which is about 10^{43} to 10^{44} erg sec^{-1} for the most luminous sources. The energy contained in the particle fluxes filling the extended radio lobes may well exceed 10^{61} erg in some cases. This would imply a total mean energy production rate of $\sim 10^{45.5}$ erg sec^{-1} if the lifetime of these sources is about 10^8 yr. Therefore we have assumed that also for radio galaxies $L/L_{radio} = 10^{2.4}$. Using

the crude conversion factors given above we now estimate 'bolometric' luminosity functions from observationally determined radio, optical and X-ray luminosity functions of active galaxies at the present epoch. The following luminosity functions (scaled to H_o = 50 km sec^{-1} Mpc^{-1}) were used to obtain the results given in Figure 3.

(i) X-ray luminosity function of Seyfert galaxies from Weedman (1979). $Log(L/erg\ sec^{-1}) = log(L_x/erg\ sec^{-1}) + 3/4$.

(ii) Optical luminosity function of QSO's at the present epoch from Braccesi et al. (1980). The total luminosity was calculated from $log(L/erg\ sec^{-1}) = log(L_{opt}/erg\ sec^{-1}) = 0.4 \times (90 - M_B)$. Note that the determination of this luminosity function depends on the cosmological evolution of the QSO luminosity function.

(iii) Space density of Seyfert nuclei (optical) from Huchra and Sargent (1973). $Log(L/erg\ sec^{-1}) = 0.4 \times (90 - M_B)$. As noted by Braccesi et al. this point agrees well with the low luminosity extrapolation of the optical (local) luminosity function of QSO's.

(iv) Radio luminosity function of E + SO galaxies from Auriemma et al. (1977). Here we have used $log\ (L/erg\ sec^{-1}) = 41 + log\ (P_{1.4}/10^{24.5}\ WHz^{-1}) + 2.4$, where $P_{1.4}$ is the radio power at 1.4 GHz. A large fraction of the radio luminosity of these sources comes from the extended radio structures at large distances from the centre of the parent galaxy. The luminosity from these radio lobes may not be representative for the luminosity of the nucleus. Therefore we also give the radio luminosity function of the cores of E+SO galaxies (determined at 5 GHz) from Fanti and Perola (1977).

The agreement between the observational functions given in Figure 3 is surprisingly good in particular for $L > 10^{44}$ erg sec^{-1}. Not only the slopes but also the absolute values of these functions are very similar. This strongly suggests that the different types of active galaxies such as Seyferts (QSO's) and radio galaxies belong to a single class of objects having a basic spectrum which can indeed be characterized by the values for α used above. The scatter in the relative portions of energy emitted at different wavelengths may then be the result of different physical conditions in galactic nuclei. The fraction of active E + SO galaxies, for instance, is comparable to the fraction of active Spirals (Seyferts). For some reason the relative amount of energy produced by active E + SO galaxies in the optical regime is lower than

that of Seyferts while active Spirals (Seyferts) are underluminous at radio wavelengths, but it seems likely that the activity in both types is caused by the same mechanism.

4.3 Comparison of the model with observations.

The crude model estimate from equation (32) (see Figure 3) was done for a single set of parameters assumed to be typical for the nucleus of a bright galaxy. Nevertheless, the agreement with observed luminosity functions in the range $10^{42.5} \leq L \leq 10^{44.5}$ is already quite good indicating that the model may be consistent with observations.

Equation (31) may be compared more directly with the observed bivariate radio luminosity function of E + SO galaxies, determined by Auriemma et al. (1977), which gives the fraction of E + SO galaxies within a certain (1 magnitude) bin of optical (stellar) luminosity having a certain radio power (at 1.4 GHz). This function has the following properties.

1. It can be described by a single power law for $P > P^* = 10^{25}$ WHz^{-1} $\simeq 10^{42}$ erg sec^{-1} (H_o = 50 km sec^{-1} Mpc^{-1}). The power law index is about −1.4, independent of the optical luminosity L_g of the galaxy.
2. P^* is not very sensitive to L_g.
3. The fraction of galaxies having a certain radio power scales as $L_g^{1.5\pm0.2}$ for $P > P^*$.
4. Below P^* the index of the power law fit to the radio luminosity function varies from ~ -0.7 for the faintest galaxies to ~ -0.1 for the most luminous galaxies in the sample. Auriemma et al. remark that the flattening may be due to some form of saturation since, for the most luminous galaxies, the fraction having radio power larger than P is already close to 1 at $P = 10^{24}$ WHz^{-1}.

These properties resemble some of the properties of $\phi(L)$ given by (31).

1. $\phi(L)$ has a power law index −1.4 for $L > L^*$, where $L^* = 5.10^{44}$ erg sec^{-1}, which would imply a radio luminosity of 10^{42} erg sec^{-1} using the conversion factor discussed in the previous section. A difficulty with the model may be that $\phi(L)$ extends over only a small range in luminosity beyond L^*.
2. Assuming $L_g \propto \sigma^4$ (Faber and Jackson, 1976) we find that $L^* \propto L_g^{0.9}$ to 0.6 for $M \propto \sigma^2$ to 4.

3. For $L > L^*$, $\phi(L)$ scales approximately as σ^4 which goes as L_g.
4. Below L^* the power law index of $\phi(L)$ is -0.75. In the merger model saturation may occur also for the most massive galaxies when the mean time between subsequent mergers is smaller for these galaxies than for the smaller ones. Observations indicate that for the brightest members of rich clusters it may be as short as $\sim 3.10^9$ yr for $f \sim 1/3$ (Hoessel, 1980). For $f = 10^{-2}$ it may then be as short as $\sim 10^8$ yr, implying a time-averaged activity of $\sim 10^{44}$ erg sec^{-1}.

4.4 Cosmological evolution.

So far we have only considered the luminosity function of active galaxies at the present epoch. This function, however, has evolved quite strongly between $z = 0$ and $z = 2.5$ (Schmidt, 1968). The observations seem to yield two global characteristics of the evolution: 1. the strong, steep spectrum radio source population has evolved most rapidly, while the evolution of fainter sources may have been very weak (e.g. Longair, 1978, Wall et al., 1980). 2. This strong evolution may have occurred during a relatively recent epoch ($0.25 \lesssim z \lesssim 1$, Katgert et al. 1979) and may be related to the epoch of cluster formation. These points suggest that the strong evolution may be correlated with the evolution of giant ellipticals in regions where the number density of galaxies is high.

The over-all merging rate in cosmological simulations declines slowly as t^{-1} and the merging rate between $z = 2.5$ and $z = 0$ differs by a factor ~ 6.5. Note, however, that this may be a lower limit to the real over-all evolution as pointed out in paper I, because the calculations are relevant only for galaxies in the field and in poor clusters. In rich clusters the merging rate declines on a dynamical time scale of a few times 10^9 yr by a factor of about one hundred due to (i) the increase in velocity dispersion of the galaxies during collapse of the cluster, and (ii) rapid tidal stripping of galaxy halos during and after collapse of the cluster (paper II). The steepest evolution may therefore be associated with the massive ellipticals that are formed in the central regions of rich clusters. These results imply that at earlier cosmological epochs cluster galaxies form a much more dominant fraction of active galaxies.

5. Discussion and Summary.

We have attempted to link the merging phenomenon and the occurrence of activity in galactic nuclei. It has often been suggested that infall of gas from outside the stellar cusp surrounding a black hole may be responsible for enhanced activity. This may be gas which is released during stellar evolution or gas which is accreted in a merger (Gunn, 1979) or cooling intra-cluster gas sinking into the potential well of the most massive galaxy (Matthews and Bregman, 1978). A serious difficulty with these models is that the fraction of the gas which is finally accreted by the hole is unknown. Large portions of the gas may form stars or be heated in shocks and expelled from the nuclear region, or even from the galaxy (Matthews and Baker, 1971). As an alternative mechanism to fuel the central engine in active nuclei we suggest here tidal disruption near the black hole of stars scattered into loss-cone orbits by the perturbing gravitational field of an intruding galaxy. Fuelling by tidal disruption of stars is attractive because (i) the stars are disrupted very close to the Schwarzschild radius of the black hole and a large fraction of their mass probably accreted by the hole, (ii) the tidal disruption rate can be estimated and predictions can be made both for the stationary case and for the case of galaxy mergers. In this paper we have given a very crude estimate of the enhanced tidal disruption rate during mergers, treating the perturbation of stellar orbits by the intruding galaxy as a random process similar to two-body relaxation. Detailed theoretical and numerical investigations are required to check this approach or to improve upon it. The results should therefore be regarded as tentative.

A vital condition for the model to yield sufficiently high tidal disruption rates during mergers is that the cusp mass should be comparable to the black hole mass. This situation can only occur if the black hole is not embedded in a galactic core with core radius larger than the cusp radius. If our galaxy contains a black hole of $\sim 5.10^6\ M_\odot$ this condition if fulfilled. In other galaxies the situation may be similar since in the cases where a resolved core may have been found it is in most cases not significantly larger than the seeing disk (Schweizer, 1979). Our assumption that central black holes in galactic nuclei are not embedded in cores has some important consequences:

(i) the tidal disruption rate in unperturbed galaxies containing a black hole of $\sim 10^7 M_\odot$ is of the order $\sim 10^{-3} M_\odot\ yr^{-1}$ yielding a time-averaged luminosity of $\sim 10^{42}$ erg sec^{-1}. If the Milky Way is representative for normal galaxies in this respect then this implies a space density of $\sim 10^6\ Gpc^{-3}\ mag^{-1}$ for galaxies having a luminosity of $\sim 10^{42}$ erg sec^{-1}. Note that this is in good agreement with the low luminosity extension of the Seyfert luminosity function given in Figure 3.
(ii) The tidal disruption rate may be considerably enhanced during mergers yielding maximum luminosities, determined by the inner edge of the cusp at r_{coll}, of $\sim 10^{46}$ erg sec^{-1} and a total energy production during mergers with small ($m_g/M_g \approx 10^{-2}$) galaxies of about $10^{59}\ [M/10^7 M_\odot]$ erg in $\sim 10^7$ yr. This is high enough to explain the maximum luminosities of active galaxies at the present epoch. Some QSO's at larger redshifts appear to have luminosities of 10^{48} erg sec^{-1}. It is possible, however, that the luminosity of these objects is overestimated if the emitted radiation is beamed or if the emission of radiation occurs in short bursts. Also, these high luminosities would imply black hole masses of $10^{10} M_\odot$, if they are radiating at the Eddington limit, yielding a cusp radius of ~ 1 kpc, which exceeds observational limits in galaxies at low redshifts.
(iii) The binary black hole, which will be formed if both merging galaxies contain a central black hole, will lose orbital energy due to three-body interactions with stars and the binary orbit may shrink to a size where loss of orbital energy via emission of gravitational waves takes over and assures merging of the holes within a Hubble time.

As pointed out by Begelman et al. (1980), a binary black hole system in active galaxies may manifest itself via its orbital period or via precession of the spin axes of the holes. Short orbital periods of ~ 10 yr and precession periods of $\sim 10^4$ yr may be most likely. The time scale for loss of orbital energy due to emission of gravitational radiation is expected to be about $10^{6\ to\ 7}$ yr. If such a system is present in 3C 273 as proposed by Begelman et al. its short evolution time suggests a causal relation between the occurrence of the binary black hole and the activity in galactic nuclei.

Another phenomenon which may be observable is the activity of the secondary component. A merger of two similar galaxies may yield a double nucleus with a separation of order 1 to 10^2 pc. Shklovskii (1978) has

proposed just such a model for the nucleus of NGC 1275. He argues that the asymmetry of the narrow emission lines and the results of radio interferometric observations (at 1.35 cm) with intercontinental baselines are both consistent with the existence of a central binary system in this galaxy, having a separation of 1 pc and total mass > 1.5 x 10^8 $M_\odot$. However, recent spectroscopic observations of Seyfert galaxies reveal that most asymmetric narrow lines extend further towards short wavelengths, which is most likely due to outflow of the radiating gas from the nucleus (Heckman et al., 1981). Double active nuclei may be found by observations with high spatial resolution, using for instance very long baseline interferometers, or, in the near future, the Space Telescope. Information on the (gas) dynamics of such systems may be obtained from the broad emission lines observed in many active nuclei.

In most cases activity must have been triggered by mergers of galaxies with mass ratio $f \approx 10^{-2}$. Such small companions are not likely to cause significant distortions of the optical appearance of galaxies. The (less frequent) interactions among galaxies having mass-ratio 0.1 to 1, however, can yield significant distortions in spiral galaxies such as rings and tails (Toomre and Toomre, 1972; Lynds and Toomre, 1976). The merging process then takes a few dynamical times and some of the distortions may still be photometrically apparent in active galaxies. Such distortions are most pronounced in spiral galaxies due to their large systematic rotation velocities and due to the presence of gas which will respond to the density perturbations and may trigger bursts of star formation. Many Seyfert galaxies are indeed disturbed (Adams, 1977; Simkin et al. 1980). It seems possible that many of the distortions in these galaxies can be explained if they have recently accreted a smaller companion. Note that the nucleus of this companion has to be at a distance smaller than $\sim$ 100 pc from the centre of the galaxy in order to enhance the tidal disruption rate significantly.

The estimated merging rate of galaxies of mass M_g with companions of mass 10^{-2} M_g is about 1 per 10^9 yr. It is interesting to note that this merging rate may be sufficient to dominate the growth of black holes in galactic nuclei. The central black hole may grow by $\sim$ 10% during a merger, yielding a growth rate of $\sim 10^{-3}\,[M/10^7\,M_\odot]\,M_\odot\,yr^{-1}$, which is comparable to the growth via tidal disruption of stars or star-star collisions in the absence of mergers. Using the above merging rate

we have estimated the fraction $\phi(L)$ of galaxies having total energy outflow from their nucleus L due to merging with smaller companions. We must be careful when comparing this luminosity function with observed luminosity functions. Firstly, because we have made a number of assumptions to arrive at equation 31. Secondly, because the relation between the energy emitted in some waveband with the total energy released near the black hole is not clear. In view of this second point we have compared several observed luminosity functions of active galaxies at the present epoch. The results suggest that there is indeed a single population of active galaxy nuclei having a basic continuum spectrum. The variations in spectral index for different types of active galaxies may then be due to different physical conditions in their nuclei. We find that our predicted luminosity function agrees well with observed luminosity functions of active galaxies at the present epoch. It is also capable of explaining some features of the bivariate radio luminosity function of E + SO galaxies.

The merger model naturally leads to a (slow) cosmological evolution in the over-all number density of active galaxies. The strong evolution of powerful steep spectrum sources may be correlated with the strong decrease of the merging rate during the collapse of rich clusters. The spectra of active nuclei may also evolve during the epoch of cluster formation, as the gas content of galaxies becomes depleted, due to mergers and sweeping. The differences between the spectra of active E and SO galaxies and of active spiral galaxies (Seyferts) could be the result of such a process.

In a previous study (paper I) it was shown that in a hierarchical clustering scenario merging among galaxies during the epoch of cluster formation may well be responsible for an evolution of galaxies along the sequence (Sc-Sb-Sa)-SO-E. In this paper we have considered some of the possible implications of the merging process for models of active galactic nuclei. Assuming conditions in galactic nuclei similar to those in the nucleus of the Milky Way, we have argued that merging with smaller galaxies will enhance the tidal disruption rate of stars near a central black hole and may yield a luminosity which is typical for a Seyfert galaxy. The fraction of galaxies that is estimated to be active at a certain level is consistent with observed luminosity functions of active galaxies. These encouraging results indicate that

enhanced tidal disruption of stars during galaxy mergers may be very important in causing activity in galactic nuclei.

Acknowledgements.

I thank Harry van der Laan and Colin Norman for their comments on the manuscript. I have also benefitted from discussions with Ernst van Groningen, Peter Katgert and George Miley.

References.

Aarseth, S.J. and Fall, S.M., 1980, Ap.J. 236, 43.

Adams, T.F., 1977, Ap.J. Suppl. 33, 19.

Auriemma, C., Perola, G.C., Ekers, R., Fanti, R., Lari, C., Jaffe, W.J., Ulrich, M.H., 1977, Astron. and Astrophys. 57, 41.

Bahcall, J.N. and Wolf, R.A., 1976, Ap.J. 209, 214.

Bardeen, J.M., 1970, Nature 226, 64.

Begelman, M.C., Blandford, R.D. and Rees, M.J., 1980, Nature 287, 307.

Blandford, R.D., 1979, in 'Sources of Gravitational Radiation', p. 199 Ed. L. Smarr, Cambridge University Press.

Braccesi, A., Zitelli, V., Bonoli, F. and Formiggini, L., 1980, Astron. and Astrophys. 85, 80.

Cohn, H. and Kulsrud, R.M., 1978, Ap.J. 226, 1087.

Ekers, R.D., Fanti, R., Lari, C. and Parma, P., 1978, Nature, 276, 588.

Faber, S.M. and Jackson, R.E., 1976, Ap.J. 204, 668.

Fanti, R. and Perola, G.C., 1977, IAU Symposium No. 74 'Radio Astronomy and Cosmology' ed. by D.L. Jauncey (D. Reidel Publishing Company, Dordrecht, Holland).

Felten, J.E., 1977, Astron. J. 82, 861.

Frank, J., 1978, M.N.R.A.S. 184, 87.

Frank, J. and Rees, M.J., 1976, M.N.R.A.S. 176, 633.

Gott, J.R. and Turner, E.L., 1979, Ap.J. Letters, 232, 79.

Gunn, J.E., 1979, in 'Active Galactic Nuclei', p. 213, eds. C. Hazard and S. Mitton, Cambridge University Press.

Hall, G. and Watt, J.M., 1976, eds. 'Modern Numerical Methods for Ordinary Differential Equations' (Clarendon Press).

Hardy, I.M., 1979, M.N.R.A.S. 188, 495.
Heckman, T.M., Miley, G.K., Van Breugel, W.J.M. and Butcher, H.R., 1981, preprint.
Heggie, D.C., 1975, M.N.R.A.S. 173, 729.
Hills, J.G., 1975a, Astron. J. 80, 1075.
Hills, J.G., 1975, Nature 254, 295.
Hills, J.G. and Fullerton, L.W., 1980, Astron. J. 85, 1281.
Hoessel, J.G., 1980, Ap.J. 241, 493.
Holmberg, E., 1969, Ark. Astronomy 5, 305.
Huchra, J. and Sargent, W.L.W., 1973, Ap.J. 186, 433.
Jones, B.J.T. and Efstathiou, G., 1979, M.N.R.A.S. 189, 27.
Katgert, P., de Ruiter, H.R. and Van der Laan, H., 1979, Nature 280, 20.
Ku, W.H.M., Helfland, D.J. and Lucy, L.B., 1980, Nature 288, 323.
Longair, M.S., 1978, in 'Observational Cosmology', p. 125, Eighth Advanced Course of the Swiss Society of Astronomy and Astrophysics, eds. Maeder, A., Martinet, L., and Tammann, G., Geneva Observatory.
Lightman, A.P. and Shapiro, S.L., 1977, Ap.J. 211, 244.
Lynds, R. and Toomre, A., 1976, Ap.J. 209, 382.
Matthews, W.G. and Baker, J.C., 1971, Ap.J. 170, 241.
Matthews, W.G., and Bregman, J.N., 1978, Ap.J. 224, 308.
Miley, G.K., 1980, Ann. Rev. Astron. and Astrophys. 18, 165.
Oort, J.H., 1977, Ann. Rev. of Astron. and Astrophys. 15, 295.
Peebles, P.J.E., 1980, 'The Large-Scale Structure of the Universe' (Princeton University Press, Princeton, New Jersey).
Rees, M.J., 1978, Nature 275, 516.
Rieke, G.H. and Lebosfsky, M.J., 1979, Ann. Rev. of Astron. and Astrophys. 17, 477.
Roos, N., 1981, Astron. and Astrophys. 95, 349 (Paper I).
Roos, N. and Aarseth, S.J., 1981, in preparation (Paper II).
Saslaw, W., Valtonen, M.J. and Aarseth, S.J., 1974, Ap.J. 190, 253.
Schmidt, M., 1968, Ap.J. 151, 393.
Schweizer, F., 1979, Ap.J. 233, 23.
Schweizer, F., 1980, Ap.J. 237, 303.
Shklovskii, I.S., 1978, Sov. Astron. Lett. 4, (6), 266.
Spitzer, L., 1962, 'Physics of Fully Ionized Gases' (New York: Interscience).
Sinkin, S.M., Su, H.J. and Schwartz, M.P., 1980, Ap.J. 237, 404.

Toomre, A. and Toomre, J., 1972, Ap.J. 178, 662.

Toomre, A., 1977, in 'The Evolution of Galaxies and Stellar Populations' p. 401, eds. B.M. Tinsley and R.B. Larson (New Haven: Yale University Observatory).

Tubbs, A.D., 1980, Ap.J. 241, 969.

Ulrich, M.H., 1981, preprint.

Valtonen, M.J., 1975, Mem. R. Astron. Soc. 80, 77.

Van Albada, T.S., 1968, Bull. Astron. Inst. of the Netherlands, 19, 479.

Van Bueren, H.G., 1978, Astron. and Astrophys. 70, 707.

Wagoner, R.V., 1975, Ap.J. Lett. 196, 63.

Wall, J.V., Pearson, T.J. and Longair, M.S., 1980, M.N.R.A.S. 193, 683.

Weedman, D.W., 1979, in 'Highlights of Astronomy', Vol. 5, p. 623, ed. P.A. Wayman (D. Reidel Publishing Company, Dordrecht, Holland).

Wollman, E.R., Geballe, T.R., Lacy, J.H. and Townes, C.H., 1977, Ap.J. Lett. 218, 103.

Young, P.J., 1979, Ap.J. 215, 36.

Zamorani, G., Henry, J.P., Maccacaro, T., Tannenbaum, H., Soltan, A., Avni, J., Liebert, J., Stocke, J., Strittmatter, P.A., Weymann, R.J., Smith, M.G. and Condon, J.J., 1980, Ap.J. in press.

SAMENVATTING

Heldere sterrenstelsels kunnen globaal onderverdeeld worden in drie klassen: spiraalstelsels (S), SO stelsels (SO) en elliptische stelsels (E). S stelsels (zoals onze Melkweg) bestaan uit een sterk afgeplatte roterende schijf welke aanzienlijke hoeveelheden gas en stof bevat. In het centrale deel bevindt zich meestal een sferisch of ellipsvormig lichaam (de "bulge"). Spiraalachtige structuren zijn meestal te onderscheiden in de verdeling van het gas en stof in de schijf en in de verdeling van de jonge, heldere sterren, die daaruit gevormd worden. SO en E stelsels bevatten weinig of geen gas. De sferische component van SO stelsels is meestal groter dan die van S stelsels. E stelsels lijken in veel opzichten op de sferische componenten van S en SO stelsels. Ze zijn echter niet omringd door een afgeplatte schijf.

Een aantal (heldere) sterrenstelsels onttrekt zich aan het bovenstaande classificatieschema. Deze "vreemdsoortige" stelsels komen vaak in groepen van twee of meer voor en vertonen sterke vervormingen. De gebroeders Toomre lieten in 1972 overtuigend zien dat dergelijke vervormingen in veel gevallen het resultaat moesten zijn van getijdekrachten die twee botsende stelsels op elkaar uitoefenen. Bij lage relatieve snelheden zijn botsingen sterk inelastisch. De stelsels zullen elkaar dan binnen korte tijd na de eerste ontmoeting invangen en een nieuw stelsel vormen (dit proces wordt ook wel aangeduid met de term kannibalisme). Toomre en Toomre opperden de mogelijkheid dat E stelsels ontstaan zijn uit "versmeltingen" (mergers) van S stelsels.

De belangstelling voor botsingen tussen sterrenstelsels steeg in de loop van de 70-er jaren toen er steeds meer aanwijzingen kwamen dat sterrenstelsels aanzienlijk groter en zwaarder zijn dan het op fotografische platen zichtbare gedeelte.

Botsingen tussen sterrenstelsels komen het meest voor in clusters. De meeste sterrenstelsels maken deel uit van groepen of clusters welke slechts enkele tot een paar duizend leden bevatten. S stelsels zijn in de meerderheid in groepen en in de buitengebieden van clusters. In de centrale delen van rijke clusters komen overwegend SO en E stelsels voor. Spitzer en Baade hebben in 1951 al laten zien dat botsingen tussen sterrenstelsels in rijke clusters geen belangrijk effect op de sterverdeling in de stelsels hebben, omdat de relatieve

snelheden van de stelsels te hoog zijn. De relatieve snelheden moeten echter veel lager geweest zijn tijdens de vorming van clusters. Onder die omstandigheden hebben botsingen een veel sterker effect. Het onderzoek van dit proefschrift is erop gericht de wisselwerking tussen sterrenstelsels tijdens de vorming van clusters van sterrenstelsels, en de consequenties daarvan voor o.a. de evolutie van sterrenstelsels nader te bestuderen.

In hoofdstuk II van dit proefschrift worden eerst de resultaten besproken van een systematische numerieke studie van het effect van botsingen tussen N-deeltjes systemen. Deze resultaten werden ingebouwd in numerieke simulaties van de vorming van kleine groepen en clusters van sterrenstelsels. De meeste botsingen die leiden tot de vorming van een nieuw, zwaarder deeltje komen voor tijdens de vorming van de clusters. Na de vorming van de clusters is de fractie gefuseerde stelsels in deze modelberekeningen in goede overeenstemming met de fractie E stelsels waargenomen in clusters. Bovendien blijken deze deeltjes, evenals E stelsels bij voorkeur in de centrale delen van clusters voor te komen.

In hoofdstuk III worden de resultaten besproken van numerieke simulaties van een expanderend heelal, waarin botsingeffecten in rekening gebracht zijn. Dergelijke berekeningen werden ongeveer gelijktijdig gedaan door Jones en Efstathiou en door Aarseth en Fall in Cambridge. Zij lieten o.a. zien dat het impulsmoment van de gefuseerde deeltjes vergelijkbaar is met dat van elliptische stelsels. In hoofdstuk III wordt vooral aandacht besteed aan de clustering eigenschappen van de deeltjes. De waarschijnlijkheid dat een deeltje met massa m_1 fuseert met een deeltje met massa m_2, blijkt ongeveer evenredig te zijn met $m_1 + m_2$. De clustering eigenschappen van elliptische stelsels kunnen zodoende verklaard worden wanneer hun massa enige malen zwaarder is dan die van spiraalstelsels. SO stelsels moeten op grond van hun clustering eigenschappen als een overgangsklasse tussen S en E stelsels gekenmerkt worden.

In het tweede deel van hoofdstuk III wordt vervolgens een model onderzocht waarin sterrenstelsels bij $z \sim 10$ als schijfstelsels gevormd worden met een helderheidsfunctie welke ongeveer gelijk is aan de huidige. Tijdens het daaropvolgende tijdperk van clustervorming groeit de sferische component van de stelsels door invang van kleinere stelsels. Sommige stelsels zullen daardoor evolueren tot SO stelsels of zelfs

E stelsels. Een belangrijke voorspelling van het model, nl. dat de fractie E stelsels toeneemt met de massa van sterrenstelsels blijkt in goede overeenstemming te zijn met recente waarnemingen.

In hoofdstuk IV wordt de evolutie van een rijke cluster van sterrenstelsels onderzocht. Massasegregatie, de relatie tussen de dichtheid van een bepaald gebied en de fractie sterrenstelsels van een bepaald type, en massaverlies van sterrenstelsels in botsingen met andere stelsels worden besproken en vergeleken met recente waarnemingsgegevens.

De algemene conclusie van hoofstukken II, III en IV is dat allerlei eigenschappen van SO en E stelsels (meer in het algemeen, van vroeg-type stelsels) op natuurlijke wijze verklaard kunnen worden in het hierarchisch clustering model wanneer sterrenstelsels oorspronkelijk gevormd worden als schijfvormige (spiraal)stelsels. Tijdens het daaropvolgende tijdperk van clustervorming zullen dan door middel van invang van kleinere stelsels de sferische componenten, en daarmee SO en E stelsels, gevormd worden.

In hoofdstuk V wordt een mogelijk verband tussen de activiteit in de kernen van sterrenstelsels en de invang van stelsels onderzocht. Tegenwoordig wordt meestal aangenomen dat de productie van grote hoeveelheden energie door actieve stelsels het gevolg is van accretie van gas door een zwart gat met een massa van 10^6 à 10^8 $M_\odot$. Ongeveer 5 tot 40% van de massa van het gas kan daarbij in energie omgezet worden. Een van de manieren waarop gas aan het gat toegevoerd kan worden is invang van een ster door het gat. Wanneer een ster het gat al te dicht nadert wordt zij door sterke getijdekrachten, uitgeoefend door het gat, uit elkaar getrokken en vervolgens verzwolgen. Onder omstandigheden zoals die heersen in de kern van onze Melkweg zal een zwart gat van 5.10^6 $M_\odot$ ongeveer één maal per 1000 jaar een ster invangen. In de meest actieve stelsels wordt echter een hoeveelheid gas gelijk aan ongeveer één zonsmassa per jaar als brandstof aan het gat geleverd. Wanneer een sterrenstelsel een kleiner stelsel invangt zal de sterverdeling rond het zwarte gat verstoord worden. Het aantal sterren dat per tijdseenheid ingevangen wordt door het gat kan daardoor waarschijnlijk toenemen en de waarde van één zonsmassa per jaar bereiken. De actieve periode van een stelsel zal ongeveer 10^6 tot 10^7 jaar duren, wanneer een klein stelseltje ingevangen wordt. De helderheidsfunctie van actieve kernen kan ook ruwweg geschat worden in dit model. Deze functie kan de

vergelijking met waargenomen helderheidsfuncties goed doorstaan.

De omstandigheden in de kernen van sterrenstelsels zijn niet goed bekend. Het onderzoek beschreven in hoofdstuk V is daarom in belangrijke mate verkennend. De resultaten van dit onderzoek wijzen er echter op dat kannibalisme onder sterrenstelsels van groot belang kan zijn voor de processen die zich in actieve kernen afspelen.

STUDIEOVERZICHT

Na het behalen van het einddiploma Gymnasium-β aan het Christelijk Lyceum te Dordrecht begon ik in 1968 mijn studie aan de Vrije Universiteit te Amsterdam. In 1976 deed ik het doktoraalexamen theoretische fysica met bijvak wiskunde. Van 1975 tot 1977 gaf ik natuurkundelessen aan de Crhistelijke Scholengemeenschap Sweelinck te Amsterdam.
Vanaf begin 1977 tot 1 januari 1981 heb ik als wetenschappelijk assistent aan de Sterrewacht te Leiden gewerkt onder begeleiding van Dr. C.A. Norman en Prof. Dr. H. van der Laan. Dit proefschrift kwam grotendeels tot stand gedurende die periode.

DANKBETUIGING

Sverre Aarseth uit Cambridge wil ik hartelijk danken voor het beschikbaar stellen van zijn computerprogramma waarmee veel van de berekeningen in hoofdstukken III en IV gedaan zijn.

Veel medewerkers van de Leidse Sterrewacht hebben bijgedragen aan de totstandkoming van dit proefschrift. Vooral Colin Norman zou ik willen bedanken voor zijn enthousiasme en voor de vele stimulerende discussies die ik met hem heb gevoerd. Marja van Haaster en Lenore de Leeuw hebben dit proefschrift grotendeels getypt. Helaas verbiedt het promotie-reglement van de Leidse Universiteit mij hen daarvoor te bedanken.