

# THE COLLAPSE AND FORMATION OF GALAXIES

R. B. LARSON

*Yale University Observatory, New Haven, Conn., U.S.A.*

## 1. Introduction: The Problems

The ultimate aim of studies of the dynamics of stellar systems is to gain an understanding of why they have the structures that are observed, how they might have formed, and how they might evolve with time. In the case of galaxies, unlike smaller astronomical systems such as stars and star clusters, the problem of understanding their present structure is inseparable from the problem of understanding their formation since, as has long been clear, the two-body relaxation time in galaxies is much longer than the age of the universe, so that properties such as the radial density distribution and the flattening of elliptical galaxies cannot have changed significantly since they were formed. Other features of galaxies, such as the presence of spiral patterns or nuclear activity, may be of a more transitory nature and not closely related to the formation process, but such phenomena probably involve mainly the gas component of galaxies and may be regarded as perturbations on the underlying stellar system, whose basic structure must still be understood in terms of the formation process.

In considering those properties of galaxies which should be explained by a theory of the formation process, it seems most profitable to concentrate first on those characteristics which show the greatest degree of regularity or similarity from case to case. Thus it is natural to consider first the elliptical (E) galaxies, whose structures appear to be very smooth and regular and whose radial surface brightness distributions  $I(r)$  are very similar to each other (Liller, 1960, 1966; van Houten, 1961; King, 1975). Thus a model of galaxy formation should be able to explain first the radial brightness profiles of E galaxies, and should provide a basis for understanding why  $I(r)$  is so similar for different E galaxies. A second readily measurable property of the E galaxies for which considerable data exist is the isophotal ellipticity  $\varepsilon(r)$  as a function of radius. Unlike the surface brightness profile  $I(r)$ , however, the ellipticity profile  $\varepsilon(r)$  shows considerable variability from case to case; thus a theory of galaxy formation should also be able to explain typical  $\varepsilon(r)$  profiles, while still allowing the possibility of variations in  $\varepsilon(r)$  from case to case.

Many galaxies contain a prominent or (visually) dominant disc component, and in such cases it is necessary to explain the existence and the relative prominence of the disc and spheroidal (halo and/or nuclear bulge) components in a galaxy; again, considerable variations exist in the relative prominence of disc and spheroidal components (Freeman, 1970), and this should ultimately be understandable on the basis of a theory of galaxy formation. On the other hand, the characteristic exponential structure claimed for many galactic discs may not be so significant for theories of

galaxy formation, since the structure of galactic discs can be substantially altered by the occurrence of bar and/or spiral instabilities (see the contributions of Bardeen and Hohl to this symposium.)

In recent years it has become clear that a very general characteristic of both elliptical and spiral galaxies is the presence of radial composition gradients; the abundances of the heavy elements are always greatest at the center, and decrease markedly with radius. This effect is observed in the radial gradients in the strength of the CN band and other metallic absorption features in the spectra of E galaxies and the bulge components of spirals (e.g. McClure, 1969; Spinrad *et al.*, 1971, 1972); in the compositions of H II regions in the discs of spiral galaxies (e.g. Shields, 1974); in the radial gradient in the CN abundance of stars in the disc of our Galaxy (e.g. Janes, 1975); and in the radial variation of the composition of globular clusters in our Galaxy (Arp, 1965) and M31 (Hartwick and Sargent, 1974). In addition, the presence of composition differences between the various stellar populations in our Galaxy has long been recognized, and generally interpreted in terms of the collapse and simultaneous metal enrichment of the protocloud from which our Galaxy formed. A similar collapse or infall process might also provide a natural explanation for the composition gradients observed in other galaxies, in which case models for the formation of galaxies should be able to reproduce the observed abundance gradients. Finally, models for galaxy formation should be able to explain the observed correlation between the metal abundances and the masses of elliptical galaxies; various investigations (e.g., Sandage, 1972; Faber, 1973) have shown that the metal abundances of E galaxies increase systematically with increasing mass.

It may also be hoped that, in addition to explaining the various characteristics of galaxies mentioned above, a complete theoretical picture of the galaxy formation process will provide a better understanding of the origin of several other important phenomena which seem to be related in some way to the formation or early evolution of galaxies. For example, there is now considerable evidence that the quasar phenomenon is related in many cases to an early stage in the evolution of galactic nuclei. Also, the still incompletely understood problem of the origin of spiral structure may eventually be clarified by a better understanding of the formation and evolution of galaxies. At any rate, another criterion for deciding between possible contending models of galaxy formation is the extent to which the models may allow such additional phenomena to be understood.

## 2. Possible Models of Galaxy Formation

### 2.1. GRAVITATIONAL COLLAPSE VS NUCLEAR EJECTION

At present, since almost nothing is known in detail about how galaxies form, it is necessary to start by considering some possible models based on simple but plausible hypotheses. Hopefully, by comparing the predictions of the various models with the observed properties of galaxies, we can then get some idea about which types of models best account for the observations. Up to now, all of the models for which

detailed calculations have been made have been based on the hypothesis that galaxies form by the gravitational collapse or condensation of diffuse matter in regions of the universe which are dense enough to separate out from the expanding background and recollapse. This general picture of galaxy formation is implied by conventional big bang cosmologies in which it is assumed that the universe began with a nearly uniform density but with small superimposed density fluctuations which grew in amplitude through gravitational instability and eventually collapsed into galaxies (e.g., Peebles, 1972). The gravitational instability picture is attractive because it offers the possibility of explaining within a common framework not only the formation of galaxies but also the origin of the observed clustering of galaxies on various scales; as Peebles (1974a) has shown, the spatial distribution of galaxies is consistent with a gravitational instability origin for all scales of clustering, including the galaxies themselves as the smallest 'clusters'. The general success of the gravitational instability picture in accounting for the large scale distribution of matter in the Universe is persuasive evidence in favor of the gravitational collapse picture of galaxy formation.

A second important reason for considering collapse models for galaxy formation is the evidence from correlations between the metal abundances, kinematics, and spatial distributions of stars in our Galaxy that our Galaxy formed by a collapse process (Eggen *et al.*, 1962). Subsequent work has further substantiated the general result that the oldest and most metal poor stars and star clusters have the most extended spatial distributions, whereas younger and/or more metal rich objects are increasingly concentrated toward the galactic center or the galactic plane. As has already been noted, similar correlations between metal abundance and spatial position exist in other galaxies, and model collapse calculations (Larson, 1974a) show that these observations are quite naturally explained by a collapse model of galaxy formation.

While the collapse picture in some form is widely accepted, an alternative hypothesis has occasionally been advocated which proposes that galaxies or parts of galaxies have been formed by the creation of matter and its ejection from nuclear singularities or 'white holes'. However, apart from the fact that many of the apparently strange phenomena (such as quasars) which this concept has been invoked to explain now seem better accounted for on more conventional grounds, this hypothesis suffers several severe drawbacks as a possible basis for understanding galaxy formation. The first is that, since it lies outside the realm of known physics, one cannot use any known physical principles to construct a model of galaxy formation based on the nuclear ejection hypothesis. Secondly, even if a theory could be found to explain the ejection of such large amounts of mass from galactic nuclei, the ejection velocity would still have to be very finely adjusted to make the ejected matter fill a galaxy-size volume without escaping to infinity. Finally, an apparently fatal difficulty is that the ejected matter cannot possibly carry the amount of angular momentum possessed by typical galaxies, yet still remain bound in a galaxy. This is a general difficulty with all theories which attempt to account for galactic phenomena as a result of nuclear ejection of

matter; the ejected matter cannot carry any significant angular momentum, so that to the extent that the phenomena in question involve angular momentum, it must have some other origin. By contrast, gravitational collapse or infall theories encounter no such *a priori* difficulty in accounting for substantial amounts of angular momentum in galaxies.

## 2.2. SIMPLE COLLAPSE VS CLUSTERING OR ACCRETION

If we accept, for the reasons discussed above, that galaxies form by some sort of gravitational collapse or condensation process, we must then consider in more detail just how the pre-galactic material becomes collected together under gravity to form a galaxy. The simplest possible picture is that a galaxy begins as a discrete, roughly spherical, and nearly uniform cloud of gas which, after it has stopped expanding, is well isolated from its surroundings and collapses more or less all together to form a galaxy. Because of its simplicity, this picture has served as the basis for most of the model calculations which have been undertaken to describe the collapse of protogalaxies. However, it is probably a considerably over-idealized view of the real situation, in which a protogalaxy is probably not such a simple and well-defined object. It is possible, for example, that a protogalaxy may originate through the gravitational clustering of smaller objects, or that it may interact continually with its surroundings and grow from small beginnings by accretion of more matter. For example, it has been suggested (e.g., Peebles and Dicke, 1968) that the first condensations to form in the expanding universe actually have the mass of globular clusters, and that protogalaxies later develop as clusters of these globular-cluster sized fundamental units. A possible mathematical description of such a gravitational clustering process has been given by Saslaw (1972).

Whatever may be the most correct description of the galaxy formation process, it is likely to involve at least some elements of continuing accretion or infall of matter in a forming galaxy. This is because even in the most idealized case of an isolated uniform spherical cloud, collapse calculations (both for interstellar clouds and for protogalaxies) always show a nonhomologous collapse characterized by the development of a dense core surrounded by an extended envelope of residual gas which continues to fall into the central condensation until some effect such as an outflowing wind intervenes to halt the infall. More realistically, we would in general expect a forming galaxy to be surrounded by additional uncondensed material with which it continues to interact, eventually accreting at least part of it; this is suggested, for example, by the numerical simulations by Peebles (1971, 1972) of galaxy or cluster formation by gravitational instability, which show that the forming protogalaxies or protoclusters do not immediately become separated from their surroundings but continue to interact gravitationally with surrounding matter even after they have begun to collapse. Also, Oort (1970) has shown via simple models that one should in general expect ambient matter to continue to fall into a galaxy for a long time after it has formed. Tidal interactions between neighboring young galaxies may also continue to exert important effects on their structure and dynamics, and galaxies may even

collide inelastically and merge to form larger systems (Toomre and Toomre, 1972; Alladin *et al.*, 1975).

In view of the possibilities mentioned above, it can hardly be expected that simple idealized collapse models will provide a fully complete or correct description of the galaxy formation process. Perhaps it is reasonable to hope that the simple models which can now be treated numerically will adequately describe the formation of the simplest systems, namely the elliptical galaxies; systems with prominent disc components, spiral structure, nuclear activity, etc., may have formed in a more complex manner involving, for example, continuing interactions with or accretion of surrounding matter.

### 2.3. STELLAR DYNAMICS VS GAS DYNAMICS

We now consider the possibilities for constructing 'simple collapse models' intended first to represent the formation of E galaxies. The conventional general view concerning the formation and evolution of galaxies is that a galaxy begins as a diffuse cloud of gas which somehow condenses and over a period of time becomes transformed into stars, leaving in the end a system composed mainly of stars. This type of picture certainly seems required to explain the evolution of the disc of our Galaxy, where it is clear that star formation has gone on continuously over the history of the system. In the case of the elliptical galaxies, however, all indications are that there has been little or no star formation since early in their history (Larson and Tinsley, 1974), so that star formation in the E galaxies must have taken place relatively rapidly. Since nothing is known directly about how soon the star formation in E galaxies is completed, one can consider as a simple limiting case the possibility that star formation is completed at a very early stage, even before significant collapse has taken place. In this case one has only to consider the dynamics of a system of stars, and the complexities of attempting to treat the gas dynamics are avoided. Since the collapse of a system of stars is the simplest problem to study and has received a fair amount of attention, we shall consider in the next section the results that have been obtained for purely stellar collapse models. Later sections will consider the probable necessity of including a gas component in a more complete model for the formation of E galaxies, and also the results that have been obtained with models which treat the dynamics of both gas and stars.

## 3. Stellar Collapse Models

It has been remarked by King (1966) and others that, despite the fact that two-body relaxation processes cannot be important in elliptical galaxies, these galaxies nevertheless 'look relaxed'; that is, their radial surface brightness distributions resemble those predicted on the basis of two-body relaxation for models of star clusters. A possible relaxation effect which might account for the apparently 'relaxed' structure of E galaxies has been described by Lynden-Bell (1967), who pointed out that the rapid changes in gravitational potential associated with the initial collapse and ap-

proach to equilibrium of a stellar system should cause large changes in the energies of the individual stars, thus producing a 'violent relaxation' effect. Assuming that this violent relaxation effect is sufficient to completely redistribute the stars in energy and bring about an approach to the most probable distribution, Lynden-Bell argued that the resulting system would be basically an isothermal sphere with a Maxwellian velocity distribution modified by a cutoff at the escape velocity. The particular form of the velocity cutoff derived by Lynden-Bell was equivalent to the 'lowered Maxwellian' approximation previously used by King (1966) as the basis for his star cluster models, and since these models were able to represent also the structure of E galaxies, it was argued that the structure of E galaxies could be explained as a result of violent relaxation. Recently, Wilson (1975) has generalized the King models to the case of axisymmetric rotating systems, and has shown that these models are capable of representing quite well a number of the basic properties of E galaxies.

However, the extent to which the violent relaxation concept and Lynden-Bell's statistical mechanics are applicable to elliptical galaxies is still somewhat unclear, in view of the fact that various numerical experiments designed to test the predictions of violent relaxation theory for one-dimensional systems (e.g., Lecar and Cohen, 1972) have shown only partial agreement with the theory: in all cases, significant numbers of particles are thrown into high energy 'halos' where relaxation is ineffective and the experimental energy distributions disagree with predictions. This result suggests that violent relaxation and Lynden-Bell's statistics may not apply to the halos of elliptical galaxies.

Recently, a more direct test of the applicability of violent relaxation theory to explaining the structure of E galaxies has been provided by the detailed numerical calculations by Gott (1973) of the collapse of an axisymmetric rotating system of stars, represented numerically by a system of 2000 discrete mass rings. Gott found that these stellar collapse calculations yielded models qualitatively resembling E galaxies in their structure, and also qualitatively consistent with the predictions of violent relaxation theory. In particular, these models have ellipticities and ellipticity profiles  $\epsilon(r)$  that are roughly consistent with those typically observed for E galaxies. However, the density distributions in the outer envelopes of E galaxies are not correctly reproduced by these models; at large radii, the density distributions in these stellar collapse models are approximately of the form  $\rho \propto r^{-4}$ , whereas the observed\* density distributions in E galaxies are considerably less steep, varying from about  $\rho \propto r^{-2.7}$  at intermediate radii to  $\rho \propto r^{-3}$  (Hubble's law) at large radii. Thus it appears that violent relaxation alone is not able to explain the observed density distribution in the envelopes of E galaxies. Furthermore, the stellar collapse models do not reproduce the highly condensed nuclei observed in the giant E galaxies (although this may in part be because Gott's calculations had to be cut off at small radii because of numerical difficulties.)

More recently, Gott (1975) has extended the stellar collapse models to the more

\* Assuming, as usual, a constant mass-to-light ratio.

realistic case where a protogalaxy is treated not as an isolated object but as a perturbation in an expanding cosmological medium. Ambient material which is gravitationally bound to the central density perturbation will then have its expansion retarded and eventually reversed, and will continue to fall into the forming galaxy for some time after the central perturbation has collapsed. The importance of continuing infall of matter as a natural consequence of the galaxy formation process has also been discussed by Oort (1970) on the basis of a slightly different picture whereby galaxies are assumed to develop from primordial velocity perturbations rather than density perturbations, and by Larson (1972a, 1974a) on the basis of hydrodynamical collapse calculations (see below). If the infalling matter is in the form of stars, as assumed by Gott, then a much more extended stellar envelope can be built up than in an isolated system where infall effects are ignored. For his 'infall' models, Gott found a density distribution of the form  $\rho \propto r^{-2.8}$ , in good agreement with that observed in the envelopes of E galaxies. Thus the infall models offer a more promising possible explanation of the structure of E galaxies; we note, however, that the density distribution in these models is no longer determined primarily by violent relaxation effects, but depends essentially on the infall effect and thus might be expected to depend to some extent on the ambient conditions present when a galaxy forms.

It is interesting that some of the gas-dynamical models of Larson (1974a) which were calculated with an expanding boundary yielded rather similar results, the density distribution at large radii being given approximately by  $\rho \propto r^{-3}$  independently of the detailed model assumptions. These expanding boundary models are basically similar to the Gott infall models in that the central region collapses first, while material farther from the center continues to expand for a longer period of time before turning around and falling back into the forming galaxy. Since the infalling matter is essentially in free fall, the dynamics is not much affected by whether it consists of stars only, as in the Gott models, or of a mixture of gas and stars, as in the Larson models; this may account for the fact that the resulting envelope density distributions are fairly similar in all of the 'infall' models. The fact that these results seem to resemble closely the density distributions in the envelopes of E galaxies suggests that the similar structure of E galaxies could be a result of cosmological infall effects, as proposed by Gott, but as will become evident, this explanation is not unique.

The rotating infall model computed by Gott (1975) has an ellipticity profile  $\varepsilon(r)$  which agrees approximately with that observed for the E4 galaxy NGC 4697, which is interesting because the same model also gives a good fit to the outer part of the observed surface brightness profile  $I(r)$  of NGC 4697. These calculations do not offer any direct explanation of the observed differences in ellipticity profiles between different E galaxies, but Gott suggests that this could be due to structural irregularities in the initial protogalaxies and to the resulting irregularities in the angular momentum distribution induced by tidal torques. Unfortunately, attempts to compare theoretical rotation curves with observations have so far been largely frustrated by the paucity of reliable data, but there appears to be an unexplained discrepancy between

the rotation curve predicted by Gott's model of NGC 4697 and the observed rotation curve of this galaxy (Bertola and Capaccioli, 1975), which shows smaller velocities than expected theoretically for any dynamical model with the same ellipticity.

In summary, the available stellar collapse models appear to enjoy some success in explaining the flattening of E galaxies and, when infall effects are incorporated, the light distribution in the envelopes of E galaxies. However, the stellar collapse models do not reproduce the highly condensed nuclei observed in giant E galaxies, nor do they offer any explanation of the observed composition gradients. For this it still appears necessary to assume that the formation of at least the central regions of E galaxies involves gaseous condensation and metal enrichment processes.

Before considering models containing a gas component, we note some additional stellar dynamical processes that could contribute to the development of dense cores in E galaxies. First, if an E galaxy begins as a cluster of smaller objects or mass concentrations (e.g., proto-globular clusters) or if it rapidly fragments into such objects, then two-body relaxation effects between these objects could cause some evolution in the structure of the system, just as in a cluster of stars. However, since it seems unlikely that a major part of the mass could remain concentrated in small objects for the  $\sim 10$ – $20$  crossing times required for two-body relaxation effects to produce a dense core, it also seems unlikely that this could be a major effect determining the structure of E galaxies, but it could play a contributing role. Secondly, Tremaine *et al.* (1975) have shown that the dynamical friction effect experienced by massive globular clusters moving in a smooth background of halo stars could cause them to spiral inward and pile up at the center to form a small dense nucleus. These authors suggest that this process might explain the apparently distinct small 'nucleus' at the center of M31, which has a radius of  $\sim 7$  pc and appears to have a slightly smaller metal abundance than its immediate surroundings (Spinrad *et al.*, 1972).

#### 4. Models Containing Gas

Even if the primordial gas in a protogalaxy is quickly exhausted and plays no role in the collapse process, it is still necessary to consider the dynamics of the recycled gas lost from evolving stars. If this recycled gas is not all lost from the galaxy, at least some of it will fall toward the center and form new stars there. Spitzer (1971) has outlined schematically how a sequence of stages of stellar mass loss, infall of gas toward the center, and formation of a new generation of stars in a region of smaller size and higher density might eventually lead to the formation of a highly condensed, metal rich nucleus. If the system has substantial angular momentum the recycled gas will presumably condense instead into a disc, and as Ostriker and Thuan (1975) have shown, galactic evolution models based on this assumption can reproduce successfully the observed distribution of stellar metal abundances in the solar neighborhood. It seems very likely, however, that primordial gas must also have been present along with the recycled gas from halo stars, since the primordial gas in a protogalaxy is almost certainly not totally converted into stars before the system has begun to col-



lapse, as assumed in the stellar collapse models. Indeed, the disc of our Galaxy must have formed at least partly from unprocessed primordial gas if the observed local deuterium abundance is of cosmological origin, as currently seems likely. Thus we are led to consider the collapse of systems of both gas and stars, in which gaseous dissipation effects play an important role in producing the final highly centrally condensed or flattened structure of the system.

In general, it is expected from experience with calculations for collapsing interstellar clouds that gaseous cooling or dissipation effects will lead quite naturally to a highly centrally condensed density distribution. However, the dynamics of the gas in a collapsing protogalaxy is difficult to treat in any precise way because the gas is probably quite 'turbulent', i.e. characterized by large density inhomogeneities and a more or less chaotic state of motion. The presence of turbulent motions in the proto-cloud from which our Galaxy formed seems required to account for the observed large velocity dispersions of the halo stars, many of which even have retrograde orbits. There are many ways in which such motions could have originated, including residual primeval turbulence, early supernova explosions or other violent events, collisions of pregalactic gas clouds, and the action of gravitational forces on an irregular mass distribution. Since the stars that form in a collapsing protogalaxy inherit the motions of the gas from which they form, it is important to keep track of the velocity dispersion of the turbulent gas and how it changes with time as a result of the collapse and dissipation processes. This means that it is necessary to specify the rate at which the turbulent motions are dissipated by inelastic collisions between the different randomly moving fluid elements. In addition, if we assume that star formation continues for a significant period of time during the collapse, it is necessary to specify the rate at which gas is transformed into stars. Needless to say, not enough is known about the physical processes involved to allow the gaseous dissipation and star formation rates to be calculated in any *a priori* way, and therefore it is necessary to resort to simple parameterized models which hopefully represent the essential features in a qualitatively correct fashion, while still leaving a few undetermined parameters which can be adjusted to achieve the best agreement with the observations. Clearly, one cannot hope in this way to produce models which are quantitatively accurate in all respects, but one can at least hope to gain some insight into the essential features of galaxy formation by collapse. The situation is perhaps analogous to the earliest days of the study of stellar structure, when many of the basic physical processes were not well understood but it was nevertheless possible to make important progress by using simple polytropic models.

One possible approach to estimating the turbulent dissipation rate, adopted by Larson (1969) and Brosche (1970), is to assume that the protogalactic gas is entirely concentrated in discrete spherical clouds occupying about one-tenth of the volume and possessing individual random motions with an isotropic velocity distribution. The rate of dissipation of the kinetic energy of cloud motions is then straightforwardly calculated in terms of the velocity dispersion and mean free path of the clouds. In order to estimate the size of the clouds, Larson (1969) assumed that the cloud mass

is equal to the Jeans mass at a temperature of  $10^4$  K, in which case the cloud masses are initially of the order of  $10^9 M_{\odot}$ ; Brosche (1970), on the other hand, estimated that a protogalaxy typically consists of  $\sim 10$  clouds, assuming that the typical angular momentum of a galaxy is the sum of the purely random angular momenta of its constituent clouds. In either case, the dissipation time is predicted to be comparable to or not much longer than the free-fall time, so that the collapse of a protogalaxy takes place within a few free-fall times, i.e. a few times  $10^9$  yr. Both investigations also arbitrarily adopted the common assumption that the star formation rate is proportional to a power of the gas density; Brosche assumed that the star formation rate varies as the square of the gas density, whereas Larson allowed both the exponent and the coefficient of proportionality in the power law to vary and studied the effects of these parameters on the resulting stellar density distribution, obtaining the best agreement with observations with an exponent of about 1.8.

Unfortunately, it is difficult to justify physically all of the assumptions described above, particularly the assumption that the star formation rate varies as a power of the gas density. Possibly some justification for this type of assumption could result if supernova explosions provide a sufficient amount of energy to strongly heat the gas in a protogalaxy and thus control the star formation rate through a negative feedback effect. As has been shown by Larson (1974b), there is a critical star formation rate proportional to the 1.76 power of the gas density such that if the actual star formation rate exceeds this critical value all of the supernova energy goes into heating the gas, whereas for lower star formation rates most of the supernova energy is radiated away and the heating effect is considerably reduced. At present the various numerical quantities involved are too uncertain to permit a definite conclusion concerning the possible effect of supernovae in controlling the star formation rate in collapsing protogalaxies, but it is intriguing that the *form* of the star formation rate predicted on this basis agrees with what is needed to produce realistic collapse models.

A simpler approach to specifying the gaseous dissipation and star formation rates has been adopted in some of the more recent model calculations of Larson (1974a). Since it is difficult to justify the assumption that neighboring gas elements or clouds in a protogalaxy have independent random motions, and since a protogalaxy may in reality contain systematic streaming motions on various scales up to that of the galaxy itself, it is possible that the dissipation time scale is more closely related to the time scale for collisions between such large (galaxy-size) gas streams, i.e. to the dynamical or free-fall time for the system. A simple example of a situation where this would be the case is the free infall of gas from the outer part of a protogalaxy into the dense inner region, where it collides with other gas and is stopped. Calculations have therefore been made with the assumption that the dissipation time is proportional to the local free-fall time, the constant of proportionality being left as a free parameter.

The star formation rate may also depend on the dynamical time scale: for example, if star formation is caused primarily by the compression of gas in the shock fronts produced when large scale gas streams collide, then the time scale for transformation

of gas into stars is related to the time scale for processing of the gas through these shock fronts, which in turn is related to the dynamical time scale for the system. Accordingly it has been assumed by Larson (1974a) that the star formation time scale is proportional to the local free fall time, with an undetermined constant of proportionality of order unity. The assumption that star formation is related to the large scale dynamics of a galaxy receives some support from the observation that star formation generally occurs in spiral arms, which evidently reflect large scale patterns in the dynamics of the gas in a galaxy, being currently interpreted as shock fronts accompanying spiral density waves. A similar type of assumption for the star formation rate was found by Oort (1974) to be successful in explaining the radial variation of the gas density and star formation rate in spiral galaxies. Undoubtedly the dependence of the star formation rate on physical parameters is much more complicated than a simple dependence on the dynamical time scale, since many small scale processes, including the effects of massive stars and supernovae, must almost certainly play a role. However, in view of our still very limited understanding of all the relevant processes, it is probably not worthwhile at present to introduce a much more elaborate representation of the star formation rate than the simple possibilities mentioned above.

### 5. Results for Spherical Gas-Dynamical Models

Here we shall briefly review the properties of the spherical collapse models calculated by Larson (1969, 1974a) on the basis of the assumptions described above. In each case the calculations begin with a uniform gaseous protogalaxy, and the system begins to collapse nearly in free fall, continually transforming gas into stars as it does so. As expected, the gas component of the protogalaxy condenses more and more strongly toward the center as the collapse proceeds; thus successive generations of stars are on the average formed in regions of successively smaller size and higher density, and the galaxy is built up from the outside in. Since the infalling gas becomes progressively more enriched in heavy elements as it flows toward the center, a composition gradient is naturally and inevitably established in models of this type.

As was found by Larson (1974a), it is possible with plausible choices of the parameters to obtain models which reproduce quite closely the surface brightness distribution of the E galaxy NGC 3379 measured photoelectrically by Miller and Prendergast (1962), and also the very similar mean  $I(r)$  profile for 14 elliptical galaxies as determined by King (1975). It is noteworthy that, unlike the stellar collapse models, these gas-dynamical models reproduce not only the surface brightness distribution in the envelope but also the dense cores of E galaxies. However, while the possibility of obtaining detailed agreement with the surface brightness observations shows that the gas-dynamical models are at least consistent with these observations, it is still not possible to discriminate between different models of this type; for example, when the parameters are appropriately adjusted, the differences between predictions based on the 'power law' star formation rate and those based on the 'dynamical time scale' star formation rate are smaller than the differences between the observed

brightness profiles of different E galaxies. When we recall that the stellar infall models of Gott are also able to reproduce the observed envelope brightness profiles, it is evident that we cannot on this basis alone decide in detail between the various mechanisms that may be responsible for establishing the structure of E galaxies, particularly in the outer regions.

A further point of comparison with the observations is in the predicted radial variation of metal abundance  $Z(r)$ . Unfortunately, the available observations refer only to the innermost 1 or 2 kpc of elliptical galaxies, and cannot be compared very quantitatively with the predictions because of the lack of accurate calibrations. The models of Larson (1974a) all predict essentially the same behavior for  $Z(r)$  in the innermost 2 kpc of elliptical galaxies, and this predicted  $Z(r)$  is qualitatively in agreement with the observations; this provides strong evidence for the formation of at least the inner regions of E galaxies by a gaseous infall process. Different gas-dynamical models make different predictions for  $Z(r)$  in the outer parts of E galaxies, but here there are no direct observations, and the indirect measure of  $Z(r)$  provided by the UBV colors is unfortunately uncertain, due to disagreements between different observers. One of the most important ways of discriminating between models for the formation of elliptical galaxies will be to obtain measurements of the metal abundances in the outer envelopes of E galaxies.

The predicted stellar velocity dispersions are even more difficult to compare with observations, owing to the complete lack of measurements except at the centers of a number of E galaxies. If it becomes possible to measure velocity dispersions in the envelopes of E galaxies, this will provide another important way of discriminating between models.

Although this result cannot yet be related to observations of elliptical galaxies, the gas-dynamical models provide a 'bonus' in that they yield a prediction of the frequency distribution  $N(Z)$  of stellar metal abundances at each radius. An important result is that in the innermost few hundred parsecs of the models most of the stars at any given radius have approximately the same metal abundance, and relatively metal poor stars are quite rare. This result comes about because the continuing infall of gas in this region allows a nearly time-independent  $Z(r)$  distribution to be established, and most of the stars are formed after this approximate steady state has been attained. The predicted frequency distribution of stellar  $Z$  values near the center of these models resembles that observed among stars in the solar neighborhood, and this suggests that the disc of our Galaxy was formed by a continuing gas infall process, like the central regions of these models. The results described below for rotating models also appear to point toward an infall picture for the formation of discs in galaxies.

The effect of supernova explosions on the early evolution of these models has been estimated by Larson (1974b). It is found that supernova explosions can deposit a substantial amount of energy in the gas component of a forming galaxy, causing some of this gas to be driven out of the galaxy in a 'galactic wind'. The importance of supernova-driven gas loss increases with decreasing galactic mass, and in a forming galaxy

whose mass is less than  $\approx 10^9\text{--}10^{10} M_{\odot}$ , much or all of the residual metal-enriched gas which would otherwise have condensed at the center to form a dense metal-rich nucleus may instead be expelled from the system. Model calculations incorporating this effect predict, in qualitative agreement with the observations, that the less massive galaxies should show less prominent nuclei and less prominent composition gradients; furthermore, a variation of overall metal abundance with galactic mass is predicted which agrees qualitatively with the observed  $Z(M)$  relation for elliptical galaxies. Thus it appears that the effect of supernova explosions on the gas dynamics can explain, at least qualitatively, the observed mass dependence of the structure and composition of E galaxies. This result provides additional evidence for the importance of gas-dynamical processes in the formation of E galaxies.

We note that the effect of supernovae in expelling most of the gas from small protogalaxies and perhaps disrupting them completely may be of even more fundamental importance in influencing the mass spectrum of galaxies. Peebles (1974a) has noted that the spatial distribution of galaxies shows no preferred scale of clustering, and suggests that the galaxies themselves may simply represent the smallest surviving scale of clustering, appearing as single objects rather than clusters of smaller objects because gas-dynamical effects have destroyed the smaller scales of clustering. One such 'gas-dynamical effect' might be the effect of supernova-driven gas loss in destroying most of the smaller systems. Suppose, for example, that a galaxy of typical mass  $\sim 10^{11} M_{\odot}$  begins as a cluster of smaller systems (eg., proto-globular clusters) of mass  $\lesssim 10^9 M_{\odot}$ ; then supernova explosions will expel most of the initial gas from these smaller systems, and many of them may be completely disrupted. However, because of its higher escape velocity, the larger system of mass  $\sim 10^{11} M_{\odot}$  may retain much of the debris from the smaller systems and thus survive as a single galaxy. This argument would predict that most of the mass in galaxies should be in systems with masses greater than  $\approx 10^9\text{--}10^{10} M_{\odot}$ , which appears to be consistent with the observations.

## 6. Rotating Gas-Dynamical Models

In view of the general success of the spherical gas-dynamical models in explaining those properties of E galaxies that are not related to their rotation, it is of considerable interest to see whether with the addition of rotation such models can also explain such properties as the ellipticity profiles  $\varepsilon(r)$  of E galaxies and the presence of disc components in spiral and S0 galaxies. Calculations with such rotating models, intended mainly to provide more realistic models for flattened and rotating E galaxies, are currently in progress, and the most significant (if still somewhat tentative) results so far obtained with these models will be briefly reviewed here; detailed results will be published in a later paper.

A model calculation must begin with a specification of the initial conditions for a collapsing protogalaxy; in particular, we are concerned here with the initial state of rotation in a system with significant angular momentum. We shall not attempt to discuss the much-debated question of the origin of the angular momentum of galaxies,

but merely mention some of the possibilities which have been suggested. Peebles (1974b) has pointed out that it is difficult to understand how the angular momentum of galaxies could originate directly from primeval vorticity in the early universe, and has advocated the point of view that protogalaxies acquired their angular momentum through the action of tidal torques (e.g., Peebles, 1971). Other authors (e.g., Icke, 1973; Binney, 1974) have attempted to understand the rotation of galaxies as having resulted from motions generated in a collapsing protocluster of galaxies. Silk and Lea (1973) have pointed out that inelastic collisions and coalescence of randomly moving primordial gas clouds or protogalaxies could plausibly produce systems with the angular momenta characteristic of galaxies. Finally, there is the possibility that large scale turbulent motions could be generated by processes such as supernova-driven winds or quasar activity occurring at about the time of galaxy formation. Evidently, there is no lack of possible mechanisms for generating rotational motions, and it is even possible that different galaxies may have acquired their angular momenta in different ways. However, all of the proposed mechanisms seem likely to leave a protogalaxy with a more or less irregular structure and state of motion, and no simple prescription can be given for the initial rotational motion of a collapsing protogalaxy.

If it is indeed the case that protogalaxies are chaotic in structure and motion, then the observed regularities in the structure of galaxies must have been produced by processes occurring during the collapse. Thus one may hope that, if the model calculations realistically describe the physical processes occurring during the collapse, one can still obtain reasonably realistic and representative models by starting with idealized initial conditions such as a uniform, spherical, and uniformly rotating protocloud. At any rate, until the details of galaxy formation are better understood, there seems to be little point at present in exploring any but the simplest models, since more complex models would contain more parameters than we would know what to do with. In this spirit, all of the model calculations so far undertaken have assumed an initially uniform and uniformly rotating protogalaxy.

Calculations of the collapse of rotating protogalaxies have been made using a two-fluid generalization (to represent both the gas and stellar components) of the axisymmetric computer code used by Larson (1972b) to calculate the collapse of a rotating cloud. The first calculations to be undertaken consisted essentially of repeating the spherical collapse calculations with the addition of rotation but without otherwise modifying the simple hydrodynamic treatment previously used for the gas and stars in a protogalaxy; this means in effect that the basic flow is assumed to be laminar and inviscid, so that each fluid element conserves its angular momentum during the collapse. Models were calculated using both of the approaches described previously for estimating the gaseous dissipation and star formation rates. However, all of these calculations produced qualitatively the same unsatisfactory result. The first stars to form are distributed in a spheroidal halo, as might be expected, but as the residual gas continues to condense inward it develops a more and more flattened distribution; this is reflected in the ellipticity  $\varepsilon(r)$  of the resulting stellar system, which

increases strongly and monotonically toward decreasing radii. In fact, the central part of the resulting system consists of a remarkably flat, uniform, and uniformly rotating disc, with a radius of the order of some kpc but without any central nuclear condensation. This disc would certainly be unstable to non-axisymmetric modes (see the contributions of Bardeen and Hohl to this symposium) and might eventually evolve into a centrally condensed structure resembling those observed in spiral or barred spiral galaxies, but we would still be left with the problem of understanding the formation of *elliptical* galaxies with realistic ellipticity profiles that do not rise sharply to very high values at small radii.

The trouble with the inviscid hydrodynamical models described above is almost certainly that the neglect of viscosity effects and angular momentum transport is not justified. Indeed, the basic conceptual picture of a highly turbulent protogalaxy implies a large associated 'turbulent viscosity' (e.g., Saslaw, 1971); for example, if a protogalaxy is imagined to consist of discrete randomly moving gas clouds, the system of gas clouds will possess a kinetic viscosity just like a classical molecular gas, except that here we have very large 'molecules' (i.e., the gas clouds) and very long mean free paths, so that the viscosity effect is correspondingly larger. In general, whatever the detailed nature of the gas motions and the dissipative processes, we might expect that associated with the gaseous dissipation will be a viscosity effect which will tend to transport angular momentum, and that the time scale for significant transport of angular momentum will be of the same order as the dissipation time scale.

Given a cloud model for the turbulent gas in a protogalaxy, it is straightforward to calculate the kinetic viscosity coefficient in terms of the mean free path of the gas clouds, and to substitute this into the appropriate generalized form of the classical Navier-Stokes equations describing the flow of a viscous fluid. If a discrete cloud model is not used but the dissipation time scale is instead related directly to the free fall time, as in the models of Larson (1974a), the viscosity coefficient can instead be expressed in terms of the mean free time of the colliding gas elements, which is approximately the same as the dissipation time. Both of these approaches have been used in the rotating collapse models described below, and they yield qualitatively similar results, albeit with significant quantitative differences. In either case, however, the use of the Navier-Stokes equations, in which the viscous stress term is proportional to the gradient of the angular velocity, provides only a crude first approximation to reality, since in the present situation the mean free paths of the fluid elements are not small compared with the size of the system, as assumed in classical kinetic theory. As an example of a system with very large mean free paths, consider an equilibrium stellar disc where, even in the presence of differential rotation, there is no viscous stress and no transfer of angular momentum, so that the Navier-Stokes equations are clearly not applicable. A more correct treatment of the viscosity effect in a turbulent protogalaxy would require solving explicitly for the relevant moments of the velocity distribution, and this would substantially increase the already great computing effort required. An attempt was made to circumvent this extra calculation

by approximating the relevant moments of the velocity distribution, and this produced results qualitatively similar in most respects to those obtained with the Navier-Stokes equations, but possibly describing more realistically the formation of a highly flattened subsystem in which the viscosity is expected to be relatively small; these results will be described briefly below in Section 7. In all cases, the effect of the viscosity is to transport angular momentum *outward*.

When viscosity effects are incorporated in any of the ways described above, the collapse calculations yield much more realistic models for elliptical galaxies. The outward transport of angular momentum allows the gas component to develop a highly centrally condensed distribution without becoming as highly flattened near the center as in the calculations without viscosity. The radial variation of the resulting stellar density distribution is similar to that previously obtained for the spherical models, and the ellipticity  $\varepsilon(r)$  now shows much more moderate variations with radius, as is observed in elliptical galaxies. The predicted rotation curves always rise rapidly to a maximum a few kpc from the center and then fall off gradually with increasing radius, closely resembling in form the rotation curves measured for early-type spiral galaxies by Roberts and Rots (1973).

It is worthwhile to consider in more detail the ellipticity profiles  $\varepsilon(r)$  predicted by the models, since at present they provide the main point of comparison between these rotating models and the observations. In general, we expect the ellipticity profile  $\varepsilon(r)$  to depend both on the initial angular momentum of the protogalaxy and on the effect of viscosity in redistributing angular momentum during the collapse. From the results of calculations made with different values of the parameters and different ways of treating the viscosity effect, it appears that the ellipticity of the outer parts of the resulting system depends mainly on the initial angular momentum of the protogalaxy, whereas the ellipticity of the inner regions depends primarily on the way in which viscosity redistributes angular momentum during the collapse. When the viscosity coefficient is calculated from the cloud model of Larson (1969) and the generalized Navier-Stokes equations are used, the resulting ellipticity profile is nearly constant at large radii and decreases near the center, a behavior which is not unusual for E galaxies. If the viscosity coefficient is reduced by a factor of two the resulting  $\varepsilon(r)$  profile is moderately peaked at intermediate radii, a characteristic which again is fairly common among E galaxies. Models which are not based on the Navier-Stokes equations but on a particular simple approximation to the moments of the velocity distribution produce  $\varepsilon(r)$  profiles which instead of decreasing at small radii tend to increase near the center.

In view of the dependence of  $\varepsilon(r)$  on details of the gas dynamics which are not well understood, detailed agreement between predicted and observed  $\varepsilon(r)$  profiles should probably not be expected and has not been sought, but it seems significant that a variety of more or less realistic profiles can be obtained with plausible values for the angular momentum and viscosity parameters. This result provides support for the general idea that turbulent gas dynamics plays an important role in the formation of galaxies, and offers a possibility of explaining the considerable variations among



the observed  $\varepsilon(r)$  profiles of elliptical galaxies. We have already noted that it is not to be expected that all protogalaxies have the same initial angular momentum distribution; nor is it to be expected that the internal dynamics would be the same in all cases, since different amounts of 'turbulence' could lead to differences in the effectiveness of viscous transport of angular momentum during the collapse. Thus the variations in observed  $\varepsilon(r)$  profiles could reflect differences in the initial conditions as well as in the dynamics of the collapse. When more data become available, it will be interesting to see whether the variations in  $\varepsilon(r)$  correspond to differences in the kinematic properties of E galaxies; if so, the E galaxies may not be quite so simple and homogeneous a class of objects as has usually been supposed.

At the least, it seems established by the results described in this section that it is not valid to assume, as has often been done, that each fluid element conserves its angular momentum during the collapse of a protogalaxy; the structure of E galaxies can only be explained if some redistribution of angular momentum takes place during the collapse. Consequently, the final distribution of angular momentum in a galaxy depends not only on the initial conditions but also on the dynamics of the collapse and formation process, and inferences that galaxies form from protoclouds with the same distribution of angular momentum are probably not justified.

## 7. Formation of Galactic Discs

As yet, there have been no model calculations which realistically represent the formation of a spiral galaxy with a prominent disc component, so our final remarks concerning the formation of galactic discs must necessarily be somewhat conjectural. Nevertheless, it appears that the rotating gas-dynamical models described above provide some useful information or constraints on the possible ways in which discs might form.

In all of these rotating models, the residual gas remaining during the later stages of the galaxy formation process gradually settles into a more and more flattened distribution, finally forming a thin, centrifugally supported disc in the equatorial plane. (Near the center of the disc, the velocity dispersion of the gas remains large compared with the rotational velocity, so that the density distribution there remains nearly spheroidal and not highly flattened.) Thus during the later stages of the galaxy formation process the stars are formed in a more and more flattened distribution, finally ending with the formation of a thin disc-like 'Population I' component in the equatorial plane. This flat component contains only a small fraction ( $\lesssim 10\%$ ) of the total mass, but in the models with higher angular momentum its presence begins to have a noticeable effect on the shape of the stellar iso-density contours, causing them to become slightly pointed in the equatorial plane. This effect is exactly what is observed in the isophotes of the more flattened E galaxies: for systems flatter than about E5, the isophotes become noticeably pointed at the ends of the major axis (e.g., Liller, 1966; King, 1975). The model results are qualitatively consistent with these observations, so that both models and observations suggest that a disc com-

ponent begins to appear in the E galaxies with higher angular momentum, i.e. those flatter than about E5. Indeed, Liller (1966) suggests that all such systems should be classified as S0 rather than E galaxies.

It is interesting to remark in passing that the presence of a flattened subsystem in the models is always much more evident in the composition distribution and the kinematics than in the density distribution. Just as is the case for the Population I component of our Galaxy, the incipient disc component found in the models has a higher metal abundance, larger rotational velocity, and smaller velocity dispersion than the halo component, and all of these quantities show quite marked gradients in the direction perpendicular to the equatorial plane. The contours of equal metal abundance, for example, are quite remarkably flattened toward the equatorial plane, much more so than the stellar iso-density contours. This result suggests that the presence of disc components in flattened E and S0 galaxies may be more conspicuous in their composition and kinematics than in the light distribution, and suggests observations to look for these effects. Observational evidence for a composition gradient perpendicular to the plane of the edge-on S0 galaxy NGC 4762 has been presented by Freeman at this symposium.

From the results described above, we can understand how a disc component might form in a rotating galaxy from residual gas left over after the formation of a spheroidal halo. However, in none of the models so far calculated does the disc component exceed a minor fraction ( $\lesssim 10\%$ ) of the total mass. Recycled gas from halo stars, not yet explicitly incorporated in the models, would contribute to the disc mass but would not add more than another  $\simeq 10\%$  of the total mass. If only this small amount of residual gas is involved in the formation of a disc, it is not even certain that in reality a disc would be formed at all, since this residual gas can easily be swept out of the galaxy by a supernova-driven galactic wind (Larson, 1974b), or possibly even by an intergalactic wind. After some experimentation with the parameters, we have tentatively concluded that simple collapse models of the type described above cannot explain the formation of galaxies with prominent disc components, the reason for this being simply that if the models are to be capable of explaining *elliptical* galaxies, then star formation must be reasonably efficient during the initial collapse and most of the gas must be transformed into stars before there has been time for it to settle into a very flat disc.

Perhaps it is reasonable that the simplest model, i.e. one that begins by assuming a uniform, spherical, and uniformly rotating protocloud should be capable of explaining only the simplest galaxies, i.e. the elliptical galaxies; this is probably all that we are entitled to expect. However, since these models do produce incipient disc components, they suggest how a less idealized model might in general yield a galaxy with a more prominent disc component. Basically, what is required to form a more massive disc is a larger amount of leftover primordial gas which does not participate in the formation of a spheroidal halo but condenses later into a disc. If the halo component of a galaxy is formed by the free fall collapse of a roughly spherical pre-galactic density perturbation, as in the above models, then we require that there

be additional ambient primordial gas around the central perturbation which does not participate in the initial free fall collapse but continues to fall in later for a more extended period of time. Such a situation is entirely plausible, and is in fact predicted by calculations which take into account the effects of cosmological expansion, as was already pointed out in connection with the infall models discussed in Section 3.

Gott (1975) assumed that the ambient infalling material would all be efficiently converted into stars, in which case it would contribute to building up the halo of the galaxy but would not form a disc. However, it is possible that once a halo of significant mass has already formed, star formation would be inhibited in the gas which continues to fall in later, especially if the density of the infalling gas is lower than the average density of the system; then the tidal disruptive forces due to the rest of the system will dominate over local self-gravitational forces in the gas, and local collapse of the gas to form stars will presumably be inhibited (unless large density enhancements are already present in the gas.) The gas will then continue to swirl around and dissipate its energy until it eventually settles into a disc, and star formation can finally take place once the local gas density in the disc has become higher than the mean density of the system (or, more precisely, once the Goldreich and Lynden-Bell (1965) criterion for local gravitational instability in a gas disc is satisfied.) Depending on the amount of ambient gas which continues to fall into a forming galaxy in this way and condense into a disc, it would seem possible to eventually build up a disc of almost any mass relative to the already formed halo component.

Thus, while no realistic models have yet been computed and many of the details remain to be clarified, there appears to be no difficulty in principle in understanding how galaxies might form with varying proportions of disc mass to halo mass; the simple collapse of a roughly uniform, spherical protocloud yields a nearly pure halo system, i.e. an E galaxy, whereas a more complicated formation process involving the accretion of varying amounts of leftover primordial material will in general result in a spiral or S0 galaxy. The models described above, together with the results described by Brahic at this symposium, indicate that the halo component of a galaxy forms relatively rapidly (within  $\simeq 10^9$  yr), whereas the formation of a disc is a more gradual process; several rotation periods are required for the gas to settle down into a flat disc, and the disc may continue to be built up over an extended period of time by continuing infall of matter. Indeed, if there is a significant amount of uncondensed intergalactic matter remaining at the present time, the discs of spiral galaxies may still be slowly gaining mass by accretion (Oort, 1970; Larson, 1972a). We note that the present picture for the formation of galactic discs is consistent with the theoretical indications (Ostriker and Peebles, 1973; Bardeen, 1975) that the discs of spiral galaxies may be embedded in halos of at least comparable mass, since the present picture requires a halo of significant mass to be present initially; we cannot yet, however, draw any quantitative conclusions about halo masses on the basis of theories of galaxy formation.

Finally, we note that if galactic discs are really built up over a period of time by the infall of leftover primordial gas, as suggested here, then this picture offers little basis

for explaining the characteristic approximately exponential structure of galactic discs as a consequence of the initial conditions, since we expect no particular regularities in the initial spatial distribution or angular momentum distribution of the matter that eventually makes up the disc. Thus, to the extent that galactic discs show any regularities in their structure, such regularities must result from the dynamical evolution of the discs themselves. Such dynamical evolution is certainly expected to occur as a result of the bar and/or spiral mode instabilities which occur in a disc when its velocity dispersion is small (Bardeen, 1975; Hohl, 1975). Calculations of the dynamics of discs show that an unstable disc characteristically develops a central bar with trailing spiral arms, and that the associated gravitational torques transport angular momentum outward and cause the disc to develop a more centrally condensed structure. According to Hohl (1971), the structure finally attained is approximately an exponential disc. Even the observed presence of spiral structure in galaxies implies the presence of gravitational torques which will tend to transport angular momentum outward and slowly change the structure of the disc. It is probably in terms of such processes that the structure of galactic discs is ultimately to be understood, rather than in terms of the initial conditions or the details of the collapse process.

### 8. Epilogue

It should be evident from the foregoing review that theorists are still groping in considerable darkness in attempting to understand the formation of galaxies. Major uncertainties remain in our understanding of the basic physical processes involved, particularly with regard to gas-dynamical processes and star formation. Until we have a better understanding of the efficiency of star formation in collapsing protogalaxies, for example, it will be difficult to determine the relative importance of gas-dynamical and stellar-dynamical processes, and it will be difficult to decide between the different possible models which can be constructed to give a good fit to the surface brightness distribution of galaxies like NGC 3379. It will be of great importance to obtain more observational data on the properties of galaxies, particularly the composition distribution and the kinematics in regions other than the nuclei of elliptical galaxies, since such data should eventually provide a powerful means of discriminating between models.

Another avenue of progress will be to look for observable phenomena which are similar to, or perhaps even identical to, those which take place during the formation of galaxies. Thus the processes of gas dynamics and star formation currently taking place in the disc of our galaxy should have a good deal in common with processes which took place at a much earlier stage of evolution, and it is clear that a better understanding of current star formation processes will be most useful, if not prerequisite, for understanding star formation in forming galaxies.

The most intriguing and perhaps most promising possibility is that of actually observing galaxy formation processes taking place in relatively nearby systems. The theoretical models strongly suggest the possibility that in some cases residual proto-

galactic gas may continue to condense into a galaxy long after the initial collapse process is completed and thus help sustain continuing star formation, either steadily or in occasional bursts. Some well known galaxies which may have experienced recent gas infall and bursts of star formation are M82, which contains many relatively young stars and a large amount of chaotically distributed gas and dust; NGC 5253, an elliptical galaxy which contains several knots of gas and young stars, and which has recently attained notoriety as a prolific producer of supernovae (an elliptical galaxy in the process of formation?); and NGC 5128, a giant spherical galaxy surrounded by a conspicuous dark band of dust, gas, and young stars (a disc system in the process of formation?) Detailed studies of such systems should eventually increase substantially our understanding of gaseous condensation and star formation processes in forming galaxies.

### References

- Alladin, S. M., Potdar, A., and Sastry, K. S.: 1975, this volume, p. 167.  
 Arp, H. C.: 1965, *Stars and Stellar Systems* 5, 401.  
 Bardeen, J. M.: 1975, this volume, p. 297.  
 Bertola, F. and Capaccioli, M.: 1975, this volume, p. 373.  
 Binney, J.: 1974, *Monthly Notices Roy. Astron. Soc.* **168**, 73.  
 Brosche, P.: 1970, *Astron. Astrophys.* **6**, 240.  
 Eggen, O. J., Lynden-Bell, D., and Sandage, A. R.: 1962, *Astrophys. J.* **136**, 748.  
 Faber, S. M.: 1973, *Astrophys. J.* **179**, 731.  
 Freeman, K. C.: 1970, *Astrophys. J.* **160**, 811.  
 Goldreich, P. and Lynden-Bell, D.: 1965, *Monthly Notices Roy. Astron. Soc.* **130**, 125.  
 Gott, J. R.: 1973, *Astrophys. J.* **186**, 481.  
 Gott, J. R.: 1975, *Astrophys. J.* (in press).  
 Hartwick, F. D. A. and Sargent, W. L. W.: 1974, *Astrophys. J.* **190**, 283.  
 Hohl, F.: 1971, *Astrophys. J.* **168**, 343.  
 Hohl, F.: 1975, this volume, p. 349.  
 Icke, V.: 1973, *Astron. Astrophys.* **27**, 1.  
 Janes, K. A.: 1975, in press.  
 King, I. R.: 1966, *Astron. J.* **71**, 64.  
 King, I. R.: 1975, in preparation.  
 Larson, R. B.: 1969, *Monthly Notices Roy. Astron. Soc.* **145**, 405.  
 Larson, R. B.: 1972a, *Nature* **236**, 21.  
 Larson, R. B.: 1972b, *Monthly Notices Roy. Astron. Soc.* **156**, 437.  
 Larson, R. B.: 1974a, *Monthly Notices Roy. Astron. Soc.* **166**, 585.  
 Larson, R. B.: 1974b, *Monthly Notices Roy. Astron. Soc.* **169**, 229.  
 Larson, R. B. and Tinsley, B. M.: 1974, *Astrophys. J.* **192**, 293.  
 Lecar, M. and Cohen, L.: 1972, in M. Lecar (ed.), *Gravitational N-Body Problem*, D. Reidel Publ. Co., Dordrecht-Holland, p. 262.  
 Liller, M. H.: 1960, *Astrophys. J.* **132**, 306.  
 Liller, M. H.: 1966, *Astrophys. J.* **146**, 28.  
 Lynden-Bell, D.: 1967, *Monthly Notices Roy. Astron. Soc.* **136**, 101.  
 McClure, R. D.: 1969, *Astron. J.* **74**, 50.  
 Miller, R. H. and Prendergast, K. H.: 1962, *Astrophys. J.* **136**, 713.  
 Oort, J. H.: 1970, *Astron. Astrophys.* **7**, 381.  
 Oort, J. H.: 1974, in J. R. Shakeshaft (ed.), 'The Formation and Dynamics of Galaxies', *IAU Symp.* **58**, 375.  
 Ostriker, J. P., and Peebles, P. J. E.: 1973, *Astrophys. J.* **186**, 467.  
 Ostriker, J. P., and Thuan, T. X.: 1975, *Astrophys. J.* (in press).  
 Peebles, P. J. E.: 1971, *Astron. Astrophys.* **11**, 377.  
 Peebles, P. J. E.: 1972, *Comments Astrophys. Space Phys.* **4**, 53.

- Peebles, P. J. E.: 1974a, *Astrophys. J. Letters* **189**, L51.
- Peebles, P. J. E.: 1974b, in J. R. Shakeshaft (ed.), 'The Formation and Dynamics of Galaxies', *IAU Symp.* **58**, 55.
- Peebles, P. J. E. and Dicke, R. H.: 1968, *Astrophys. J.* **154**, 891.
- Roberts, M. S. and Rots, A. H.: 1973, *Astron. Astrophys.* **26**, 483.
- Sandage, A. R.: 1972, *Astrophys. J.* **176**, 21.
- Saslaw, W. C.: 1971, *Monthly Notices Roy. Astron. Soc.* **152**, 341.
- Saslaw, W. C.: 1972, *Astrophys. J.* **177**, 17.
- Shields, G. A.: 1974, *Astrophys. J.* **193**, 335.
- Silk, J. I. and Lea, S. M.: 1973, *Astrophys. J.* **180**, 669.
- Spinrad, H., Gunn, J. E., Taylor, B. J., McClure, R. D., and Young, J. W.: 1971, *Astrophys. J.* **164**, 11.
- Spinrad, H., Smith, H. E., and Taylor, B. J.: 1972, *Astrophys. J.* **175**, 649.
- Spitzer, L.: 1971, in D. J. K. O'Connell (ed.), *Nuclei of Galaxies*, American Elsevier, New-York, p. 443.
- Toomre, A. and Toomre, J.: 1972, *Astrophys. J.* **178**, 623.
- Tremaine, S. D., Ostriker, J. P., and Spitzer, L.: 1975, *Astrophys. J.* **196**, 407.
- Van Houten, C. J.: 1971, *Bull. Astron. Inst. Neth.* **16**, 1.
- Wilson, C. P.: 1975, *Astron. J.* **80**, 175.

## DISCUSSION

*Lynden-Bell*: I am interested in the actual value of your kinematic viscosity and its ratio to the typical angular momentum per unit mass of a star. Is this ratio  $10^{-3}$  or what?

*Larson*: I'm not sure exactly how the kinematic viscosity that I assumed compares with the angular momentum per unit mass of the stars, but in any case it is a large effect: in terms of the mean free path of the colliding fluid elements, the mean free path is of the same order as the radius, and in terms of the time scale for significant redistribution of angular momentum, this time scale is comparable with the collapse time.

*Hunter*: I agree that some mechanism for the transfer of angular momentum is very important for this problem. What justification can you give for supposing the relevant mechanism to be a turbulent viscosity, rather than gravitational torques that arise when axisymmetry is broken? Other work tells us of the importance of bar instabilities, and these could transfer significant angular momentum. Also, do you have a turbulent viscosity for the stars as well as for the gas?

*Larson*: Perhaps I should clarify first that I use the terms 'turbulence' and 'turbulent viscosity' merely as convenient labels, without implying any close connection with turbulence as understood (?) in other contexts. The physical picture that I have in mind is that a protogalaxy consists of some collection of gas elements, in general having a chaotic distribution and state of motion. Then, as was pointed out some time ago by McCrea in connection with his floccule theory of star formation, the gas elements which happen because of their random motions to have little angular momentum will fall toward the center, where they will remain if their motions are rapidly dissipated. On the other hand, the elements with large angular momentum will remain in the outer part of the system, and the net result will be a redistribution of angular momentum in which angular momentum is transported outward. It seems to me that any system with random motions will show this effect, and this is what I have attempted to simulate in my fluid dynamical calculations. This separation of the low- and high-angular momentum material will be important during the initial collapse, but probably not after the gas has settled into a flattened layer in approximate centrifugal equilibrium. Then non-axisymmetric instabilities to bar and/or spiral modes would likely set in and might well lead to further important outward transport of angular momentum.

It might even be that gravitational torques could be important during the initial collapse, since differential rotation during the collapse could wind the density inhomogeneities into spirals, associated with which would be gravitational torques. However, I don't know how this effect could be estimated or taken account of in my models.

The models that I have calculated so far have no viscosity for the stars, although such an effect would probably exist also for the stars. My guess is that the viscosity effect for the stars would be less important than for the gas.

*Brosche*: What are the assumptions on star formation in your models?

*Larson*: I tried two different approaches to treating the star formation rate: one was to relate the star formation rate to a power of the gas density, in which case a star formation rate proportional to about the

1.8 power of the gas density is found to give models in good agreement with observations; and the second approach, possibly better in that it involves fewer free parameters, was to assume that the time scale for transformation of gas into stars is proportional to the dynamical time scale for the system. The latter approach may be roughly justified on physical grounds if star formation is initiated by large scale dynamical processes such as the compression of the gas in shock fronts caused by collisions between gas clouds or streams. This point of view would be consistent, for example, with current attempts to explain star formation in spiral shock fronts in spiral galaxies.

*Freeman*: Turbulence plays a fundamental part in your picture and your results. Do you have any physical external reason for invoking it?

*Larson*: The strongest reason for assuming that the gas had large chaotic motions is perhaps an 'internal' one and is the fact that halo stars in our Galaxy are observed to have a very large velocity dispersion (including many retrograde orbits), and this is most reasonably explained in terms of large random motions in the gas from which the stars formed. The 'external' reasons are harder to make precise, but the various mechanisms which have been proposed for imparting large scale motions (eg. rotation) to galaxies all seem likely to produce a more or less chaotic state of motion in a protogalaxy. This is true in varying degrees of tidal torques, primordial turbulence, motions generated during the collapse of protoclusters, and random motions generated (eg. by supernova explosions) during the formation of galaxies. If any (or all) of these processes operate, it would seem rather difficult to produce a protogalaxy whose internal motion is extremely regular, such as pure uniform rotation with no random motions superimposed.

*King*: The  $Z(r)$  observations that you mentioned were made with an old generation of equipment. A new generation now exists that will be able to extend the curve into the region that you want.

*Miller*: Could you provide a guide to the observers as to what observations are most effective in discriminating among various models? It appears that the models differ most strongly in the faint outer regions, and that details near the center are less sensitive.

*Larson*: It is unfortunately true that some of the important observational tests to discriminate between models involve the properties of the faint outer envelopes of galaxies. For example, if it turns out that the metal abundance is constant in the outer envelopes of elliptical galaxies, then a model in which star formation is rapid compared to the collapse time (such as Gott's models) might be viable, whereas if the metal abundance  $Z$  continues to decrease with  $r$  even in the halo, as seems to be indicated by recent observations of NGC 3379 by Burkhead and Kalinowski, it seems clear that a model involving the collapse and simultaneous metal enrichment of a gas cloud is required. Also, different models make somewhat different predictions for the kinematics of the outer regions of galaxies, while they may predict much the same thing in the central regions. In general, it is my impression that observations of the composition and kinematics of galaxies and especially of their variation with radius may provide more information that is useful for discriminating between models than observations of the light distribution.