PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Galaxy Formation after z = 1000

B. J. T. Jones

Phil. Trans. R. Soc. Lond. A 1980 **296**, 289-298
doi: 10.1098/rsta.1980.0171

Phil. Trans. R. Soc. Lond. A **296**, 289–298 (1980) [289]
Printed in Great Britain

Galaxy formation after $z = 1000$

By B. J. T. Jones
*Institute of Astronomy, Madingley Road, Cambridge CB*3 0*HA, U.K.*

The 'hierarchical clustering' and 'pancake' theories for galaxy formation are reviewed. In spite of the considerable difference between these two schemes it is difficult to offer observational tests that might discriminate whether galaxies or clusters of galaxies formed first. Recent observations of the microwave background radiation spectrum suggest that we may be looking back to the time of galaxy formation, and future isotropy measurements below 1 mm may provide vital clues.

1. Introduction

Barrow (this symposium) has clearly reviewed our understanding of the early Universe up until the time the hydrogen recombined at a red shift $z \sim 1000$. If we follow the canonical hot big bang picture (see, for example, Weinberg 1972), it seems that we have good reason for believing that the Universe never deviated very much from the ideal state of homogeneity and isotropy. Of course the deviation needed to form galaxies is small but finite, and as yet we have no explanation for the origin of these deviations. Nevertheless, our growing understanding of physics near the origin of time, as exemplified by the work of Ellis *et al.* (1979) and Weinberg (1979), gives us grounds for hoping that the problem of the initial conditions for galaxy formation may be within our grasp in the not too distant future. Until such a time we must infer the likely conditions for galaxy formation from a detailed study of galaxy formation theories. The arguments against the cosmic turbulence theory of galaxy formation (see Jones (1976) for a review of this) and the Omnes-type baryon symmetric cosmogonies, for example, are now such as to present a serious challenge to anyone wishing to take these ideas further. (See the articles by Barrow (this symposium) and Steigman (1977) for further details. Although no single argument is conclusive, viewed as a whole these theories seem to pose more questions than they provide answers.) In comparison, the gravitational instability theories for galaxy formation have in recent years shown much promise as a framework in which to understand galaxy formation. Most of the impetus here has come from attempts to understand the details of galaxy clustering, an aspect of the subject that has been extensively and clearly reviewed by Fall (1979*a*, *b*).

There are in fact two quite distinct ideas as to how gravitational forces act on inhomogeneities to form galaxies. This is because there are two quite different possible kinds of primordial density perturbation in the pre-recombination Universe. There are density perturbations such that the entropy per baryon of the Universe is everywhere constant: such *adiabatic* perturbations behave like sound waves if their wavelength is small enough for pressure gradients to dominate over gravitational forces. There are also density perturbations such that the temperature is everywhere constant. These are referred to as *entropy* perturbations, or isothermal perturbations, and have the property that they do not evolve relative to the background Universe as the Universe expands. An important property of adiabatic perturbations is that before and during the recombination period, dissipative processes wipe out perturbations on mass scales smaller than

$$M_s \approx 5\times 10^{13}(\Omega h^2)^{-\frac{5}{4}} M_\odot, \tag{1}$$

where Ω denotes the ratio of the present density of the Universe to the density that would be required to just close the Universe; and h denotes the present Hubble constant in units of hundreds of kilometres per second per megaparsec; currently popular values are $h = 0.5$ and $\Omega = 0.1$ (see, for example, Peebles & Yu 1970; Bonometto *et al.* 1977). Adiabatic perturbations are thus to be thought of as the progenitors of galaxy clusters rather than galaxies. The Moscow group of Zel'dovich, Doroshkevich, Novikov and Shandarin has been mostly responsible for the development of this idea (see, for example, Doroshkevich *et al.* 1978), and have expounded a theory in which the galaxies form during the collapse of protoclusters having mass greater than M_s. Because such collapse generally leads to a flattened structure, this theory is often referred to as the 'pancake' theory. By way of contrast, isothermal perturbations are not wiped out by viscous processes and what emerges from recombination strongly reflects the initial conditions in the big bang itself. Hogan (1978) has given a fine exposition of the behaviour of such perturbations. At present, there seems to be no *a priori* reason to prefer any particular mass scale or amplitude spectrum for such inhomogeneities. The Jeans mass M_J just after recombination for a density perturbation of density contrast δ, where $\delta = \rho/\langle\rho\rangle - 1$, is

$$M_J \approx 10^6(\Omega h^2)^{-\frac{1}{2}}\, \delta^{-\frac{3}{2}} M_\odot, \tag{2}$$

and so, given a spectrum of inhomogeneities extending over all mass scales, we expect this scale to leave its imprint on the galaxy formation process. Of course, the spectrum of isothermal perturbations need not extend to such small mass scales, and it could well be that the present galaxy masses are a direct reflexion of the initial spectrum (see, for example, Binney & Silk 1978).

Broadly speaking, the situation we have at present is that there are two entirely different theories for the origin of galaxies via gravitational instability. In the isothermal density perturbation theory, galaxies are built up hierarchically in a way that depends on the unknown spectrum of inhomogeneities, though by working backwards we can infer from present observation what the initial spectrum should have been (Fall 1979*b*). In the adiabatic density perturbation theory, it is the fragmentation of the collapsing protoclusters that leads to galaxy formation. It is perhaps surprising that two theories that are so different seem equally able to describe the process of galaxy formation. The important question is: by what means could we distinguish these theories on the basis of what is currently known about galaxies? Moreover, are there any observations that we could make that would enable us to discriminate? We may, for example, ask whether both theories really make the same predictions about the observed large scale clustering of the Universe, or the masses and angular momenta of galaxies. On the observational side there is the possibility of directly observing young galaxies (Sunyaev *et al.* 1978), or studying the isotropy and spectrum of the microwave background radiation spectrum.

2. Hierarchical clustering models of galaxy formation

The theories of galaxy formation in which structure is built up by continual agglomeration of matter are exemplified by the picture of White & Rees (1978). In this theory it is assumed that a large fraction of the mass of the Universe forms into 'dark material' at some early stage.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The precise nature of the dark material is irrelevant; it could be black holes, 'Population III' stars or anything else that does not contribute to the present light in the Universe. It is only assumed that this material is inhomogeneously distributed throughout the Universe, and thus provides a set of potential wells into which the remnant gas can fall. Two physical processes operate simultaneously. The dark matter clusters under the influence of the forces of gravity: thus neighbouring lumps of dark material continually merge to form larger lumps of dark material. At the same time the remnant gas cascades into these potential wells, forming stars when its density becomes great enough for it to become self-gravitating. It is these stars that compose the luminous component of the observed Universe. An important aspect of this theory is that it provides an explanation for the apparent upper limit on galaxy masses. This upper limit arises because there is a maximum mass above which a gas cloud that is in thermal equilibrium takes longer to cool than its gravitational timescale. Such a cloud can only deflate on the cooling time for the gas. This maximum mass is on the order of $10^{12}\,M_{\odot}$, and smaller masses of gas can never be quasi-statically pressure-supported unless they are cooler than 10^4 K. This important fact is due to Silk (1977) and Rees & Ostriker (1977).

The use of computer simulations of the gravitational clustering phenomenon by using gravitational N-body programs (Aarseth 1979) has greatly increased our understanding of how structure might be built up hierarchically in the Universe. Of course, there is no dissipation of energy in these models, and so they can only be used to describe situations where gas dynamic effects are not of importance. Thus while we may have the White–Rees theory in mind when examining these simulations, they cannot be used in any straightforward way to model the 'pancake' theory. The simulations enable us to study the growth of clustering (see Gott & Rees 1976; Aarseth *et al.* 1979; Efstathiou 1979*a*, *b*; for a review, see Fall 1979*a*) and the way in which lumps (galaxies?) acquire their angular momentum via tidal interactions (Peebles 1969; Efstathiou & Jones 1979*a*).

The situation with regard to understanding the clustering of galaxies via these numerical simulations is well reviewed by Fall (1979*a*, *b*). It certainly seems that we can understand the present clustering as measured by the two-point correlation function $\xi(r)$ in terms of simple power law spectrum for the initial inhomogeneity of the Universe together with non-dissipative gravitational clustering. There are some technical problems concerned with interpreting the N-body simulations: particle discreteness effects dominate the small scale clustering on scales less than the mean interparticle separation (Fall 1978), and there are problems arising out of the fact that the simulations start off in a non-clustered state (see also Efstathiou 1979*b*).

The situation regarding angular momentum is particularly interesting. Peebles (1969) had made an analytical estimate of the transfer of angular momentum via tidal torques. The important quantity is the induced rotation velocity compared with the velocity required to maintain virial equilibrium, and this is usually expressed as the dimensionless number

$$\lambda = \frac{\mathscr{H}}{M}\left(\frac{E}{GM^3}\right)^{\frac{1}{2}}, \tag{3}$$

where $\mathscr{H}$ is the angular momentum, E is the total energy and M is the mass of the protogalaxy. Peebles's calculation of $\lambda \approx 0.06$ has since been confirmed by the 1000-body simulations of Efstathiou & Jones (1979*a*) who obtained a median value

$$\lambda = 0.07 \pm 0.02, \tag{4}$$

independent of the protogalaxy mass. Of course, any dissipation that takes place during the collapse of protogalaxies increases λ. It seems significant therefore that present observations of the rotations of elliptical galaxies (see, for example, Illingworth 1977; Schechter & Gunn 1979; Davies 1978) are consistent with this low value of λ, since this implies that elliptical galaxies formed by dissipationless collapse. This is a problem for the White–Rees scheme since in that theory the visible parts of elliptical galaxies are formed by the dissipative collapse of gas onto the halo. On the other hand, for disk-type galaxies where $\lambda \sim 1$, dissipation is needed to increase λ from the tidally induced value. Without a halo the Galaxy would have had to have collapsed from a radius of 500 kpc to obtain the necessary increase in λ. This is somewhat implausible since the collapse time at the radius of maximum expansion would have been *ca.* 2×10^{10} years. However, with a massive halo the gas is not initially self-gravitating and a gas mass: halo mass ratio of 1:7 is sufficient to account for the present rotation of the disk if it collapsed from a radius of only 100 kpc. The White–Rees scheme is therefore consistent as a picture of spiral galaxy formation, but something else is needed to explain elliptical galaxies.

3. The 'pancake' theory

The 'pancake' theory, wherein galaxies form as a consequence of the collapse and fragmentation of protoclusters of galaxies, is rather more difficult to assess. This is largely because gas-dynamic processes are more difficult to evaluate quantitatively, and proper gas-dynamic simulations are not as yet available. The fragmentation problem here is no less complex than fragmentation in other astrophysical contents, but the important physical processes have been recognized by Sunyaev & Zel'dovich (1972) and by Doroshkevich *et al.* (1978). The collapse is to a planar configuration, a 'pancake', because of the initial irregular shape of the protocluster, and it is supersonic, so the pancake is cool (*ca.* 10^4 K) and dense. Gravitational instability probably has the most important effect on fragmentation, and this gives rise to fragments of mass

$$M_{\dagger} \sim 2 \times 10^8 \, n_0^{-\frac{1}{2}} \, M_{\odot} \tag{5}$$

(Jones & Wyse 1979), where n_0 is the particle number density in the protocluster at the time that starts to collapse. Typically we are talking about fragments in the range 2×10^8 to $2 \times 10^{10} \, M_{\odot}$ and so we must build galaxies by coalescing these fragments: this galaxy-building process is, however, not as straight forward as in the other version of the gravitational instability theory (see, for example, Doroshkevich *et al.* 1977, 1978). It is therefore difficult to assess how much angular momentum these galaxies have. Jones *et al.* (1979) have given arguments to suggest that, in spite of dissipation in the cold layer, λ for the galaxies has to be the same as λ for the cluster in which they form, i.e. *ca.* 0.07, since the clusters acquire their angular momentum via tidal interactions (see also Jones 1979).

Perhaps surprisingly, the clustering correlation function has not been evaluated in any detail for this theory. Since the characteristic mass scale $M_s \sim 5 \times 10^{13} \, (\Omega h^2)^{-1} \, M_{\odot}$ is imposed on the clustering distribution by primordial dissipative processes, we might expect this scale to produce a feature in the two-point correlation function (S. A. Bonometto & S. M. Fall, personal communication). The issue is not, however, clear-cut and it would be useful to attempt a simulation. The facts that the density run in galaxy clusters falls off as r^{-2}, and that the sky is seen in projection, could obscure any observational evidence for this preferred scale.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The microwave background radiation spectrum

The cosmic microwave background radiation discovered by Penzias & Wilson (1965) is the strongest piece of evidence for a hot singular origin for the Universe (Dicke *et al.* 1965). In a perfectly homogeneous and isotropic Friedman Universe, the spectrum of the radiation is Planckian and the radiation is isotropic at all frequencies. Distortions of the spectrum shape (as first discussed by Weymann (1965, 1966)) and deviations from isotropy therefore provide us with a powerful tool for studying the Universe at the earliest epochs, and put constraints on theories of galaxy formation.

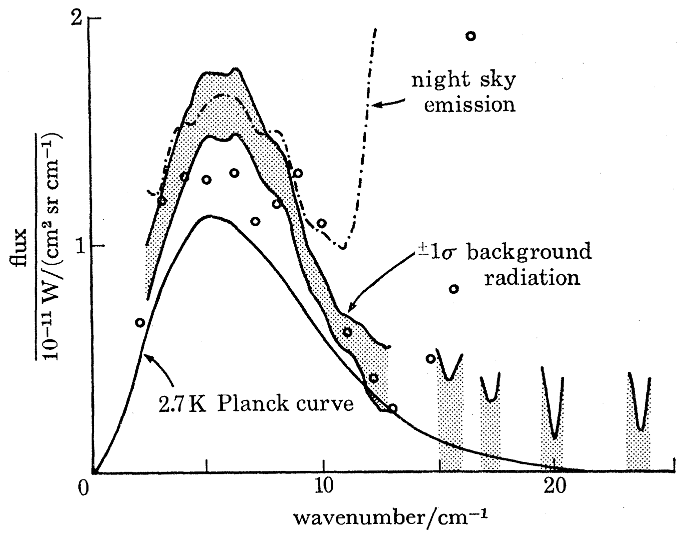

Figure 1. The data on the background radiation spectrum taken from Woody & Richards (1979). Also shown are the data of Robson *et al.* (1974) ○, a 2.7 K Planck function (this fits the Rayleigh–Jeans observations; see figure 2), and the contribution from night sky emission. The error bars for the data from Robson *et al.* are about twice the size of those for the Woody–Richards data. It must be remembered that there may be calibration errors that could lead to a vertical shifting of the various data sets relative to one another. The spectrum of the 'excess' radiation is sharply peaked, particularly on the high-frequency side; that is what will be most difficult to understand about these data.

Distortions of the Planckian shape have been considered in a variety of galaxy formation theories. If energy is injected into the Universe between a red shift $z_b \sim 5 \times 10\, \Omega^{-\frac{6}{5}}$ and the epoch of recombination, we should expect to see a deviation from the Planckian shape that depends on the amount of energy added. (The significance of the red shift z_b is that the part of the spectrum in the currently observed frequency range can be rapidly rethermalized at epochs earlier than z_b). The radiation field is heated up in the sense that the Planck curve is shifted to higher frequencies, but no photons are created to fill in a blackbody spectrum at the higher temperature if the energy is injected at $z < z_b$. The subject has been discussed at various levels of sophistication and applied to different galaxy formation theories by Sunyaev & Zel'dovich (1969, 1970), Zel'dovich *et al.* (1972), Chan & Jones (1975*a–d*) and Jones & Steigman (1978). The paper of Chan & Jones (1975*d*) displays detailed spectra of the kind of distortion expected in the cosmic turbulence theory. All of these distortion calculations have one feature in common: there is an *excess* flux at wavelengths shorter than the peak of the spectrum compared with what would have been expected on the basis of the ν^2 Rayleigh–Jeans part of the spectrum on the longer wavelength side of the peak.

It was therefore a matter of some surprise that the recent measurements of Woody & Richards (1979), with a balloon-borne spectrophotometer, revealed a spectrum having an apparent deficit of infrared photons relative to that expected on the basis of observations of the Rayleigh–Jeans part of the spectrum. Their data are shown in figure 1, and the experiment has been discussed in detail by Muehlner (1979). The data can, of course, be regarded as a blackbody spectrum with excess emission around the peak, and, indeed, Woody & Richards (1979) remark that the data are consistent, at the 80% confidence level, with a 2.79 K blackbody curve having an emissivity of 1.27. The spectral range covered by the experiment ranges from 250 μm to 6 mm; the problem of subtracting sky emission accurately over that range is a difficult one, so it is vital to check this result. But what are we to make of this spectrum if it is indeed the spectrum of the cosmic background radiation?

Interpreting the Woody–Richards data as an excess of radiation in the vicinity of the peak of the spectrum, relative to the Rayleigh–Jeans part of the spectrum, we might attribute this excess to starlight from bright primaeval galaxies which has been re-emitted in the infrared by the dust associated with the star-forming process. If the dust in our Galaxy and nearby galaxies provides any basis for judgement, the peak of the radiation from the dust would be in the vicinity of 100 μm, and the primaeval galaxies we are seeing in this way are at red shifts in the range 10–20. Of course, on this hypothesis it is a coincidence that the peaks of the spectra of the cosmic background radiation and the re-radiated starlight are at roughly the same frequency. It is not an unlikely coincidence since this is not the only reason for putting the red shift of galaxy formation in the range $z = 10$–20. Since the emissivity of dust is wavelength dependent, there will be a frequency ν_1 at which the cosmic background becomes optically thick to the dust and this hypothesis implies a significant optical depth in the dust at wavelengths not far short of the peak of the Planck curve. Moreover, the fact that the emissivity of dust is wavelength dependent (proportional roughly to λ^{-1}) means that the well observed Rayleigh–Jeans part of the spectrum is not significantly perturbed by the extra emission from the dust. (The long wavelength part of the spectrum of radiation from dust is at least as steep as ν^3 and this is one of the main reasons why it is difficult to produce the whole of the background radiation spectrum from dust alone unless we postulate dust with somewhat unusual properties.) The spectrum shortward of the peak depends in detail on the relation between optical depth and temperature in the dust at those frequencies, and it may well be that this part of the spectrum will be difficult to understand in terms of the kind of dust seen in our Galaxy.

Of course, if we ignore the measurements at wavelengths $\lambda < 1$ mm on the grounds that at such wavelengths the microwave background radiation is only a small fraction of the total signal (see figure 1), we could try to argue that we are seeing a classical Compton distortion of the spectrum resulting from the injection of energy into the Universe at early times. The observations of the spectrum over the wavelength range $0.33 \text{ cm} < \lambda < 75 \text{ cm}$ yields a thermodynamic temperature

$$T = 2.72 \pm 0.08 \text{ K} \tag{6}$$

(Chan & Jones 1975*b*), and it is clear that the Woody–Richards observation around the peak of the spectrum show an excess flux relative to this temperature. For small Compton distortions, the brightness temperature T_b at frequency ν is related to the Rayleigh–Jeans temperature T_{RJ} by

$$T_b(\nu) = T_{RJ}[1 + Y(h\nu/kT_{RJ}^{-2})], \tag{7}$$

Y being a parameter measuring the strength of the heating effect (see, for example, Jones & Steigman 1978), Thus at frequencies significantly higher than 4 cm^{-1} the brightness temperature should rise linearly with frequency. The plot of brightness temperature against frequency is shown in figure 2: It is clear that (7) is a rather poor fit since $T_b(\nu)$ in fact falls off towards higher frequencies. Nevertheless for $T_{RJ} = 2.72$ K the data out to 8 cm^{-1} imply $Y < 0.15$, while out to 13 cm^{-1} we have $Y < 0.05$. It is the nature of the deviation from a blackbody of the Rayleigh–Jeans temperature $T_{RJ} \approx 2.7$ K that makes Compton distortions an unattractive hypothesis.

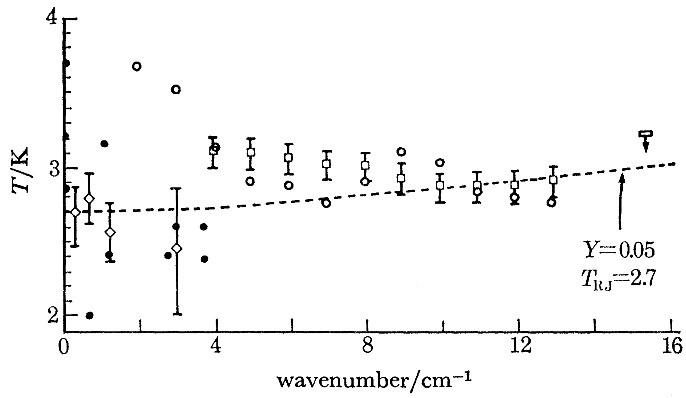

FIGURE 2. Thermodynamic temperatures for observations of the background radiation spectrum. The diamonds are the best data in the radio frequency range and these experiments all involved Dr D. Wilkinson (this argues for a certain level of consistency). The filled circles are other data from the compilation of Danese de Zotti (1977). The squares are the Woody–Richards data, and open circles the data of Robson *et al.* (1974). It is arguable that the data are consistent with a 2.95 K black body. The trend of the data for a typical Compton-distortion is shown by the broken line. That this line has the wrong trend makes the classical distortion a somewhat unattractive explanation for the observations.

Let us now turn attention briefly to the question of the small-scale anisotropy of the microwave background radiation. Just how far we are looking back to when observing the cosmic microwave background radiation depends on how much gas there is around at various epochs and its state of ionization. If a fraction f of the Universe is in the form of gas having ionization x_e, optical depth unity to electron scattering is achieved at red shift

$$z_{\tau=1} \sim 7(fx_e)^{-\frac{2}{3}} (\Omega h^2)^{-\frac{1}{3}}. \tag{8}$$

The angular scale (in minutes of arc) subtended by a sphere of mass M at red shifts $z \gg \Omega^{-1}$ is

$$\theta \approx (M/10^{12} M_{\odot})^{\frac{1}{3}} (\Omega h^2)^{\frac{2}{3}} h^{-1}. \tag{9}$$

Thus if temperature variations $\Delta T/T$ are detected over a range of angular scales from, say, 1′ to 10′, we shall have a good picture of the Universe at $z_{\tau=1}$ on scales relevant to galaxies and galaxy clusters. Since the amplitude of the temperature fluctuation $\Delta T/T$ on a given scale depends on the velocities relative to the Hubble flow on that scale, we shall also have information on the evolution of the clustering if we can decide where $z_{\tau=1}$ is.

The microwave background radiation is not the only radiation field that can be used to probe the early Universe. Hogan & Rees (1979) have recently discussed the possibilities of looking out to red shifts $z \approx 10$ with the use of the 21 cm emission from any neutral hydrogen in the Universe, and out to $z \approx 5$ looking in visible wavelengths for red-shifted Lyα emission.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

5. Final remarks

The prospects for observational evidence on the evolution of galaxies and galaxy clusters look good. It may be that the microwave background radiation near $\lambda \sim 1$ mm has a substantial component from starlight that has been rethermalized by dust at the time of galaxy formation. In that case we have observed the galaxy formation epoch already and must now wait for better spectrum measurements and observations of the small-scale anisotropy of the radiation at these wavelengths. (See Caderni *et al.* (1977) for measurements with $\theta = 25'$ at $\lambda = 1.0$–1.4 mm.)

There is a lack of predictive power in the 'pancake' theory at the moment, expecially when compared with the hierarchical clustering kind of theory, and so evaluating the relative merits of these two ideas is rather difficult at present. Of course it might well be that galaxy-size lumps were the first thing to condense out of the Universe (cf. Binney & Silk 1978): the only argument against this 'onion-skin' model is that it is difficult to explain the apparent strong rise in the galactic mass function below masses *ca.* $10^{11}\,M_{\odot}$ unless the spectrum of inhomogeneities on these scales is different from the spectrum on cluster scales ($M \gg 10^{11}\,M_{\odot}$). Larson's earlier work on modelling galaxy collapse (see, for example, Larson 1974) was based on this kind of picture, and if we change over to either the hierarchical clustering theory or the pancake theory in which galaxies are built from smaller units, we shall have to change the picture of the more recent phases of galaxy evolution (see, for example, White & Rees 1978; Tinsley & Larson 1979).

We may be successful in explaining the origin of galactic spin, the nature of the clustering of galaxies, and perhaps even the galactic mass function, but there still remain some major problems, not the least of which is why there appear to be two species of galaxy: elliptical systems and disk-like systems (see Efstathiou & Jones (1979*b*) for a detailed review of this problem). In terms of the 'onion-skin' model there were several plausible suggestions. For example, a protogalaxy might become an elliptical or form a disk depending on the ratio of the timescale for star formation to the collapse timescale, or depending on whether the galaxy became optically thick before becoming rotationally supported. (If indeed the present ideas on the contribution of dust to the microwave background radiation spectrum are correct there may be a case for reviving the latter idea.) However, there is still the problem of why larger elliptical galaxies are found preferentially in richer galaxy clusters. That fact suggests that galaxy formation is to some extent environment dependent.

There are environment dependent effects in the hierarchical clustering theory, in the sense that galaxy building seems to proceed most rapidly in the cluster environment. This phenomenon has shown up in 'merger' simulations using gravitational N-body programs that have been modified to model the coalescence of substructure in an efficient way (Aarseth & Fall 1979; Jones & Efstathiou 1979; Roos & Norman 1979). It turns out that the merged galaxies occur preferentially in galaxy clusters, and in about the right numbers to be interpreted as elliptical galaxies. Because in these simulations the mergers occur in bound nearly linear orbits, the angular momentum of the merger products is low ($\lambda \sim 0.07$). This is highly encouraging, but it must be remembered that many of the merging galaxies will already have formed disks that are rapidly rotating ($\lambda \sim 1$) and it is by no means clear that the merging of rapidly rotating disks can produce a slowly rotating elliptical galaxy. The following remarks may help to lend plausibility to this somewhat complex picture.

If the merging of N systems having disks takes place, the mean angular momentum per unit mass of the luminous material will *decrease* as $N^{-\frac{1}{2}}$ (on the reasonable assumption that the disk angular momenta are uncorrelated). Thus the merging of many spiral systems could produce a low λ luminous system from the stars that made up their stellar disks. What happens to the gas? Since the mergers take place preferentially in the cluster environment there is the possibility that the gas will be stripped from the system by the ram-pressure of the intergalactic gas, an idea recently exploited by Gisler (1979) to explain the Butcher–Oemler effect (Butcher & Oemler 1978). In the absence of gas-stripping, the gas dissipates and settles down into a new disk system, which might even get stripped at a later stage to leave an S0-type galaxy.

Where should we look for evidence of this process? The merging rate in rich galaxy clusters is presently very low because the velocity dispersion of the galaxies is high relative to the velocity dispersion of the stars in the galaxies. To judge on the basis of the merger-simulations, most mergers take place during the early collapse phase of groups and clusters of galaxies. Hence we ought to examine sparse groups with crossing times on the order of the Hubble time. It is highly suggestive that the gas-containing elliptical galaxies NGC 1052, NGC 4278, and NGC 5128 (CenA) are found in just this kind of environment. There are, however, some major questions that must be answered before claiming a definitive explanation of the Hubble sequence along these lines. How are colour gradients in elliptical galaxies created? Why do some ellipticals have no colour gradients at all? What is the fate of the globular cluster populations? Indeed, what is the origin of the globular cluster population?

My thanks go particularly to G. Efstathiou for countless discussions on galaxy formation and evolution, and to P. Martin for discussions on the spectrum of the microwave background radiation.

References (Jones)

Aarseth, S. J. 1979 In preparation.
Aarseth, S. J. & Fall, S. M. 1979 In preparation.
Aarseth, S. J., Gott, J. R. & Turner, E. L. 1979 *Astrophys. J.* **228**, 664.
Binney, J. & Silk, J. 1978 *Comments Astrophys. Space Phys.* **5**, 139.
Bonometto, S. A., de Zotti, G. F. & Lucchin, F. 1977 *Astron. Astrophys.* **54**, 405.
Butcher, H. & Oemler, A. 1978 *Astrophys. J.* **219**, 18.
Caderni, N., de Cosmo, V., Fabbri, R., Melchiorri, B., Melchiorri, F. & Natale, V. 1977 *Phys. Rev.* D **16**, 2424.
Chan, K.-L. & Jones, B. J. T. 1975*a* *Astrophys. J.* **195**, 1.
Chan, K.-L. & Jones, B. J. T. 1975*b* *Astrophys. J.* **198**, 245.
Chan, K.-L. & Jones, B. J. T. 1975*c* *Astrophys. J.* **200**, 454.
Chan, K.-L. & Jones, B. J. T. 1975*d* *Astrophys. J.* **200**, 461.
Danese, L. & de Zotti, G. 1977 *Riv. nuovo Cim.* **7**, 277.
Davies, R. 1978 Ph.D. thesis, University of Cambridge.
Dicke, R. H., Peebles, P. J. E., Roll, P. G. & Wilkinson, D. T. 1965 *Astrophys. J.* **142**, 414.
Doroshkevich, A. G., Saar, E. M. & Shandarin, S. F. 1977 In *The large scale structure of the Universe* (I.A.U. Symp. no. 79) (ed. M. S. Longair & J. Einasto), p. 423. Dordrecht: D. Reidel.
Doroshkevich, A. G., Shandarin, S. F. & Saar, E. 1978 *Mon. Not. R. astr. Soc.* **184**, 643.
Efstathiou, G. 1979*a* *Mon. Not. R. astr. Soc.* **187**, 117.
Efstathiou, G. 1979*b* Ph.D. thesis, Durham University.
Efstathiou, G. & Jones, B. J. T. 1979*a* *Mon. Not. R. astr. Soc.* **186**, 133.
Efstathiou, G. & Jones, B. J. T. 1979*b* *Comments Astrophys. Space Phys.* (In the press.)
Ellis, J., Gaillard, M. K. & Nanapoulos, D. V. 1979 *Phys. Lett.* B **80**, 360.
Fall, S. M. 1978 *Mon. Not. R. astr. Soc.* **185**, 165.
Fall, S. M. 1979*a* *Rev. mod. Phys.* **51**, 21.
Fall, S. M. 1979*b* In *Proc. 9th Texas Symposium on Relativistic Astrophysics. Ann. Phys.* (In the press.)

Gisler, G. 1979 *Astrophys. J.* **228**, 385.
Gott, J. R. & Rees, M. J. 1976 *Astron. Astrophys.* **45**, 365.
Hogan, C. 1978 *Mon. Not. R. astr. Soc.* **185**, 889.
Hogan, C. & Rees, M. J. 1979 *Mon. Not. R. astr. Soc.* **188**, 791.
Illingworth, G. 1977 *Astrophys. J.* **218**, L43.
Jones, B. J. T. 1976 *Rev. mod. Phys.* **48**, 107.
Jones, B. J. T. 1979 In *Proc. 9th Texas Symposium on Relativistic Astrophysics. Ann. Phys.* (In the press.)
Jones, B. J. T. & Efstathiou, G. 1979 *Mon. Not. R. astr. Soc.* **189**, 27.
Jones, B. J. T. & Steigman, G. 1978 *Mon. Not. R. astr. Soc.* **183**, 585.
Jones, B. J. T. & Wyse, R. 1979 In preparation.
Jones, B. J. T., Palmer, P. L. & Wyse, R. 1979 In preparation.
Larson, R. B. 1974 In *Dynamics of stellar systems* (I.A.U. Symp. no. 69) (ed. A. Hayli), p. 247. Dordrecht: D. Reidel.
Muehlner, D. J. 1979 In *Proc. 9th Texas Symposium on Relativistic Astrophysics. Ann. Phys.* (In the press.)
Peebles, P. J. E. 1969 *Astrophys. J.* **155**, 393.
Peebles, P. J. E. & Yu, J. T. 1970 *Astrophys. J.* **162**, 815.
Penzias, A. A. & Wilson, R. A. 1965 *Astrophys. J.* **142**, 419.
Rees, M. J. & Ostriker, J. P. 1977 *Mon. Not. R. astr. Soc.* **179**, 451.
Robson, E. I., Vickers, D. G., Huizinga, J. S., Beckman, J. E. & Clegg, P. E. 1974 *Nature, Lond.* **251**, 591.
Roos, N. & Norman, C. 1979 *Astron. Astrophys.* **76**, 75.
Schechter, P. L. & Gunn, J. E. 1979 *Astrophys. J.* **229**, 472.
Silk, J. I. 1977 *Astrophys. J.* **211**, 638.
Steigman, G. 1977 *A. Rev. Astron. Astrophys.* **14**, 339.
Sunyaev, R. A. & Zel'dovich, Y. B. 1970 *Astrophys. Space Sci.* **7**, 20.
Sunyaev, R. A. & Zel'dovich, Y. B. 1972 *Astron. Astrophys.* **20**, 189.
Sunyaev, R. A., Tinsley, B. M. & Meier, D. L. 1978 *Comments Astrophys. Space Phys.* **7**, 183.
Tinsley, B. M. & Larson, R. B. 1979 *Mon. Not. R. astr. Soc.* **186**, 503.
Weinberg, S. 1972 *Gravitation and cosmology*. New York: Wiley.
Weinberg, S. 1979 *Phys. Rev. Lett.* **42**, 850.
Weymann, R. 1965 *Phys. Fluids* **8**, 2112.
Weymann, R. 1966 *Astrophys. J.* **145**, 560.
White, S. D. M. & Rees, M. J. 1978 *Mon. Not. R. astr. Soc.* **183**, 341.
Woody, D. P. & Richards, P. L. 1979 *Phys. Rev. Lett.* **42**, 925.
Zel'dovich, Y. B. & Sunyaev, R. A. 1969 *Astrophys. Space Sci.* **4**, 301.
Zel'dovich, Y. B., Illarionov, A. F. & Sunyaev, R. A. 1972 *Soviet Phys. JETP* **35**, 643.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES