THE EARLY HISTORY OF OUR GALAXY: DYNAMICS

K. C. FREEMAN

Mount Stromlo and Siding Spring Observatory, Research School of Physical Sciences, Australian National University, Canberra, Australia

Abstract. Dynamical aspects of the early history of our Galaxy are discussed, with emphasis on the two-component nature of disk galaxies. Topics include the Eggen, Lynden-Bell and Sandage picture of the Galaxy's collapse, dynamical problems of the disk, and the bulge component.

1. Introduction

The conventional picture of galaxy formation is that the protosystem, with dynamical time $\sim 10^8$ yr, condensed out of the intergalactic medium and became a stellar system. A key problem in galactic dynamics is to describe and understand how all this happened. There are two complementary approaches to the problem. One is the work on galaxy formation from the expanding Universe. The other is through galactic dynamics, probing backwards to the collapse epoch from what we know now about the internal kinematics and the chemical and mass distributions in galaxies. This knowledge is a composite of detailed data for the Milky Way, plus large-scale data for external systems, and we should not use either in isolation.

We start with the two component nature of the disk galaxies (Figures 1 and 2). We accept the two component picture because, as de Vaucouleurs pointed out (p. 1), the ratio of these two components in a particular galaxy can have any value from zero to infinity. The *disk* we believe is roughly in centrifugal equilibrium, and chemically fairly uniform: probably the only useful information left in the disk from the collapse phase is the angular momentum $\mathscr{H}$ and perhaps the mass-angular momentum distribution $\mathfrak{M}(h)$ (the mass with angular momentum per unit mass $< h$). The *bulge* or halo component has a wide range in chemical abundance, and a significant part of its support is from random stellar motions. The correlation observed of chemical and kinematical properties for halo stars in the Milky Way gives us hope that some residual information about the collapse phase remains.

We are searching for a formation picture: we do not have one yet, and we will be no closer at the end of this talk. The dynamical and chemical aspects of galaxy formation are intertwined: I will attempt to discuss mainly dynamical topics and Peimbert will talk about the chemical aspects.

2. Eggen, Lynden-Bell and Sandage (1962)

This classic paper is largely responsible for our present picture of galaxy formation. Eggen *et al.* (ELS, 1962) collected data for 221 well-observed nearby dwarf stars; they showed that e (orbital eccentricity), W (the velocity in the z-direction) and h (the angular momentum per unit mass) correlate with the ultraviolet excess $\delta(U-B)$, as shown in
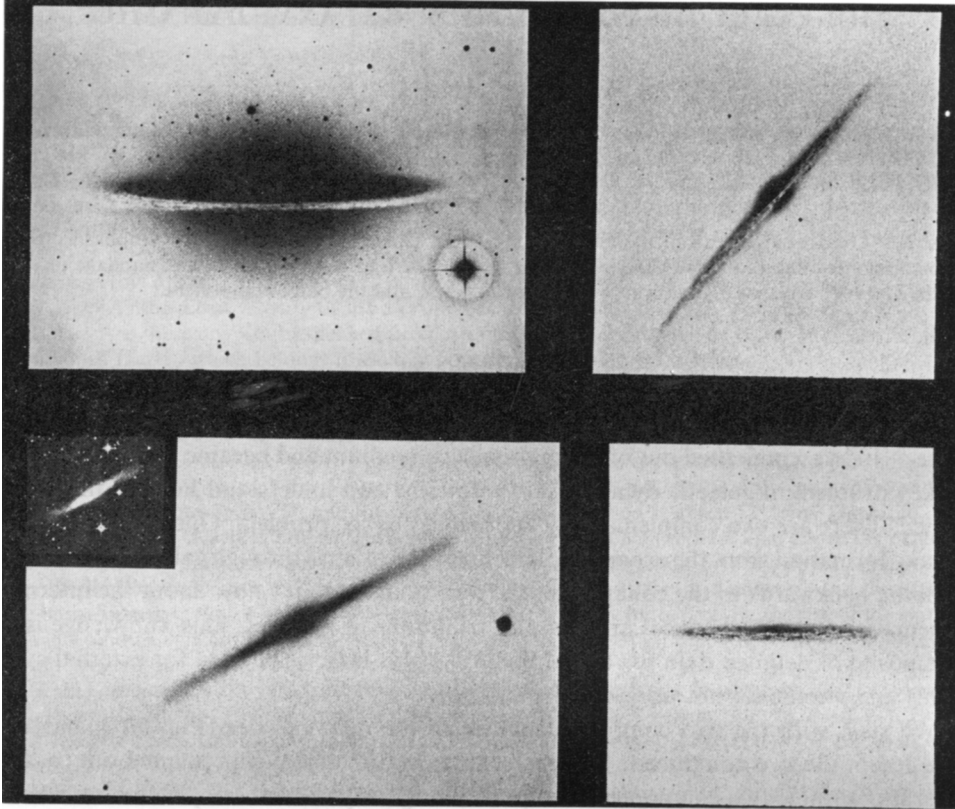
Fig. 1.   Four edge-on disk galaxies from the *Hubble Atlas*, to show their two-component nature.

Figure 3. Ultraviolet excess is a measure of metal abundance, and so presumably of age. In particular, for $\delta(U-B)<0.15$, $e<0.5$, $|W|<50$ km s^{-1} and $h>1500$ kpc km s^{-1} while stars with $\delta(U-B)>0.15$ have large e, low h and a wide range of $|W|$ values. They infer that the younger objects, with smaller $\delta(U-B)$, formed near the plane while the older metal-weak stars formed at any height: this leads to a formation picture in which the Galaxy collapsed to a disk, during or after the birth of the oldest stars. They also infer that this collapse was rapid; because e is an adiabatic invariant, a slow collapse from a rotating system would not produce the highly eccentric orbits seen in Figure 3. The scale of the collapse, in z, can be estimated from the $|W|-\delta(U-B)$ diagram. It is at least a factor 25, from $|z|_{max}\approx10$ kpc for the oldest stars to $|z|_{max}\approx400$ pc for the younger ones.

    There are some uncertainties in the ELS picture. A rapid collapse is not the only way to produce the observed correlation of kinematical and chemical properties. Also, it is not clear how unique the $\delta(U-B)$-age relation is, even for a sample of nearby stars. For example, star formation would probably begin earlier in the dense inner parts of the protogalaxy, and there could have been several generations of star formation with
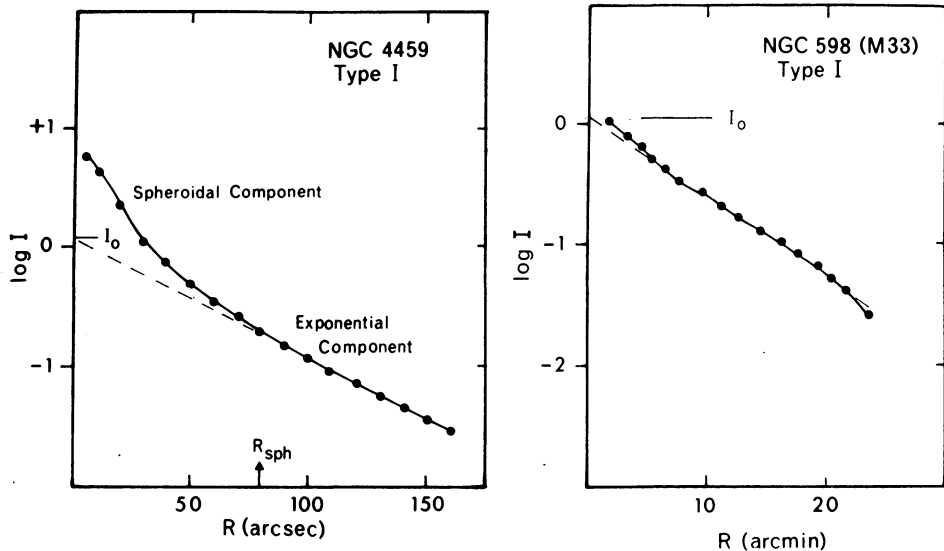
Fig. 2. Two examples of luminosity profiles. I is the surface brightness and R the radius. Note how M33 has no significant spheroidal or bulge component.

consequent chemical enrichment before star formation even began in the outer parts of the system. This picture is reinforced by our knowledge that there is a chemical gradient in the halo (see Kinman, 1959a), and means that some *metal-rich* objects, formed in the inner parts of the system, could be *older* than some *metal-weak* stars seen now in the solar neighbourhood. This may be more plausible than having stars forming first in the most diffuse outer parts of the protosystem. We could even extend the picture to suggest that the low $\delta(U-B)$, low velocity objects of Figure 3 could have formed locally in the dense disk *before* the high $\delta(U-B)$ halo objects formed. In summary, I am saying that $\delta(U-B)$ could be a measure not of age only, but of age and place of formation.

Returning briefly to the possibility that some disk stars formed before the high-velocity halo stars, I want to point out that there is no obvious age problem in this concept. Sandage's (1970) estimate for globular cluster ages is $(10 \text{ to } 12) \times 10^9$ yr, while the old disk cluster NGC 188 has an estimated age of $(8 \text{ to } 10) \times 10^9$ yr: Eggen (1970) has found evolved disk stars which are certainly older than NGC 188. We are haggling about what happened in a time interval which is probably only a few $\times 10^8$ yr: nevertheless, details of the *order* of events within this short time interval are vital for understanding the galaxy formation process, so it is worth questioning whether the metal-weak stars really did form first.

Taking this argument one stage further, consider now galaxies like M33 and the LMC, in which the disk is dominant and the bulge component very weak indeed. In the ELS picture, the chemical enrichment of the disk stars occurs through element building in the first generation (halo) stars. However, in M33 and the LMC, the chem-
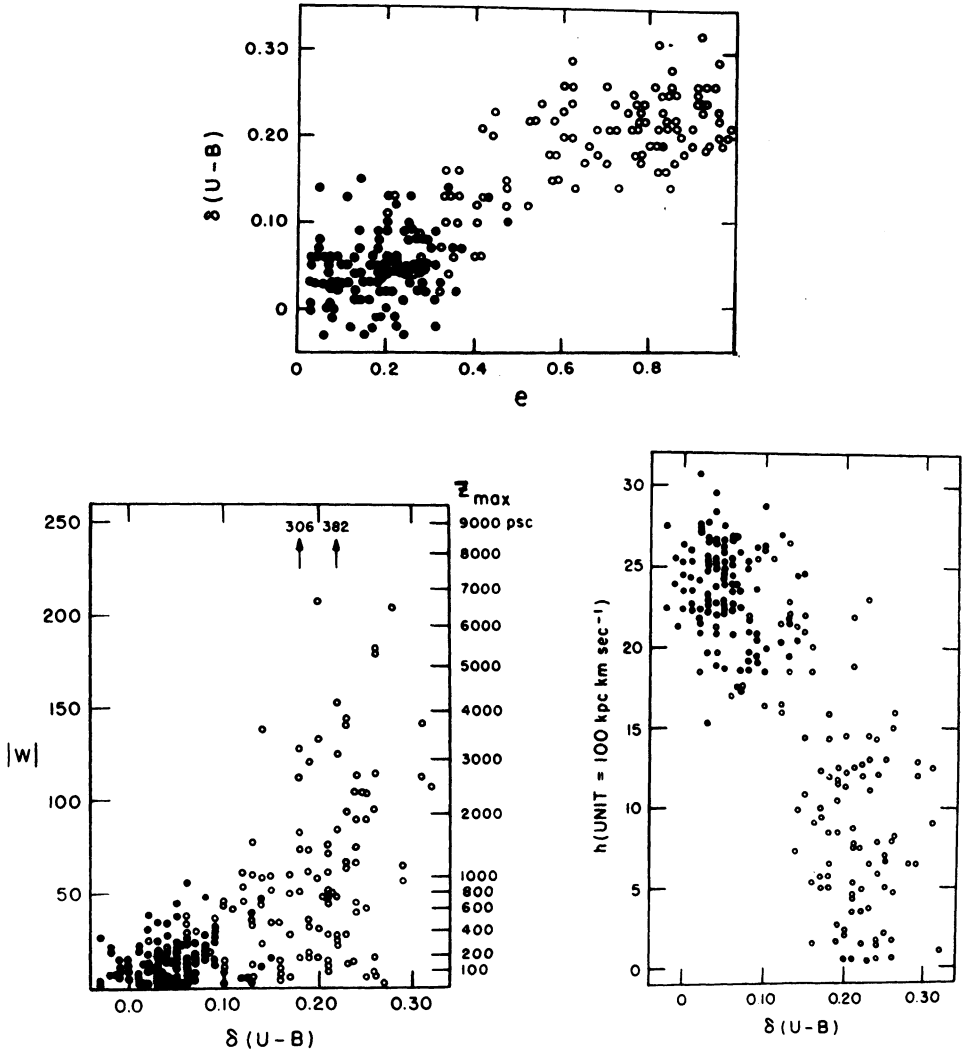
Fig. 3.   Eccentricity, W and h against $\delta(U - B)$ for the ELS stars.

ical abundances in disk objects appear to be fairly close to the solar value: because their bulges are so weak, the disks in these galaxies probably processed their own heavier elements. Given that these disks can do it, maybe the disk of the Milky Way also did its own chemical enrichment independent of the halo population.

3. Dynamical Problems of the Disk

Now we consider the disk of the disk-like galaxies. Much of its interest for galaxy formation lies in its angular momentum content and distribution, which reflects that of the protocloud *if* the collapse was axisymmetric. Discussion of the disk usually

begins from a stellar disk near centrifugal equilibrium. While there is plenty of evidence that the extreme Population I component in the disks of many galaxies is close to circular motion, such evidence for the massive *old disk* component is meagre, and depends on the small asymmetric drift of the old disk stars in the solar neighbourhood. Anyhow, for a start we consider the disk to be cold and flat.

Surface photometry of disk galaxies shows that the disk has *almost always* the exponential surface-brightness distribution seen in Figure 2. From the uniformity of the colour in the disk, it seems likely that the surface *density* distribution $\mu(R)$ is also exponential, i.e. $\mu(R) = \mu_0 e^{-\alpha R}$ in the disk. The origin of this characteristic distribution is not yet really understood. It has two interesting features for formation theory: (i) The density scale μ_0 is observed to be approximately constant from galaxy to galaxy (Freeman, 1970): since the disk mass $\mathfrak{M} = 2\pi\mu_0/\alpha^2$ and its angular momentum $\mathscr{H} = 1.109 \, (G\mathfrak{M}^3/\alpha)^{1/2}$, it follows that $\mathscr{H} \propto \mathfrak{M}^{7/4}$: since $\mathscr{H}$ and $\mathfrak{M}$ are probably conserved through the collapse, this 7/4 law would hold also for the protodisks. Note that the exponent is close to the value 5/3 which follows from elementary theory. (ii) For the exponential disk alone, the mass-angular momentum distribution $\mathfrak{M}(h)$ is almost identical to that for a rigidly rotating sphere of uniform density (or to the rigidly rotating disk D with surface density $\mu(R) = \mu(0)\,(1 - R^2)^{1/2}$ that results from projecting this sphere on to the plane of rotation). Figure 4 shows $\mathfrak{M}(h)$ for the two disks and also their surface density distributions: the two disks have the same $\mathfrak{M}$ and almost the same $\mathscr{H}$ and $\mathfrak{M}(h)$. This shows how $\mathfrak{M}(h)$ alone does not uniquely define the disk, as Mestel (1963) has pointed out: the particular $\mathfrak{M}(h)$ alone does not establish the exponential distribution. Two comments on $\mathfrak{M}(h)$: (a) If $\mathfrak{M}(h)$ is primordial and
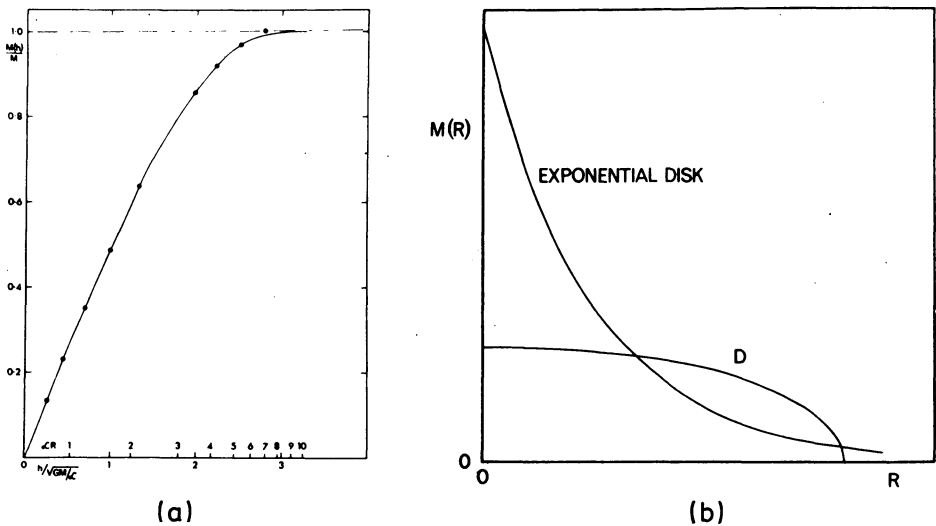

(a)   (b)

Fig. 4.   (a) The full curve is $\mathfrak{M}(h)$ for the exponential disk. The points show $\mathfrak{M}(h)$ for the disk D described in the text. (b) Surface density distribution for an exponential disk and a disk D with the same $\mathfrak{M}$ and almost the same $\mathscr{H}$ and $\mathfrak{M}(h)$.

the collapse was axisymmetric, then all protospirals had a fairly similar $\mathfrak{M}(h)$. On the other hand, if $\mathfrak{M}(h)$ was established by internal torques during the collapse phase, then there existed a very efficient process that drove $\mathfrak{M}(h)$ to the sphere-like distribution during this phase. (b) There do exist a few spiral galaxies in nearby rigid rotation. These would have the characteristic $\mathfrak{M}(h)$ if their mass distributions were like the disk D. We are studying some of these systems now at Mount Stromlo.

All this discussion is partly irrelevant if the disks are not really close to centrifugal equilibrium. The evidence suggesting that the disk could be hot comes from theory. Hohl's (1971) N-body study began with 10^5 particles distributed with the same $\mu(R)$ as the disk D but with enough velocity dispersion σ to suppress local axisymmetric instabilities according to Toomre's rule: $\sigma = \sigma_{min} = 3.36\, G\mu/\varkappa$, where $\varkappa$ is the local epicyclic frequency. He found that the disk was unstable to a large-scale bar-like mode and, by the time the system had settled down, the velocity dispersion had increased to (2 to 6) σ_{min}; Ostriker and Peebles (1973) considered this further, asking: *how can we account for the apparent stability of the Galaxy?* They collected various examples of rotating disk-like systems, including the MacLaurin spheroids, Kalnajs' (1972) exact self-consistent disk, and their own three-dimensional 150 to 500-body integrations. They pointed out that each of these examples was stable only for the ratio $T_{mean}/|\Omega| \lesssim$ $\lesssim 0.14$: here Ω is the potential energy and T_{mean} is the kinetic energy associated with ordered motion, so from the virial theorem the ratio must lie between 0 and 0.5. This means that these systems are stable only if a large part of their kinetic energy comes from random motions. There are then two ways to stabilize the galactic disk: (i) the disk is really hotter than we believe, with random stellar velocities at least as large as the rotational velocity; (ii) the halo mass *interior* to the disk is comparable to the disk mass (this increases $|\Omega|$).

Is either of these ways readily compatible with observation? We see flat systems, like the one in Figure 1, with no significant visual bulge component. Their flatness argues against a high velocity dispersion in the disk, unless this dispersion is very anisotropic with the ratio of component in the plane to component normal to the plane typically 200 km s^{-1} to 30 km s^{-1}. The smallness of the bulge component argues against the presence of a massive stabilizing bulge, unless the M/L ratio for the bulge is very large, well in excess of 100. There is evidence that M/L for the bulge is not so large: (i) the Cambridge results for M31, reported by Baldwin at this meeting (p. 139), suggest that its bulge has $M/L \approx 10$. (ii) Illingworth's unpublished results for globular clusters, which are probably fairly typical examples of the halo population, give M/L values of about 2, even for the most concentrated clusters which would be least affected by the galactic tidal field.

In summary, while theory demands $T_{mean}/|\Omega| \lesssim 0.14$ for the stability of disk systems, we see systems in which the bulge is apparently weak and the velocity dispersion apparently small (compared to the rotational velocity): their value of $T_{mean}/|\Omega|$ would then be significantly larger than 0.14. This is obviously an urgent problem which will hold back our understanding of the formation and dynamics of the disk component until it is sorted out.

4. The Bulge Component

If the two-component picture for disk galaxies is valid, we can regard the bulge component as an elliptical galaxy modified by the presence of the disk: the primary evidence for this is the $R^{1/4}$-dependence of the surface-brightness distribution in the bulges of some nearby systems. Since the galactic halo contributes so much to our present fragmentary knowledge about galaxy formation, it seems vital to get some understanding of the present dynamical state of spheroidal systems. Carrick at Mount Stromlo is now completing a systematic program of surface photometry and dynamical model construction to see if the bulges can be understood as modified elliptical systems.

Some further encouragement for interpreting bulges as modified ellipticals comes from comparing their stellar content. We know that (i) the spectral energy distributions for the bulges of the brighter spirals and the giant ellipticals are similar; (ii) the M/L value of 10 for the bulge of M31 mentioned earlier is close to the value for most normal ellipticals; (iii) the line strengths and colours change radially for many bulges and ellipticals in a way comparable with the abundance gradient in the galactic halo.

It would be very useful to know the $\mathscr{H}/\mathfrak{M}$ ratio for classes of halo objects: note that a large asymmetric drift for a class of objects does not necessarily mean a low value of $\mathscr{H}/\mathfrak{M}$, because the halo population is so extended spatially. Estimates for RR Lyrae stars and globular clusters are probably the most reliable. Kinman (1959b) found an asymmetric drift of 167 ± 30 km s^{-1} for the globular cluster system, and deduced that its $\mathscr{H}/\mathfrak{M}$ was close to the mean value for the whole Galaxy. On the other hand, the most metal-weak RR Lyrae stars have a drift of 220 ± 23 km s^{-1} (see Oort, 1965), and it seems unlikely that their $\mathscr{H}/\mathfrak{M}$ could be so high. The $\mathscr{H}/\mathfrak{M}$ ratio could be relevant to the reason why these objects formed so early in the life of the Galaxy.

Finally, I have a brief comment on the nature of the collapse of the bulge, and how it halted. There appear to be two main pictures. Firstly the Lynden-Bell (1967) picture, in which star formation occurs while the system is far from equilibrium; the bulge then relaxes to its present distribution through mean field relaxation. Secondly, the Larson picture (p. 191), in which the protosystem gets hot enough before star formation to be in equilibrium. Both pictures have their difficulties; again, until we understand better what triggers off star formation, it will be difficult to decide which is right.

5. Conclusion

There are two immediate problems. The first, from Section 3, is the reason for the apparent stability of the disk. The second is the old problem of the process that triggers off star formation in the protogalaxy. There may be a new clue. We now know that the bright blue clusters in the LMC, such as NGC 1866, are like globular clusters in every way but in their stellar content: they are a mere 10^7 to 10^8 yr old. There are no young clusters like these in the Galaxy. If we could understand why these globular clusters can form now in the LMC but not in the Galaxy, we would be closer to a solution of the star formation problem.

References

Eggen O. J.: 1970, *Vistas in Astronomy* **12**, 367.

Eggen, O. J., Lynden-Bell, D., and Sandage, A.: 1962, *Astrophys. J.* **136**, 748.

Freeman, K. C.: 1970, *Astrophys. J.* **160**, 811.

Hohl, F.: 1971, *Astrophys. J.* **168**, 343.

Kalnajs, A.: 1972, *Astrophys. J.* **175**, 63.

Kinman, T. D.: 1959a, *Monthly Notices Roy. Astron. Soc.* **119**, 538.

Kinman, T. D.: 1959b, *Monthly Notices Roy. Astron. Soc.* **116**, 559.

Lynden-Bell, D.: 1967, *Monthly Notices Roy. Astron. Soc.* **136**, 101.

Mestel, L.: 1963, *Monthly Notices Roy. Astron. Soc.* **126**, 553.

Oort, J. H.: 1965, in A. Blaauw and M. Schmidt (eds.), *Galactic Structure*, vol. V of *Stars and Stellar Systems*, University of Chicago Press, p. 455.

Ostriker, J. P. and Peebles, P. J. E.: 1973, *Astrophys, J.* **186**, 467.

Sandage, A.: 1970, *Astrophys. J.* **162**, 841.

DISCUSSION

Larson: It seems to me that the Eggen-Lynden-Bell-Sandage (ELS) argument for a collapse time of 2×10^8 yr is incorrect, since it is based on the assumption that stars are formed in essentially circular orbits. If the stars are formed in a cloud collapsing on a free-fall time-scale, they would have large radial velocities and be in highly eccentric orbits: if so, this would destroy the basis of the ELS argument. The time 2×10^8 yr is inconsistent with the initial radius of 50 kpc suggested by ELS, which implies a free-fall time of about 5×10^8 yr.

My second comment relates to the question of whether the collapse was axisymmetric. I don't know of any reason why the collapse should be axisymmetric, but even if it were, I have some preliminary calculations of axisymmetric protogalaxy collapse which show that the result is the formation of a thin, uniform disk which would be unstable to non-axisymmetric modes. Thus strict axisymmetry would be destroyed in any case, as would be the strict conservation of angular momentum for each mass element.

Oort: I do not know of any observational evidence opposing the idea that a large fraction of the mass of a galaxy is contained in a halo. From the data on stars in the vicinity of the Sun we see that the only *known* stars that contribute significantly to the mass of the halo are intrinsically very faint subdwarfs. As pointed out already in my article (p. 455) in Vol. V of *Stars and Stellar Systems*, there is nothing to oppose the idea that subdwarfs below the limit to which they have so far been found could make up a halo that would contain most of the mass of the Galaxy.

Freeman: The work of Rodgers *et al.* (in preparation) suggests that the mass contribution of objects from the L.F.T. (Luyten's five-tenths) catalogue down to $M = +11$ is very small. These objects have halo-type chemistry and kinematics. It doesn't *prove* anything but is indicative.

Oort: Who knows what the missing mass is made up of? It may be very faint sub-dwarfs.

Freeman: How faint would you want to go?

Oort: Perhaps to $M = +15$.

Freeman: Baldwin's work on the rotation curve of M31 (p. 139) is also indicative.

E. M. Burbidge: From the work of Murray and Sanduleak (*Monthly Notices Roy. Astron. Soc.* **157**, 273, 1972) and others at the Royal Greenwich Observatory, it looks as though the missing mass in the solar neighbourhood may be accounted for by M dwarfs in the disk population.

van der Kruit: If the exponential disks have the same μ_0 for all galaxies, it is only necessary to specify α to find the rotation curve parameters. Is this in contradiction with Brosche's result that the maximum rotation velocity and the radius at which this occurs are almost completely independent?

Freeman: No. The rotation curve is determined by the field of the exponential plus bulge components, and Brosche's result follows because the mass ratio of these two components varies from galaxy to galaxy.

Mestel: Two comments: (i) Miller, Prendergast, and Quirk (in W. Becker and G. Contopoulos (eds), *The Spiral Structure of our Galaxy*, *IAU Symp.* **38**, 365, 1970) found that in order to keep their disk models cool they needed to have powerful dissipation, e.g. by shock formation in gas clouds, which thermalize and radiate away much of the gravitational energy released during instability. Could

this be the answer to the difficulty of Ostriker and Peebles? (ii) In discussing star formation during collapse of the proto-galaxy, one should distinguish between the initial collapse and the final settling down into a flattened structure. A local density increase can amplify by its self-gravitation (cf. Christopher Hunter's work (*Astrophys. J.* **135**, 594, 1962; **139**, 570, 1964) on fragmentation in a collapsing sphere) while still *retaining* its energy of z-motion, presumably forming ultimately a globular cluster, with its characteristic large-amplitude z-motion. The disk population by contrast would form from gas which has lost the bulk of its z-energy before break-up into stars.

Wright: In 1969 Mrs Heidmann (*Astrophys. Letters* **3**, 153) showed that $H \propto M^{5/3}$, while your result indicates that $H \propto M^{7/4}$. Would you like to comment on this?

Secondly, if one puts both these results in a mass-radius form then they are respectively $M \propto R^3$ and $M \propto R^2$. Additionally, for elliptical galaxies Fish (*Astrophys. J.* **139**, 284, 1964) has shown that $M \propto R^2$ (essentially). It seems, therefore, rather dismaying that centrifugal equilibrium and the rise of electron opacity (as considered by Fish) lead to such similar results even though the physical processes are so different.

Freeman: $(7/4-5/3)/(5/3)$ is only 0.05, and I doubt whether my data or Mrs Heidmann's could differentiate between these two exponents. You realise also that there is a near-indeterminateness in the *observational* estimations of this exponent which makes these estimations yet more uncertain than they appear.

Miller: There are some technical points that must be treated carefully in interpreting stability results from n-body calculations. In particular, numerical instabilities may mask physical effects. Our own n-body calculations gave similar results to Hohl's; I discussed these points in some detail three years ago at Cambridge (*IAU Colloq.* **10**), so won't go into detail here. However, recent numerical experiments have shown that we can sort out some of these effects. A strictly axisymmetric n-body integration showed a stability threshold and confirmed Toomre's numerical values, leading to the conclusion that we can build stable n-body integrations and that the failure to show a stability threshold in our (and Hohl's) calculations is most likely due to non-axisymmetric modes.

Some care is required in applying these stability criteria. The paper of Ostriker and Peebles is based on maintenance of stability in an axisymmetric system. But no-one believes that our Galaxy is axisymmetrical. It is possible that a non-axisymmetric system is more stable than an axisymmetrical system and one could argue that spiral features arise from some such process. A crucial question is what are the consequences of an instability? Does a system gently shift to a more stable configuration, or does something catastrophic happen? So far, we have no answers to these questions.

Gott: Traditionally, the velocity envelope of the highest velocity stars has been taken to indicate the escape velocity for the Galaxy, and then the ratio of escape velocity to circular velocity is found to be less than the Keplerian value of $\sqrt{2}$. From this it has been concluded that the mass in the Galaxy is primarily in a disk configuration. Since the presence of a massive spheroidal halo would presumably push this ratio toward the Keplerian value, it may be possible to place upper limits on the mass of a spheroidal halo component. It would be worthwhile in this context to re-examine the usual assumption that the velocity envelope of the highest velocity stars does in fact indicate the escape velocity.