Mon. Not. R. astr. Soc. (1976) **176,** 31–52.

MODELS FOR THE FORMATION OF DISC GALAXIES

Richard B. Larson

Yale University Observatory, Box 2023 Yale Station, New Haven, Connecticut 06520, USA and IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA

(Received 1976 February 13)

SUMMARY

Previous calculations of the collapse of protogalaxies with rotation and axial symmetry are here extended to galaxies possessing both a spheroidal component, and a disc component containing a substantial fraction of the total mass. The predicted disc:bulge ratio depends mostly on the assumed star formation rate and how it varies with time: the formation of a spheroidal component requires an early phase of rapid star formation, but the formation of a disc requires a later stage of much slower star formation which allows residual gas to settle to a disc before forming stars. The rapid mode of star formation may involve strongly clumped gas and the slow mode relatively diffuse gas, and the degree of clumping may depend on the strength of turbulence or collisions in the gas; the Hubble sequence can then be understood at least in part as resulting from differences in the initial density or velocity dispersion in protogalaxies. Other general predictions of the models include stellar metallicity distributions and kinematic properties that are in qualitative agreement with those of the stellar populations in our own and other spiral galaxies, and a long time scale for the formation of the outer part of the disc, implying a significant gas content and star formation rate even after 10^{10} yr.

1. INTRODUCTION

The most fundamental fact of the structure and classification of galaxies is the existence of two major morphological types, namely the elliptical and spiral galaxies. Elliptical galaxies appear to be basically spheroidal in structure, whereas spiral and So galaxies possess a flat disc component in addition to a spheroidal bulge component. Up to now, most calculations of the collapse and formation of galaxies have aimed at explaining the simpler elliptical galaxies, and both spherical models and flattened rotating models have been studied in some detail (Gott 1973, 1975; Larson 1969, 1974, 1975b, hereafter Papers I, II, III). In Paper III it was found that rotating models of initially gaseous, turbulent protogalaxies in which star formation takes place continuously during the collapse are able to reproduce satisfactorily not only the radial density profiles of typical elliptical galaxies, but also their isophotal shapes, ellipticity profiles, and (qualitatively) their metal abundance distributions. An additional feature of some of these models is that a small amount of residual gas condenses into a thin layer and forms an incipient disc component during the final stages of the collapse; this suggests that the same qualitative collapse picture may also be able to explain the formation of spiral galaxies, if a larger fraction of the initial gas can condense into a disc before forming stars. In the present paper we describe some results of further model

calculations of this type in which the assumptions have been varied in an attempt to understand how systems with substantial disc components might be formed.

Any attempt to model spiral galaxies is made difficult by the great diversity in their properties and by the difficulty of describing them in quantitative terms. The traditional Hubble classification scheme is based on the appearance of the spiral arms and the relative prominence of the nuclear bulge, but Sandage *et al.* (1970) argue that the most fundamental parameter distinguishing between different Hubble types is the relative gas content, presumed to have been determined at the time of formation. Faber & Gallagher (1976) suggest that the gas content is related to the relative prominence of the disc and bulge components, and that the disc: bulge ratio is a more fundamental parameter. Van den Bergh (1976) has proposed a two-dimensional classification scheme in which the disc:bulge ratio and the gas content are independent parameters; the disc:bulge ratio is determined by the formation process, but the gas content may depend at least in part on environmental factors. We have assumed that the most important parameter requiring explanation by a model of the formation process is the disc:bulge ratio, and a primary aim of the present work has been to determine how this ratio depends on the various model assumptions.

Recent studies of the stability of disc systems suggest that spiral galaxies possess unseen halos of significant mass which are required to stabilize their discs (Bardeen 1975); in addition, some estimates of the masses of spiral galaxies (e.g. Turner 1976) yield high mass:light ratios which suggest massive unseen halos. The nature of such halos, if they exist, is presently mysterious, and this poses a problem for the galaxy builder. If galaxies possess massive halos, then their light distributions are not representative of the mass distributions, and we have little firm information about the mass distributions which models should aim to reproduce. Moreover, the large differences in $M:L$ between different parts of galaxies imply large differences in the star formation processes involved. In order to partly avoid these problems, we have considered the present models as representing the structure and formation of only the visible parts of galaxies; any dark halo that may be present is imagined to form separately and not to have much effect on the formation of the rest of the galaxy. This assumption may be partially justified since the objects constituting any dark halo probably emit little or no mass or energy, and therefore may be regarded as forming an inert background which plays little role in the further evolution of the system, except for its gravitational field.

2. ASSUMPTIONS

2.1 *Initial and boundary conditions*

As in Paper III, we assume that a protogalaxy begins as a uniform, uniformly-rotating sphere of gas, with a frictionless spherical boundary of fixed radius. All of the models described here have a total mass of $10^{11}\,M_\odot$, although a few models were calculated with different masses. The boundary radii vary from 30 to 100 kpc, implying initial densities from $8{\cdot}8\times10^{-4}$ to $2{\cdot}4\times10^{-5}\,M_\odot\,\mathrm{pc}^{-3}$ and free-fall times from $2{\cdot}7\times10^{8}$ to $1{\cdot}66\times10^{9}$ yr. The initial velocity dispersion is given by the virial theorem without surface pressure term, and ranges from 55 to 30 km s^{-1}. The initial angular velocity has been treated as a free parameter, with values as given below.

It may be questioned whether the simple initial and boundary conditions described above are realistic for spiral galaxies. For example, it is possible that proto-spiral galaxies are already quite inhomogeneous when they begin to collapse, and that the denser parts condense first to form a spheroidal component whereas the less dense parts condense later to form a disc. Also, since spiral galaxies occur mostly in the field or in loose and probably unbound groups, the boundary of a proto-spiral galaxy may expand with time, lengthening the time scale for infall of the outermost regions. However, qualitatively similar effects occur even in models with simpler initial conditions: the collapse is always non-homologous and rapidly produces a centrally condensed spheroidal system, after which the less dense peripheral gas more gradually falls in and settles into a disc. Thus it should be possible to explore some of the important qualitative features of the collapse of protogalaxies even with the present simple models.

2.2 *Dissipation and turbulent viscosity*

The gaseous dissipation rate has been calculated using the same assumptions as in Papers I–III. In most cases the dissipation rate is assumed to be given by the colliding cloud model of Paper I, but in a few models the dissipation time scale is related to the local free-fall time as in equation (1) of Paper II. Both assumptions give similar results during the early stages of the collapse, but later the cloud collision time becomes long compared with the free-fall time and the two assumptions give divergent results for the dissipation rate. However, this difference does not affect any of the qualitative features of the models, and alters only the time scale for the formation of the disc component.

In Paper III it was found that in order to obtain realistic centrally condensed models of rotating elliptical galaxies it is necessary to include a turbulent viscosity effect in the calculations. The turbulent viscosity transports angular momentum outward, allowing the system to become highly condensed without becoming highly flattened. If the viscosity remains large during the later stages of the collapse, the residual gas spirals rapidly towards the centre without forming a thin disc; however, if the viscosity decreases, as in Model 4 of Paper III, the residual gas settles towards a plane and begins to form a stellar disc component. Since the latter situation seems more promising as a start toward modelling the formation of spiral galaxies, we have here used the same treatment of turbulent viscosity as was used in Model 4 of Paper III: the viscosity is given by equation (A15) of Paper III, with the parameter C_δ usually set equal to 0·4. In some models the viscosity has been set equal to zero during the later stages of the collapse, as discussed below (Section 3.2).

2.2 *Star formation*

The most important factor determining the structure of the present models is the assumed star formation rate and the way in which it depends on the gas density and other variables. In Paper III it was found that models with a simple power-law dependence of star formation rate (SFR) on gas density of the form

$$\frac{d\rho_s}{dt} = A\rho_g{}^n \qquad (1)$$

with $n \sim 1{\cdot}85$ can provide a generally satisfactory representation of the observed properties of elliptical galaxies. If the SFR follows a power law with $n > 1{\cdot}5$, this

might also make it possible to explain different galaxy types as resulting from different initial conditions: since the ratio of star formation time to free-fall time increases when the initial density is decreased, less dense protogalaxies might be expected to form less massive spheroidal components during early stages of the collapse, and to form more massive disc components from the remaining gas (Gott & Thuan 1976). We have used a power law as the basic representation of the SFR in the present models, with the various modifications described below.

As will be seen, if a massive disc component is to be formed, the SFR must decrease more strongly during the later stages of the collapse than is predicted by the power law (1). A possible effect which might act to reduce the SFR during the later disc-forming stages of the collapse would be a dependence of the SFR on the velocity dispersion of the gas; for example, if star formation results from compression of the gas in cloud collisions, the SFR will depend on the collision frequency and perhaps also on the velocity with which the clouds collide. The decrease in velocity dispersion occurring as the gas dissipates its random motions and settles towards a disc would then cause a decrease in the SFR. To simulate this effect, we have in one case (Model 3) adopted a star formation law of the form

$$\frac{d\rho_s}{dt} = A\rho_g{}^n\alpha_g \tag{2}$$

where α_g is the mean square random velocity of the gas.

A second effect that might more strongly suppress star formation during the later stages of the collapse is the tidal force exerted on the residual gas by the already formed spheroidal component (Larson 1975a). If the gas density at radius r is less than the average density of matter within a sphere of radius r, the tidal disruptive forces exceed local self-gravitational forces in the gas and prevent collapse of the gas into stars, unless sufficiently large density enhancements are already present. Since the gas is almost certainly quite inhomogeneous, it is probably not realistic to assume that star formation is cut off completely when the average gas density falls below this critical density; instead, there is probably a more gradual reduction in the SFR with decreasing density, the reduction factor depending on the small scale density distribution of the gas. After some experimentation with different possibilities, we have adopted the following simple functional form to represent tidally inhibited star formation:

$$\frac{d\rho_s}{dt} = A\rho_g{}^n\left[1+\frac{B\bar{\rho}(r)}{\rho_g}\right]^{-1} \tag{3}$$

where $\bar{\rho}(r)$ is the average density of matter inside a sphere of radius r and B is a constant of order unity. According to equation (3), the SFR is reduced by approximately a factor $\rho_g/B\bar{\rho}(r)$ when ρ_g is less than $B\bar{\rho}(r)$; this factor might be justified if, for example, a fraction $\rho_g/B\bar{\rho}(r)$ of the gas has density greater than $\bar{\rho}(r)$ and can still collapse into stars despite the tidal effect. There is, of course, little theoretical justification for such an assumption, so the results obtained (e.g. Models 4 and 5) must be considered as only illustrative of the possibilities.

If the protogalactic gas has a two-phase structure with dense clouds and less dense intercloud gas, it is possible that the dense clouds collapse rapidly to form stars during early stage of the collapse, whereas the intercloud gas does not form stars until it has settled to a thin disc. The SFR in a protogalaxy then depends on the fraction of the gas which is in dense clouds, and this fraction may depend on

the initial conditions; also, it may vary with time as the dense clouds become transformed into stars, leaving only less dense intercloud gas during later stages of the collapse. In the absence of a proper multi-component model, we have attempted to simulate such possible effects by artificially varying the coefficients A and B in equation (3), as in Models 6 and 7.

Finally, we have considered the possibility that the rapid early star formation required to form a spheroidal component is caused by strong compression of the gas during an early phase of strong turbulence or rapid cloud collisions, and that the SFR decreases markedly during later stages of the collapse as collisions become less important or as less turbulent outlying gas settles into the system. More specifically, we have assumed that the SFR depends on the ratio of the cloud collision time to the free-fall time and is given by equation (1) as long as the collision time is shorter than the free-fall time, but is later reduced by a factor depending on the ratio of these time scales when the collision time becomes longer than the free-fall time. For the cloud model of Paper I, the cloud collision time is

$$t_{\rm coll} \simeq 3{\cdot}1 \times 10^2 (\rho_g \alpha_g)^{-1/2}\ {\rm Myr}, \tag{4}$$

and since the free-fall time is

$$t_{\rm ff} = 8{\cdot}09\ \bar{\rho}(r)^{-1/2}\ {\rm Myr}, \tag{5}$$

we have

$$\frac{t_{\rm coll}}{t_{\rm ff}} \simeq 38 \left[\frac{\bar{\rho}(r)}{\rho_g \alpha_g}\right]^{1/2}. \tag{6}$$

By analogy with the formula used in equation (3) to represent tidally inhibited star formation, we have assumed the following simple dependence of the SFR on the ratio $t_{\rm coll}/t_{\rm ff}$ as given by equation (6):

$$\frac{d\rho_s}{dt} = A\rho_g{}^n \left[1 + \frac{C\bar{\rho}(r)}{\rho_g \alpha_g}\right]^{-1}. \tag{7}$$

In the limit of frequent collisions ($t_{\rm coll}/t_{\rm ff} \ll 1$), equation (7) approaches the power law $A\rho_g{}^n$ which results in satisfactory models for spheroidal systems, and in the opposite limit ($t_{\rm coll}/t_{\rm ff} \gg 1$) the SFR is reduced by a factor $\rho_g \alpha_g / C\bar{\rho}$ which is similar to the previous tidal reduction factor except that it depends also on the velocity dispersion of the gas, in the same way as does the SFR of equation (2).

2.4 *Gas recycling and metal enrichment*

Gas loss from stars has been calculated using the same simple analytic approximation as was used in Paper III (equation (5)). The metal abundances of both gas and stars are also calculated as in Paper III, using the instantaneous recycling approximation. In the present models, the steep gradient of metal abundance perpendicular to the plane of the disc and the coarse grid resolution result in a relatively low numerical accuracy for the metal abundance of the gas at late times, but the spatial distribution of average stellar metal abundance is only slightly affected by this inaccuracy.

3. RESULTS

3.1 *Models with a power-law star formation rate*

A number of models were first calculated with the simple power-law SFR of equation (1) in order to determine whether it is possible by any variation of the

parameters to obtain systems with substantial disc components. Here we present results for a few of these models which illustrate the range of characteristics found. In all cases the results are given for a time at least four free-fall times after the start of the collapse, when at least 85 per cent of the gas has been transformed into stars and the stellar system has closely approached its final equilibrium state.

As remarked in Section 2.3, a lower initial density might be expected to lead to a higher flattening or a larger disc:bulge ratio for the resulting system. Model 1 has been calculated with assumptions that are similar to those of Model 4 of Paper III except that the boundary radius has been increased from 30 to 100 kpc, decreasing the initial density by a factor of 37; also, the initial balance of forces has been maintained by decreasing the initial velocity dispersion from 55 to 30 km s^{-1} and the initial angular velocity from 2×10^{-3} to $3{\cdot}3 \times 10^{-4}$ Myr^{-1}. The parameters assumed in the star formation law (equation (1)) are $A = 0{\cdot}55$, $n = 1{\cdot}85$; the gaseous dissipation rate is assumed to be given by equation (6) of Paper I; and the viscosity parameter C_δ has been set equal to 0·4. The density distribution obtained for Model 1 is illustrated in Fig. 1, and is only slightly different from that of Model 4 of Paper III. (As in Paper III, we show only the spatial density distribution, since the projected surface density distribution is very similar in appearance.) This result is typical, and shows that the structure of the models is less sensitive to the initial density than might be expected from simple arguments (e.g. Gott & Thuan 1976). This is because (1) a decrease in initial density increases not only the star formation time but also the dissipation time relative to the free-fall time, and therefore the ratio of star formation time to dissipation time, which is what primarily determines the structure of the system, is not greatly altered; also, (2) the assumed turbulent viscosity continually transfers angular momentum outward and prevents the inner part of the system from becoming very flattened, without much affecting the outer parts.

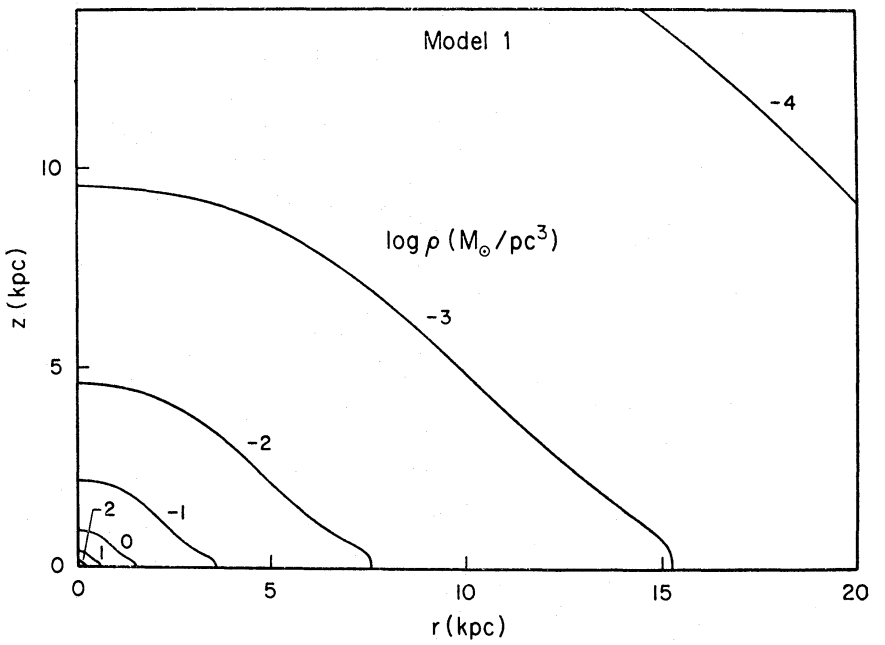

FIG. 1. *Contours of constant stellar density ρ in one quadrant of a meridional plane of Model 1; curves are labelled with the logarithm of the stellar density in units of* $M_\odot$ pc^{-3}. *This model is similar to Model 4 of Paper III except that the boundary radius is 100 kpc instead of 30 kpc.*

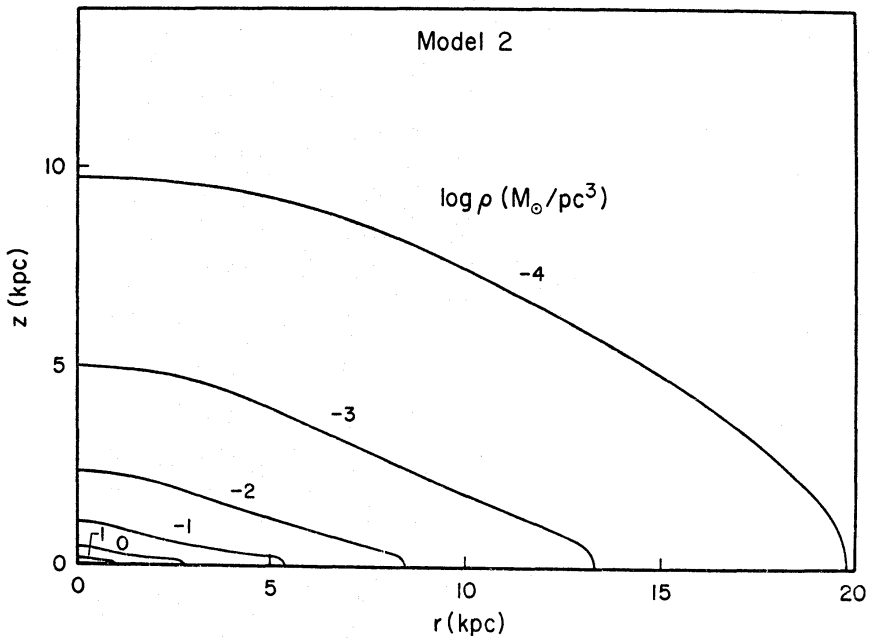

FIG. 2. *Density contours for Model 2, in which the star formation rate coefficient has been reduced by a factor of* 10.

Several further calculations in which the initial angular velocity and the viscosity parameter C_δ were varied, as well as the initial density, yielded results very similar to those for Model 1, except for the overall flattening which increases with increasing initial angular velocity or with decreasing viscosity. In all cases, however, the density contours differ from those of Fig. 1 mainly by scale factors in the vertical and horizontal directions, and there is no significant variation in the relative prominence of a disc-like component. Thus there appears to be no prospect of accounting for the Hubble sequence just by varying the initial conditions or the assumed turbulent viscosity in models of this simple type.

It is conceivable that the coefficient A in the star formation law is for some reason much smaller in proto-spiral galaxies than in proto-ellipticals, for example if the gas is initially much less clumpy. The effect of reducing A by a factor of 10 to 0·055 is illustrated by Model 2 (Fig. 2), which is identical in other respects to Model 4 of Paper III except that the initial angular velocity is 3×10^{-3} instead of 2×10^{-3} Myr^{-1}. In this case the gas condenses into a much flatter distribution before forming stars, and the inner part of the resulting system might be described as disc-like in structure. However, in this case there is little indication of any distinct spheroidal component, and the system still does not resemble a spiral galaxy with distinct bulge and disc components.

The effect of assuming a star formation law which depends on the velocity dispersion as well as the density of the gas is illustrated by Model 3 (Fig. 3), in which the SFR is given by equation (2) with $A = 1{\cdot}5 \times 10^{-5}$ and $n = 1{\cdot}6$; here the value of n has been adjusted to give a radial density distribution which is still in reasonable agreement with that of a typical elliptical galaxy. The other (less important) assumptions of Model 3 are a radius of 60 kpc, an initial velocity dispersion of 55 km s^{-1}, an initial angular velocity of 10^{-3} Myr^{-1}, and a dissipation time scale equal to twice the local free-fall time. The resulting system shows more of a two-component structure than previous models, with a flattened disc-like

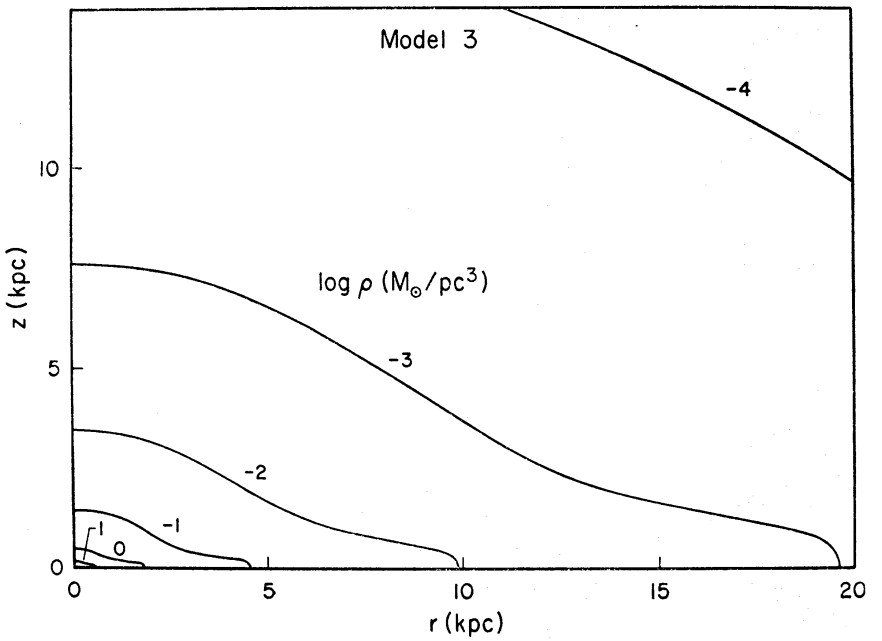

FIG. 3. *Density contours for Model 3, in which the star formation rate depends on the velocity dispersion as well as the density of the gas as in equation* (2).

component superimposed on a more spheroidal bulge-like component. The flattened component stands out more distinctly here because of the decrease in SFR caused by the decrease in velocity dispersion of the gas as it settles toward a plane; this allows more of the gas to condense into a thin disc before forming stars. Comparison of Fig. 3 with an isophotal map of the So galaxy NGC 3115 (Miller & Prendergast 1968) suggests that such a system would appear visually to possess a quite definite disc component and would probably be classified as an So galaxy. However, even in this case the disc component contains only a minor fraction of the mass, and more drastic modifications of the assumptions are required to obtain more massive disc components.

3.2 *Models with tidally inhibited star formation*

We now present results for some models in which the possible effect of tidal forces in inhibiting star formation is taken into account as in equation (3). Figs 4 and 5 show the density distributions obtained for Models 4 and 5, which have the same star formation law ($A = 0{\cdot}55$, $B = 0{\cdot}25$, $n = 1{\cdot}85$), but different boundary radii of 30 and 100 kpc respectively. Model 4 has an initial velocity dispersion of 55 km s^{-1} and an initial angular velocity of 2×10^{-3} Myr^{-1}, while Model 5 has a velocity dispersion of 30 km s^{-1} and an angular velocity of $3{\cdot}3 \times 10^{-4}$ Myr^{-1}; with these parameters the initial ratios of gravity, pressure and centrifugal force are the same in both models. Models 4 and 5 develop substantial flat disc components because the large reduction in SFR during the later stages of the collapse allows a large part of the residual gas to condense to a thin disc before forming stars. The disc component contains approximately 25 per cent of the total mass in Model 4 and 45 per cent of the mass in Model 5; the disc mass is larger in Model 5 because the lower initial density means that less gas is transformed into a spheroidal component during early stages of the collapse, and more remains to condense later when star formation is tidally inhibited, and thus form a disc component.

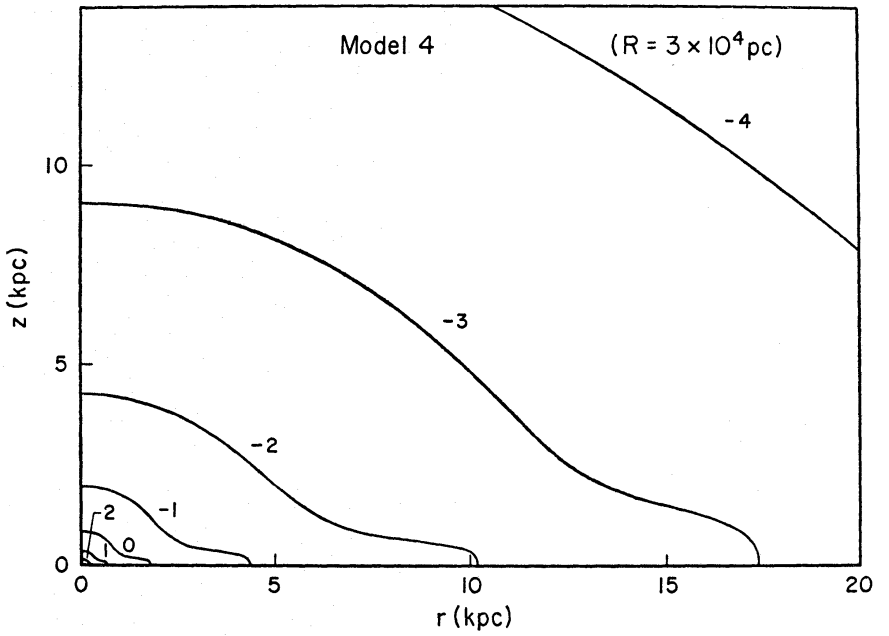

FIG. 4. *Density contours for Model 4, in which star formation is tidally inhibited in accordance with equation* (3) *with* $B = 0{\cdot}25$.

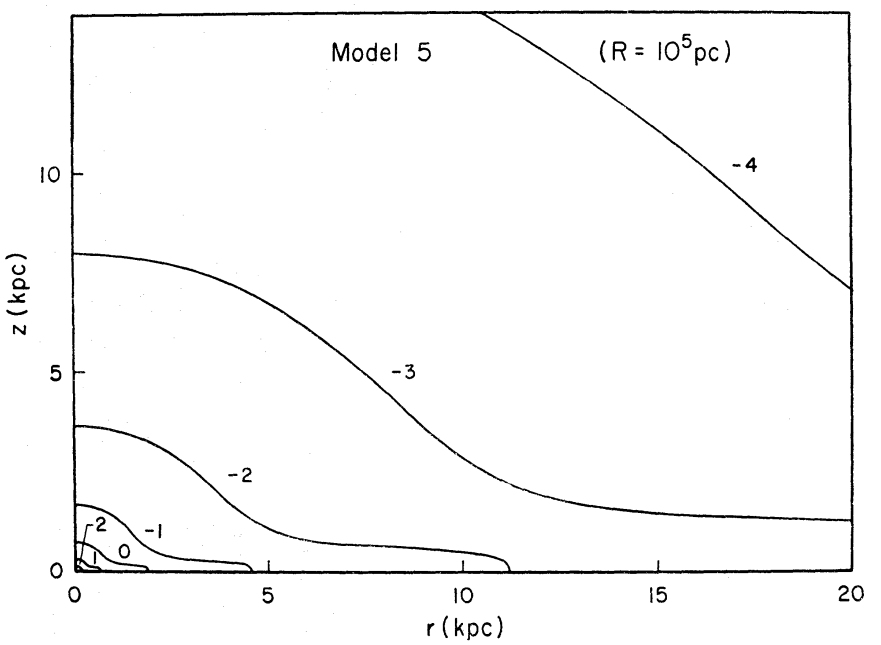

FIG. 5. *Density contours for Model 5, in which the star formation law is the same as in Model 4 but the boundary radius is* 100 *kpc instead of* 30 *kpc.*

In the above models and all of those which follow, the turbulent viscosity has been set equal to zero after the formation at an early stage of a dense core with density greater than $10\,M_{\odot}\,\mathrm{pc}^{-3}$. This has been done because, as discussed in Paper III, the viscosity due to random cloud motions is expected to be important only during the initial near free-fall stage of the collapse. Also, if the gas which eventually forms a disc is less turbulent than that which forms the spheroidal component, viscosity effects should be less important for this gas. In any case, the

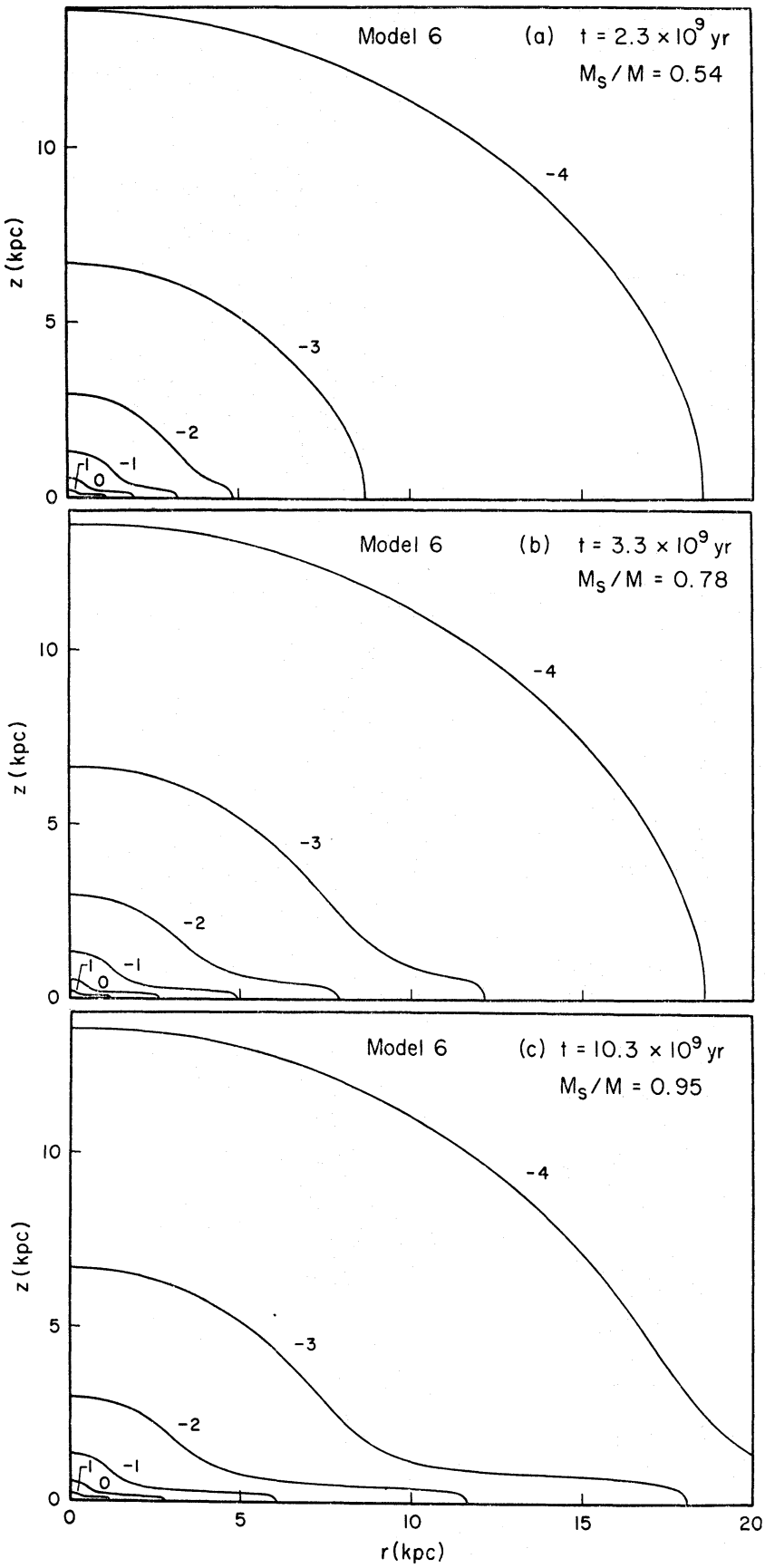

FIG. 6. *Stellar density contours for three stages in the evolution of Model* 6, *in which star formation is tidally inhibited as in equation* (3) *but the constant B is increased from* 1 *to* 10 *when the central density rises above* 10 $M_\odot\,pc^{-3}$. *The time measured from the beginning of cosmological expansion and the fractional mass in stars at each stage are* (*a*) $2{\cdot}3 \times 10^9$ *yr*, 0·54; (*b*) $3{\cdot}3 \times 10^9$ *yr*, 0·78; (*c*) $10{\cdot}3 \times 10^9$ *yr*, 0·95.

exact treatment of viscosity during the later stages of the collapse appears not to be very important, as long as the viscosity is not large.

From the results for Models 4 and 5 and other similar calculations, it appears that models of this type can produce disc components containing up to roughly one-half of the total mass, and that it is possible to some extent to explain the observed variation in disc:bulge ratio as resulting from differences in the initial conditions. The most important parameter governing the disc:bulge ratio in these models is the initial density; other model parameters turn out to have relatively little effect on this ratio. A larger initial angular velocity, for example, gives a model with a more flattened bulge component and a more extended disc, but almost the same disc:bulge ratio. In order to obtain more massive disc components and a greater range of disc:bulge ratios, as is required to explain the Hubble sequence, further modifications of the assumed star formation law are probably required.

If the protogalactic gas has a two-phase structure, the SFR may decrease rapidly after most of the dense cloud gas has condensed into stars, leaving only the less dense intercloud medium. The effect of a large drop in the SFR following the formation of a spheroidal component is illustrated by Model 6 (Fig. 6), in which the constant B in equation (3) is initially equal to 1 but is increased to 10 when the central density rises above 10 $M_\odot$ pc^{-3}, thus reducing the SFR by almost a factor of 10. The radius of Model 6 is 60 kpc, the initial velocity dispersion 39 km s^{-1}, and the initial angular velocity $7{\cdot}1 \times 10^{-4}$ Myr^{-1}, all intermediate between the values for Models 4 and 5. About 70 per cent of the total mass of Model 6 eventually ends up in the disc component.

Fig. 6(a)–(c) show a typical evolutionary sequence illustrating how the disc grows with time. After the formation of a spheroidal component, which is essentially completed by the time $t \sim 2{\cdot}0 \times 10^9$ yr, residual gas with low angular momentum condenses first to form the innermost part of the disc, followed by gas with higher and higher angular momentum which settles into the equatorial plane at larger and larger radii, causing the disc to grow outward with time. In the outer regions the time scale for the growth of the stellar disc becomes very long ($\gtrsim 10^{10}$ yr) and is dominated by the long time scale required for star formation at the low gas densities present in the outer parts of the disc. This long time scale for the formation of the outer parts of the disc is characteristic of all of the present models.

A larger disc:bulge ratio can be obtained if the SFR during the initial bulge-forming stage of the collapse is further reduced, as might be the case if only a small fraction of the gas is initially in dense clouds. The most extreme disc model calculated is Model 7 (Fig. 7), which differs from Model 6 in that the coefficient A in the star formation law is reduced by a factor of 10 to 0·055 and the initial angular velocity is increased to $1{\cdot}0 \times 10^{-3}$ Myr^{-1}. In this case a much smaller fraction of the gas is turned into stars in a spheroidal component, and about 90 per cent of the total mass eventually condenses into the disc.

Although the disc:bulge ratios of Models 6 and 7 are comparable with those of the visible matter in some spiral galaxies, we mention two problems that were encountered in models with massive discs. First, when the disc mass exceeds something like one-half of the total mass, the disc shows a tendency to clump into concentric rings. This may be a numerical effect, but it is also possible that axisymmetric instabilities of physical origin are present. Since it is difficult to decide to what extent this is a numerical phenomenon, we have assumed in presenting the results that the rings are entirely numerical and have smoothed them out. The

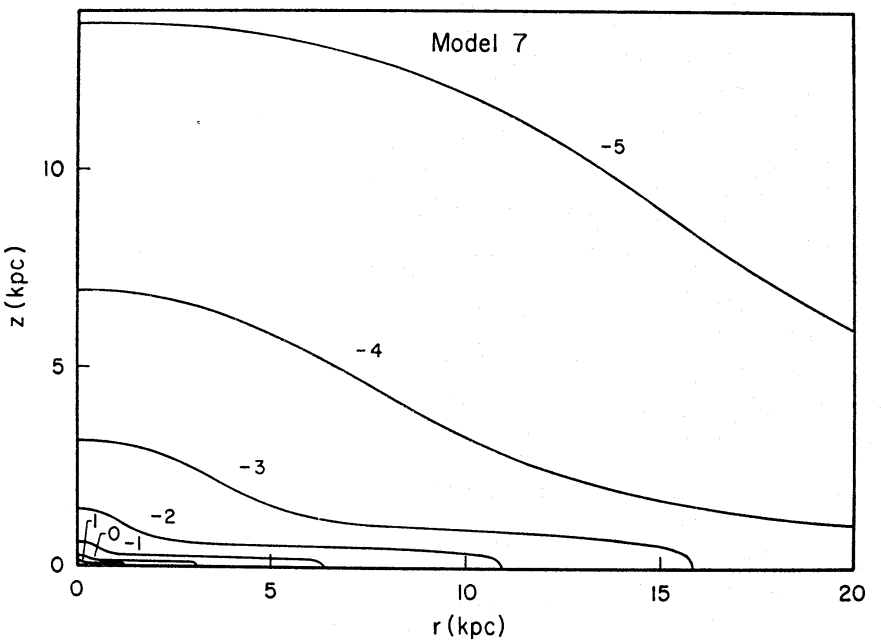

FIG. 7. *Density contours for Model 7, which is similar to Model 6 except that the star formation rate coefficient A is reduced by a factor of* 10.

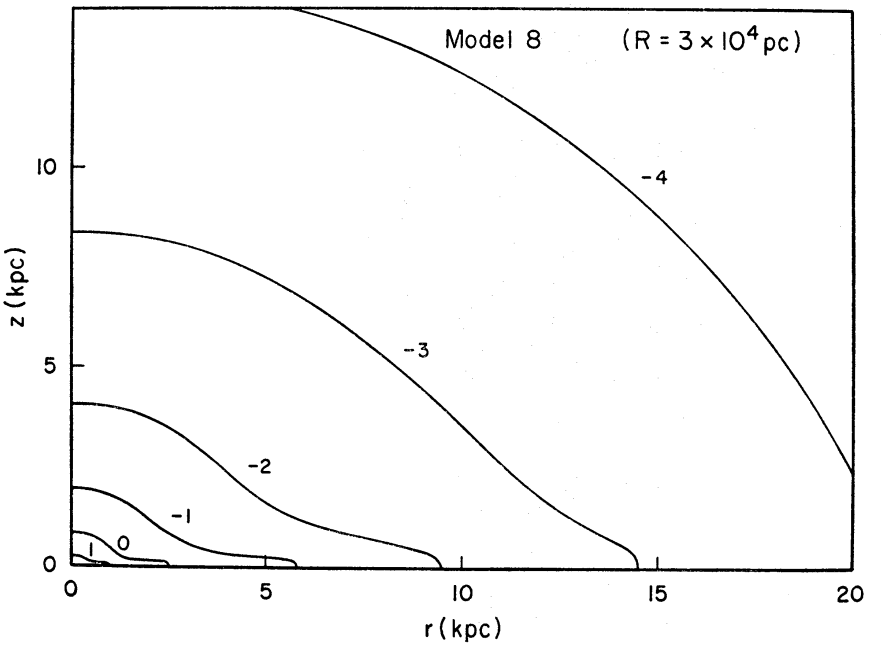

FIG. 8. *Density contours for Model 8, in which the star formation rate depends on the ratio of cloud collision and free-fall times as in equation* (7) *with* $A = 1{\cdot}0$, $n = 1{\cdot}85$, $C = 5000$.

second problem is a tendency to develop too large a concentration of mass at the centre; in the most extreme case, Model 7, this is reflected in a rotational velocity which reaches nearly 1000 km s^{-1} at a radius of 60 pc, clearly an unrealistically high value. If the models are otherwise realistic, this suggests that processes not represented in the models, such as some form of violent nuclear activity, must operate in the nuclei of galaxies to prevent the accumulation of too much mass there.

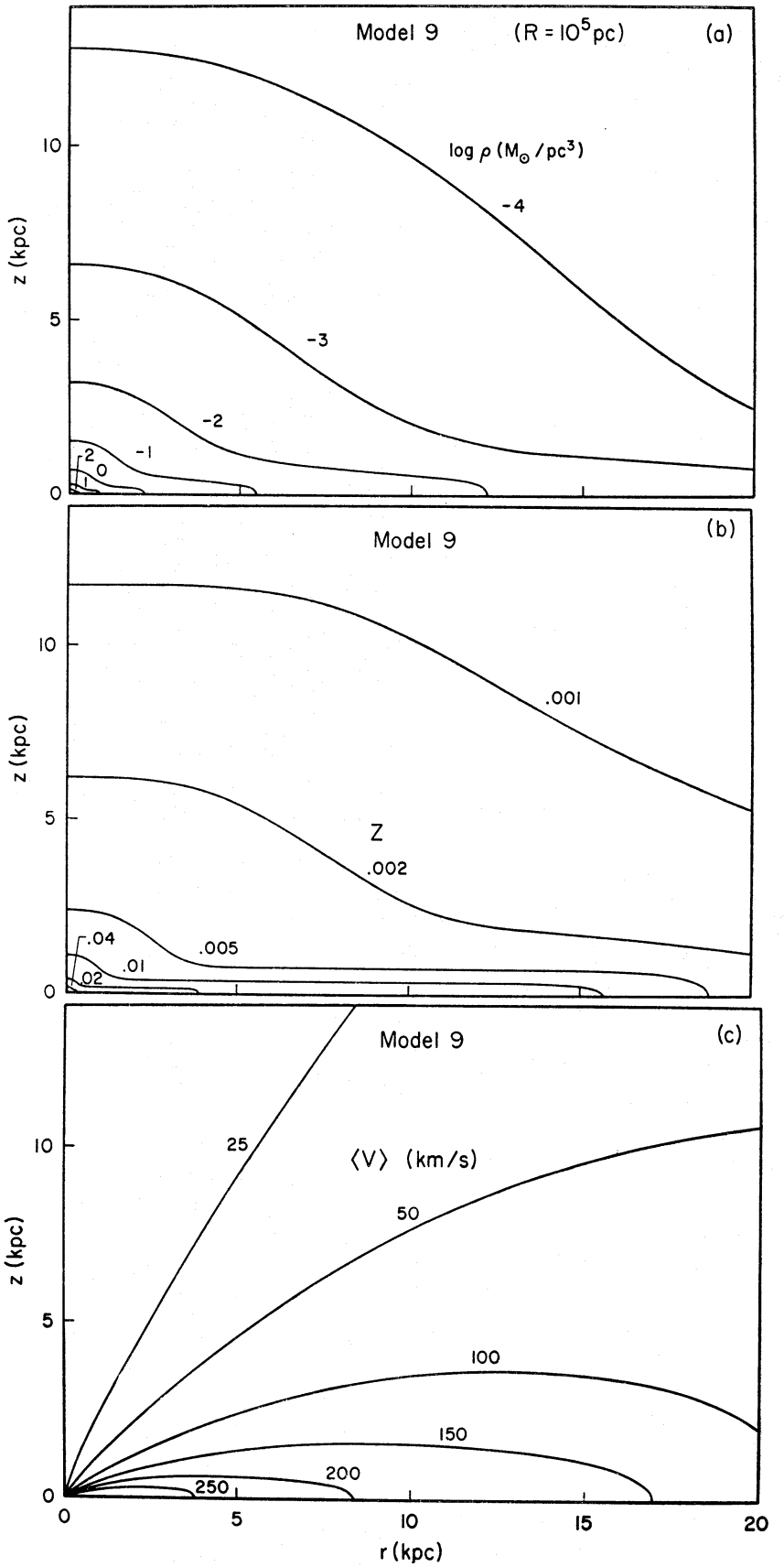

FIG 9(a,b,c)

FIG. 9. (*a*) *Density contours for Model* 9, *in which the star formation law is the same as in Model* 8 *but the boundary radius is* 100 *kpc instead of* 30 *kpc.* (*b*) *Contours of average stellar metal abundance Z for Model* 9. (*c*) *Contours of mean stellar rotational velocity* $\langle V \rangle$ *in km* s^{-1} *for Model* 9. (*d*) *Contours of stellar velocity dispersion disp* (*V*) *in km* s^{-1} (*assumed isotropic*) *for Model* 9.

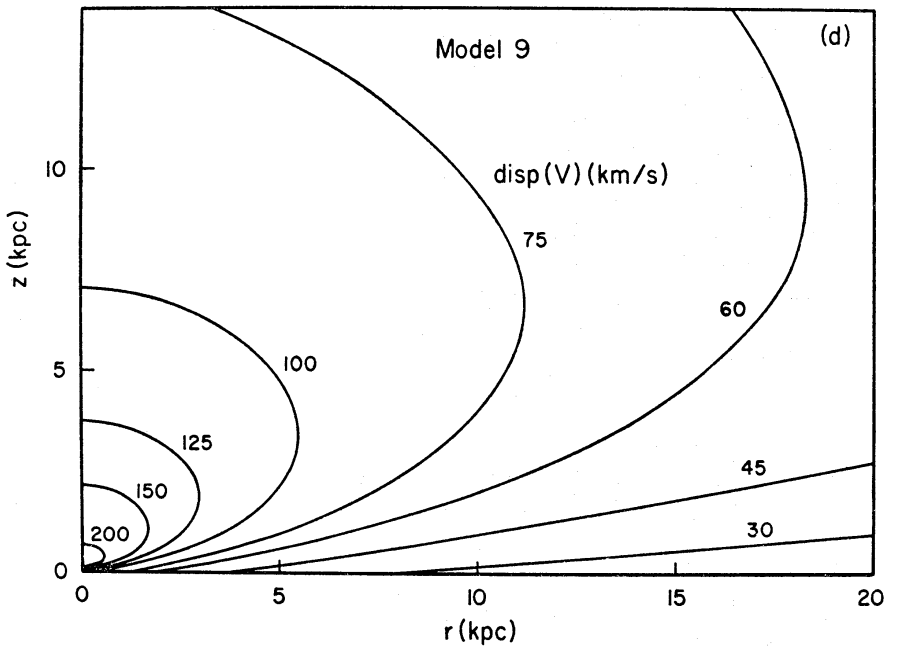

FIG 9(d)

3.3 *Models with SFR depending on a cloud collision time*

Finally, we present some results for models in which the SFR depends on the ratio of the cloud collision time to the free-fall time, and is given by equation (7). Figs 8 and 9(a) show the density distributions obtained for Models 8 and 9, which are based on the same assumptions as Models 5 and 6 except that the SFR is given by equation (7) with $A = 1{\cdot}0$, $n = 1{\cdot}85$ and $C = 5000$ km^2 s^{-2}. With this choice of parameters, the second term in square brackets in equation (7) is already important initially, and the SFR depends on the velocity dispersion as well as on the density of the gas. Both of these effects make the ratio of star formation time to collapse time greater in Model 9 than in Model 8, so that a larger fraction of the mass eventually condenses into a disc in Model 9. The fraction of the total mass contained in the disc component is about 40 per cent in Model 8 and 60 per cent in Model 9.

Since Model 9 is perhaps the most satisfactory model for a spiral galaxy, some further properties of this model are shown in Fig. 9(b)–(d). The metallicity distribution is given in Fig. 9(b), and reflects clearly the presence of two stellar subsystems: a spheroidal component with a radial Z gradient similar to that previously obtained in models for elliptical galaxies (Papers II and III), and a relatively metal-rich disc component with a weaker radial Z gradient but a strong gradient perpendicular to the disc. The distribution of average rotational velocity $\langle V \rangle$ is illustrated in Fig. 9(c) and also shows a distinction between a rapidly differentially-rotating disc component and a much more slowly and more uniformly-rotating halo. The velocity dispersion illustrated in Fig. 9(d) shows a strong gradient perpendicular to the disc, ranging from relatively small values in the plane of the disc to much larger values in the halo region. Similar results are obtained for all models with substantial disc components.

The following properties of Model 9 at a radius of 9 kpc in the equatorial plane may be of interest for comparison with the properties of the solar neighbourhood

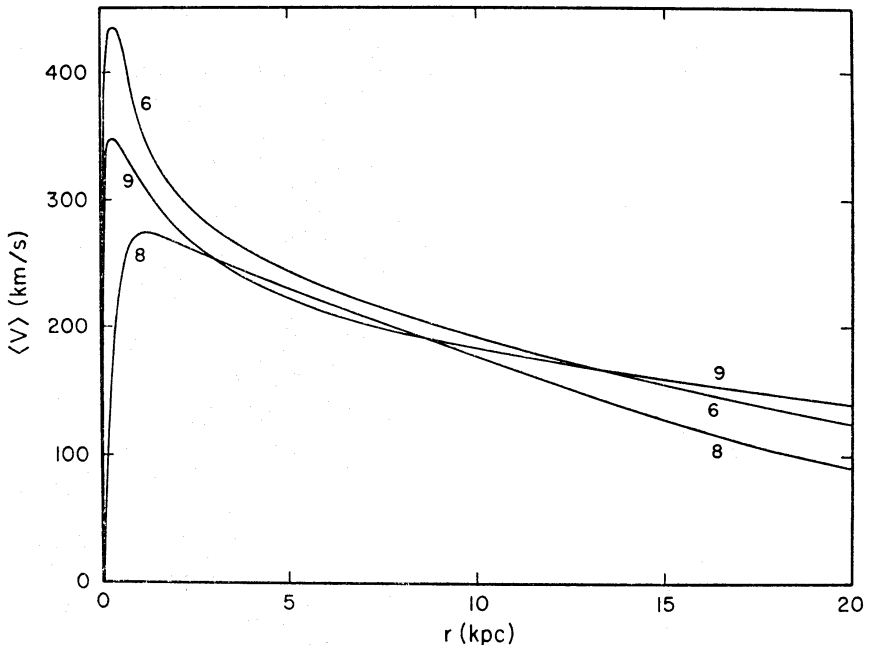

FIG. 10. *Rotation curves giving the mean stellar rotational velocity* $\langle V \rangle$ *in the equatorial plane for Models* 6, 8 *and* 9.

in our Galaxy, although no attempt has been made to make the model reproduce any of the observed properties of our Galaxy. The average rotational velocity in the model is about 195 km s^{-1} and the velocity dispersion is about 30 km s^{-1}, both in fair agreement with the values for the old disc population in the solar neighbourhood. The space density in the model is about 0·03 $M_\odot$ pc^{-3}, about five times smaller than the solar neighbourhood value of $\sim$0·15 $M_\odot$ pc^{-3}, but the model surface density of 45 $M_\odot$ pc^{-2} is closer to the solar neighbourhood value of $\sim$80 $M_\odot$ pc^{-2}, showing that most of the difference between the model and the galaxy is that the scale height of the disc is larger in the model. In the vertical direction, the model rotational velocity decreases by a factor of 2 and the velocity dispersion increases by a factor of about 2·7 between the equatorial plane and a height $z \simeq 3$ kpc; for comparison, we estimate from data given by Oort (1965) that in our Galaxy similar changes occur between $z = 0$ and $z \simeq 1{\cdot}5$ kpc. Again, the scale height of the model appears to be greater than that of the galaxy by about a factor of 2. Even this difference may be partly of numerical origin, since the model is not able to represent accurately the vertical structure of a thin disc because of the coarse grid resolution and the use of only a single stellar component with an isotropic velocity distribution.

Rotation curves for Models 6, 8 and 9 are shown in Fig. 10. As previously noted, the rotation curves tend to be too sharply peaked near the centre, reflecting an unrealistically large concentration of mass at the centre. The rotation curves also continue to decline more rapidly at large radii than those of some spiral galaxies, indicating that the models contain less mass at large radii. The model rotation curves appear to resemble most closely those of spiral galaxies of the earliest Hubble types, such as the Sab galaxy M81 (Roberts & Rots 1973); this is consistent with the fact that the models contain prominent bulge components and thus would probably be classified as early-type galaxies.

The radial profiles of projected surface density for Models 8 and 9 are shown in

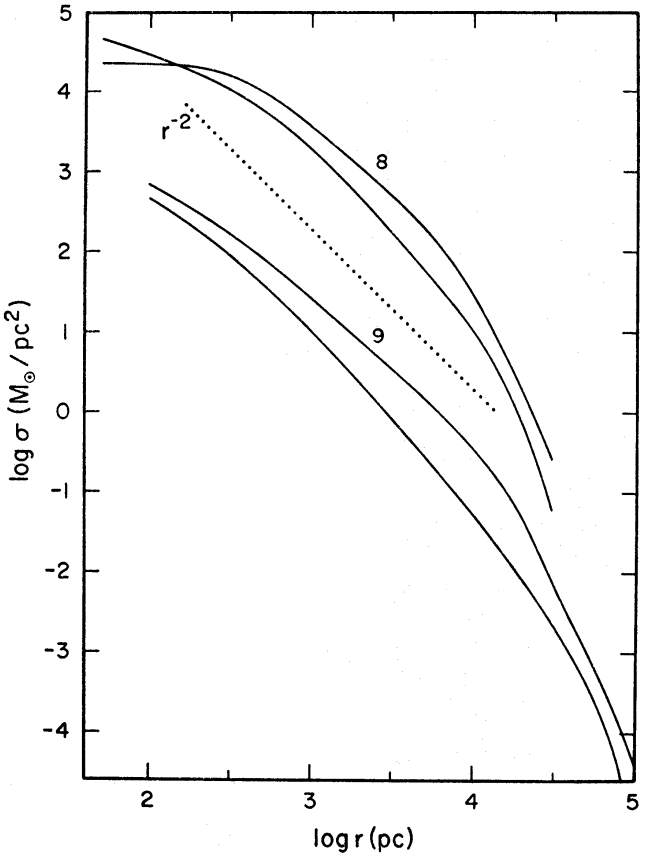

FIG. 11. *Profiles of projected stellar surface density σ(r) for Models 8 and 9. In each case the upper curve is the profile for a pole-on projection and the lower curve is the profile measured along the minor axis for an edge-on projection. The curves for Model 9 are displaced downward by 2 units in log σ.*

Fig. 11. In each case the lower curve is the minor axis profile for an edge-on projection and represents primarily the bulge component, whereas the upper curve is the profile for a pole-on projection and thus represents both bulge and disc components. The radial structure of the bulge component is similar to that of the elliptical galaxy models of Papers II and III, except that the density distribution is slightly steeper, approximating r^{-2} instead of $r^{-1.65}$. The slight hump in the pole-on profile is caused by the disc component, whose mass is mostly concentrated in a smaller range of radii than the spheroidal component. Although this is qualitatively what would be expected for a disc with an exponential surface density profile, as has been claimed to be characteristic of many spiral galaxies (Freeman 1970; Schweitzer 1976), the disc components in the present models are not closely exponential in structure but tend to be intermediate between an exponential and a power law.*

3.4 *Evolution of the models*

The time dependence of the total star formation rate dM_s/dt for Models 6, 8 and 9 is shown in Fig. 12, where the zero-point of time is taken as in Paper II to

* According to A. Oemler (private communication), a critical examination of surface photometry for many spiral galaxies shows a considerable variation in the forms of the surface brightness profiles, and only a minority are well approximated by an exponential law.

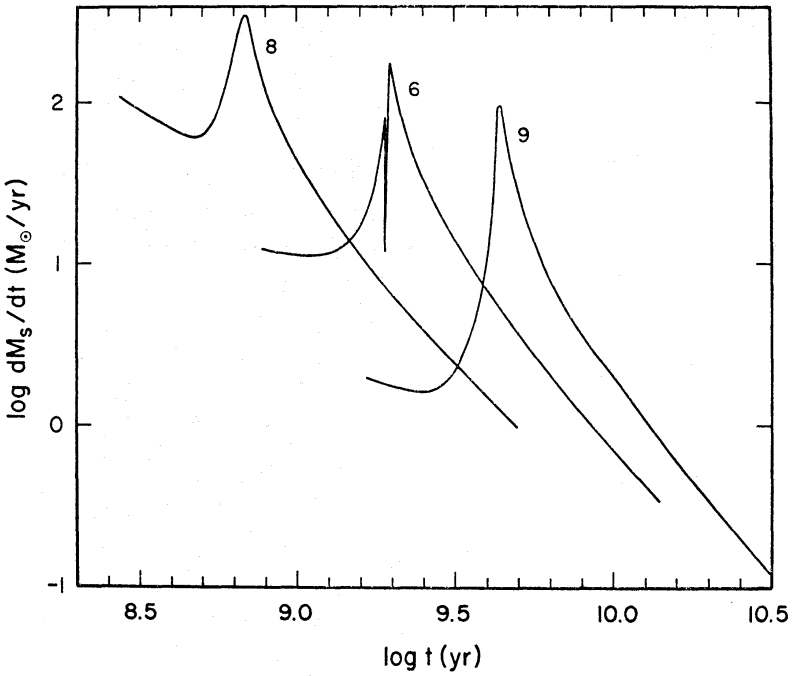

FIG. 12. *The total star formation rate dM_s/dt vs time for Models* 6, 8 *and* 9. *The zero point of time is taken as the beginning of cosmological expansion, and the model calculation is assumed to begin one free-fall time later.*

be the beginning of cosmological expansion, and the collapse is assumed to begin one free-fall time later. These results are qualitatively similar to those for the spherical models of Paper II except for a sharper peak in the SFR associated with the formation of a denser core. The cases illustrated in Fig. 12 differ mainly in time scale, which depends on the assumed radius of the model. The free-fall times for Models 8, 6 and 9 are $2{\cdot}7 \times 10^8$, $7{\cdot}7 \times 10^8$ and $1{\cdot}66 \times 10^9$ yr, respectively, and the times of peak SFR are $6{\cdot}9 \times 10^8$, $2{\cdot}0 \times 10^9$ and $4{\cdot}4 \times 10^9$ yr, about 1·5–1·6 free-fall times later in each case. We note that in Models 6 and 9 the SFR at $t = 12 \times 10^9$ yr is of the order of 1 $M_\odot$ yr^{-1}, comparable with that of a typical Sa galaxy, as inferred from its *UBV* colours (Larson & Tinsley 1974).

Fig. 13 shows the time dependence of the SFR at several radii in the equatorial plane of Model 9. The time scale for this model is relatively long, but would be reduced (i.e. the curves shifted to the left) if either the assumed radius of 10^5 pc is too large or the mass of 10^{11} $M_\odot$ is too small. In the outer part of the disc the SFR has two peaks, corresponding to two main phases of star formation which are not well separated at small radii but become progressively more distinct at larger radii: (1) an initial burst of star formation which produces the spheroidal component, and (2) a later, more extended phase of star formation which occurs as gas settles into the equatorial plane and builds up a disc component by an outward-progressing wave of star formation, which is reflected in the successively later times of maximum SFR at larger radii. Similar results are obtained for all models, although quantitative details such as the time delay between the two peaks of the SFR depend on how rapidly the residual gas settles to a disc, and hence on the assumed gaseous dissipation rate.

Although the models are probably too crude to bear a detailed comparison with the evolution of the solar neighbourhood in our Galaxy, we note that the SFR

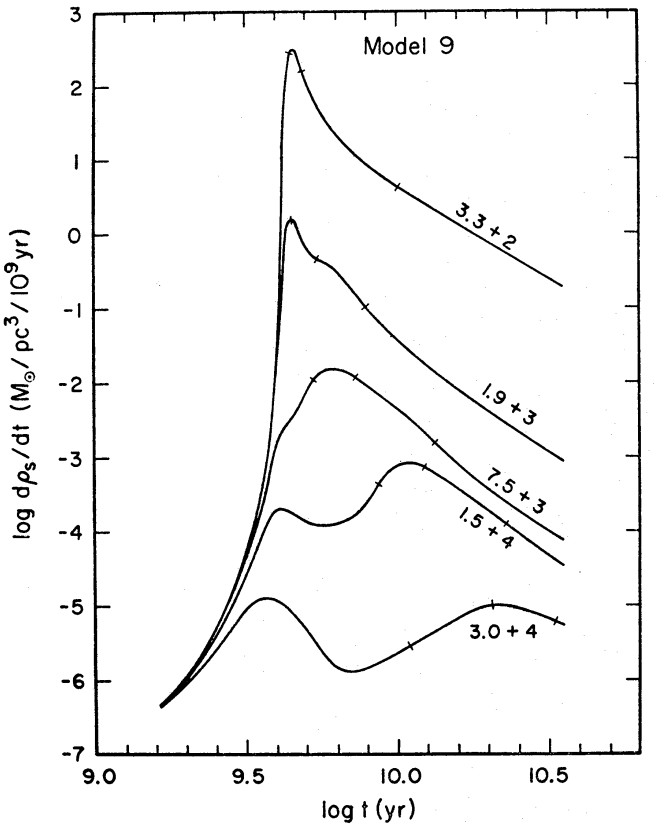

FIG. 13. *The star formation rate per unit volume vs time at five radii in the equatorial plane of Model 9; the radius is marked in pc on each curve. The three ticks marked on each curve indicate the times at which* 10, 50 *and* 90 *per cent of the final density of stars is attained.*

in the outer parts of Model 9 does not vary greatly during the period when most of the stars are being formed; this is consistent with indications that the SFR in the solar neighbourhood has not varied strongly during its past history (Tinsley 1976). (Although the predicted SFR refers to a unit volume and the ' observed ' SFR to a unit area of the galactic plane, the comparison is roughly valid since most of the star formation in the model occurs within one grid zone of the equatorial plane, and an average density in this cell is equivalent to a surface density in the plane.) The model results also suggest the possibility of a delay of several 10^9 yr between the formation of the galactic halo and the occurrence of significant star formation in the outer parts of the disc, a possibility which is at least consistent with present age determinations for disc and halo objects in our Galaxy.

The time dependence of the metal abundance Z_g of the gas at several radii in the equatorial plane of Model 9 is shown in Fig. 14. The metal abundance of the gas is not accurately calculated during the later stages, and therefore we have extended the numerical results using rough analytical estimates (dashed curves) based on the simple infall model of Larson (1972), which predicts that the metal abundance asymptotically approaches the assumed yield of 0·02 in a region where continuing inflow of metal-poor gas maintains a nearly constant gas density. There is an initial rapid rise in Z_g associated with the formation of the spheroidal component, and at the larger radii there is a continuing slow increase in Z_g at later

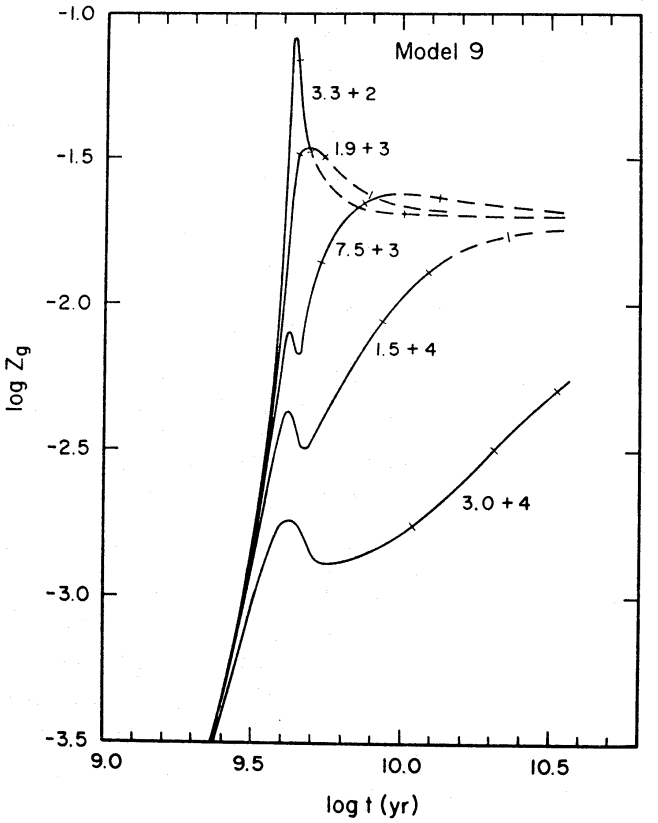

FIG. 14. *The metal abundance Z_g of the gas vs time at five radii in the equatorial plane of Model* 9, *indicated in pc on the curves. The dashed sections of the curves are estimated extrapolations which replace inaccurate numerical results. The three ticks marked on each curve indicate the times at which* 10, 50 *and* 90 *per cent of the final density of stars is attained.*

times associated with continuing slow star formation in the disc. A radial gradient of Z_g is present in the outer part of the disc because the time scale for star formation and metal enrichment increases with radius and becomes very long in the outer part of the disc; however, there is little gradient of Z_g in the inner part of the disc because here the enrichment time scale is very short and Z_g rapidly approaches the yield of 0·02 which is (by assumption) independent of radius. Although the chemical evolution of the present models is influenced by continuing inflow of gas into the disc, the results are similar to those obtained for closed models in which the only continuing gas input is from evolving stars (e.g. Talbot & Arnett 1975); the radial gradient of Z_g results in either case from the radial gradient in the star formation time scale, which increases outward in any reasonable model.

4. CONCLUSIONS

The main conclusion of this work is that the formation of a spiral galaxy with distinct bulge and disc components must proceed in two stages characterized by star formation processes that operate at very different rates: first, there must be a rapid star formation process that forms a spheroidal component within the first ~ 1 or 2 free-fall times, and later a much slower star formation process that allows most of the remaining gas to settle to a disc before forming stars. The required strong decrease in the star formation rate between early and late stages of the

collapse is greater than is predicted by a simple power-law function of gas density or velocity dispersion, and only models in which star formation is somehow more strongly suppressed during the later stages of the collapse yield substantial disc components.

One effect that may produce distinct bulge and disc components is the action of tidal forces in inhibiting star formation during the later stages of the collapse. Models in which this effect is simulated by multiplying the star formation rate by a reduction factor depending on the ratio of gas density to mean density (e.g. Models 4 and 5) develop distinct disc components containing up to one-half of the total mass. The fractional mass in the disc increases with decreasing initial density, but the variation is not great enough to account for the entire Hubble sequence on this basis alone.

A further possibility is that the protogalactic gas has a two-phase structure with dense clouds that rapidly form stars in a spheroidal component, and less dense intercloud gas which does not form stars until it has settled to a disc. The disc:bulge ratio of the resulting system then depends on the fraction of the gas which is initially in dense clouds. Models in which the star formation rate is further reduced to simulate a situation in which only a small fraction of the gas is in dense clouds (e.g. Models 6 and 7) develop disc components containing up to ~ 90 per cent of the total mass.

A plausible way of forming dense condensations and causing rapid star formation in protogalaxies is through the compression produced by high velocity collisions between randomly moving gas clouds or streams. If radiative cooling is sufficiently effective, as seems likely, compression by several orders of magnitude in density is possible when gas clouds collide at velocities of the order of 100 km s^{-1} or more; this degree of compression is much greater than that attainable in galactic discs where the random velocities are typically only ~ 10 km s^{-1}, and this difference may account for the relatively rapid early star formation required to form spheroidal systems. Spheroidal systems may thus form from gas which experiences strong turbulence or cloud collisions, and disc systems may form from more quiescent residual gas in which collisions are less important. Models in which the star formation rate is reduced when the cloud collision time becomes longer than the free-fall time (e.g. Models 8 and 9) develop distinct massive disc components in qualitatively the same way as models with tidally inhibited star formation. The disc:bulge ratio of these models increases with decreasing initial density and decreasing velocity dispersion; this suggests that elliptical galaxies and early-type spirals may form in regions of unusually high density or large velocity dispersion, while late-type spirals with large disc:bulge ratios may form in regions of lower density or smaller random velocities. This expectation is consistent with the fact that elliptical galaxies generally occur in clusters or groups with high density and high velocity dispersion, whereas spiral galaxies occur in the field or in groups with lower density and lower velocity dispersion.

In all models the time scale for the formation of the disc is much longer than the time scale for the formation of the spheroidal component, and increases with increasing distance from the centre. The inner part of the disc always forms first, and the outer parts form progressively later as gas with higher and higher angular momentum settles into the equatorial plane at larger and larger radii. The star formation time scale also increases with radius, so that even after 10^{10} yr the outermost part of the disc is still mostly gaseous. The fraction of the total mass

which at this time is still in the form of gas is ≈ 10 per cent and the star formation rate is $\approx 1\ M_{\odot}\ \mathrm{yr}^{-1}$, in reasonable agreement with the properties of typical spiral galaxies.

The present models are probably too crude to provide fully realistic predictions for the detailed structure and evolution of the discs of spiral galaxies. The radial variation of disc surface density in the models tends to be intermediate between an exponential and a power law, in qualitative but not quantitative agreement with the observed light profiles of many spiral galaxies. The structure of the model discs depends on the angular momentum distribution in the initial spherical protogalaxy and on the assumed viscous transfer of angular momentum during early stages of the collapse, neither of which may be very realistic for most spiral galaxies. Furthermore, the models neglect the probable role of non-axisymmetric instabilities and gravitational torques in redistributing angular momentum in the disc after it has formed; these effects, rather than the initial conditions, may be primarily responsible for establishing the radial structure of galactic discs (Hohl 1975). Similarly, the model rotation curves may not be realistic in detail if the mass distributions in galaxies are affected by processes not considered in the models (such as, for example, violent nuclear activity which prevents the accumulation of too much mass at the centre).

The chemical evolution and the metal abundance distributions of the spheroidal components in these models are similar to those obtained for the elliptical galaxy models of Papers II and III. The early chemical evolution of a large part of the disc is to a first crude approximation described by an infall picture (e.g. Larson 1972) in which star formation and continuing gas inflow proceed simultaneously and produce an approach to a constant metal abundance which is approximately equal to the assumed yield of 0·02. The metal abundance of the innermost part of the disc is higher than this because of prior enrichment of the infalling gas before it reaches the equatorial plane, and the metal abundance of the outermost part of the disc is smaller because of the relatively long time scale for star formation and metal enrichment, but in general the radial abundance gradient in the disc is moderate compared with that in the halo, and the disc is relatively metal rich. In all cases a strong metallicity gradient perpendicular to the plane of the disc is produced by vertical inflow of gas during the formation of the disc, just as radial inflows produce radial abundance gradients in the elliptical galaxy models of Papers II and III. These predictions are in general agreement with the available data on spatial variations of metal abundance in our own and other galaxies. In addition, the predicted continuing inflow of gas during the formation of the disc provides a simple way of explaining the observed paucity of metal-poor stars in the solar neighbourhood (Tinsley 1975).

In summary, models of the present type with appropriate assumptions concerning star formation appear to be capable of reproducing the qualitative distinctions between galaxies of different Hubble types and of explaining in qualitatively correct fashion the structure and stellar population characteristics of typical spiral galaxies. More quantitative predictions will require a more detailed understanding of the many processes involved in the formation and evolution of galaxies, especially the process of star formation. To a large extent, the problem of understanding galaxy formation is the problem of understanding star formation; it will be important, for example, to know more about the circumstances under which stars are formed and the factors which determine the rate of star formation and the stellar

mass spectrum before galaxy formation models with more detailed predictive power than the present simple parameterized models will be possible.

ACKNOWLEDGMENTS

The work described here has benefited greatly from discussions with numerous colleagues, including S. M. Faber, J. R. Gott, I. R. King, A. Oemler, J. P. Ostriker, M. J. Rees, B. M. Tinsley and C. P. Wilson. Special thanks are due to Dr Gordon Lasher for his interest in this project and for arranging my stay at the IBM Watson Research Center, where all of the computing for this project was carried out.

REFERENCES

Bardeen, J. M., 1975. *Dynamics of stellar systems*, IAU Symp. No. 69, p. 297, ed. A. Hayli, Reidel, Dordrecht.
Faber, S. M. & Gallagher, J. S., 1976. *Astrophys. J.*, **204**, 365.
Gott, J. R., 1973. *Astrophys. J.*, **186**, 481.
Gott, J. R., 1975. *Astrophys. J.*, **201**, 296.
Gott, J. R. & Thuan, T. X., 1976. *Astrophys. J.*, **204**, 649.
Hohl, F., 1975. *Dynamics of stellar systems*, IAU Symp. No. 69, p. 349, ed. A. Hayli, Reidel, Dordrecht.
Larson, R. B., 1969. *Mon. Not. R. astr. Soc.*, **145**, 405 (Paper I).
Larson, R. B., 1972. *Nature Phys. Sci.*, **236**, 7.
Larson, R. B., 1974. *Mon. Not. R. astr. Soc.*, **166**, 585 (Paper II).
Larson, R. B., 1975a. *Dynamics of stellar systems*, IAU Symp. No. 69, p. 247, ed. A. Hayli, Reidel, Dordrecht.
Larson, R. B., 1975b. *Mon. Not. R. astr. Soc.*, **173**, 671 (Paper III).
Larson, R. B. & Tinsley, B. M., 1974. *Astrophys. J.*, **192**, 293.
Miller, R. H. & Prendergast, K. H., 1968. *Astrophys. J.*, **153**, 35.
Oort, J. H., 1965. *Galactic structure*, p. 455, eds A. Blaauw & M. Schmidt, University of Chicago Press.
Roberts, M. S. & Rots, A. H., 1973. *Astr. Astrophys.*, **26**, 483.
Sandage, A. R., Freeman, K. C. & Stokes, N. R., 1970. *Astrophys. J.*, **160**, 831.
Talbot, R. J. & Arnett, W. D., 1975. *Astrophys. J.*, **197**, 551.
Tinsley, B. M., 1974. *Astrophys. J.*, **197**, 159.
Tinsley, B. M., 1976. *Astrophys. J.*, **208**, 5.
Turner, E. L., 1976. *Astrophys. J.*, in press.
Van den Bergh, S., 1976. *Tercentenary Symposium of the Royal Greenwich Observatory*, in press.