SECTION III.5

THE OLD POPULATION

Friday 3 June, 1045 - 1140

Chairman: K.C. Freeman

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

Freeman (right) and Allen talking over beer. Background: Fujimoto and Paris Pişmiş. GSS

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

THE FORMATION AND EARLY EVOLUTION OF THE MILKY WAY

S. Michael Fall

Institute of Astronomy, Cambridge, England

I. INTRODUCTION

In broad outline, the traditional picture for the formation of the Milky Way can be summarized as follows. The proto-galaxy consisted of a slowly rotating cloud of metal-free gas that cooled by bremsstrahlung and recombination radiation. As the internal pressure of the gas decreased, it collapsed in stages with smaller dimensions, faster rotation velocities and flatter shapes until it reached centrifugal support in a fundamental plane. At the same time, the gas was progressively depleted by the formation of stars and enriched with heavy elements by the ejecta from previous generations. The result is a general correlation between the kinematic properties, chemical compositions and relative ages of the stellar populations within the Galaxy. This picture was formulated at the Vatican symposium by Oort (1958) and others and was elaborated by Eggen, Lynden-Bell & Sandage (1962), Sandage, Freeman & Stokes (1970), Gott & Thuan (1976), Larson (1976) and others. Much of the recent work on galaxy formation has been an attempt to extend these ideas to a more comprehensive picture that includes large quantities of dark matter. The purpose of this article is to review several topics concerning the collapse phase in the evolution of the Milky Way.

In the following discussion, the Galaxy is assumed to consist of three main components: an exponential disc with a scale-radius of about 4 kpc, a de Vaucouleurs spheroid with an effective radius of about 3 kpc and an isothermal halo with much larger dimensions. This description, which relies heavily on extragalactic studies, is consistent with the available star counts and kinematic data for the Milky Way (Bahcall & Soneira 1980, Caldwell & Ostriker 1981). The traditional members of population II are provisionally assigned to the spheroid, although a distinction between globular clusters and field stars may be useful for some purposes. Suitable candidates for the dark matter in the halo include stellar remnants, low-mass stars and a variety of elementary particles. Following the arguments of White & Rees (1978), Fall & Efstathiou (1980), Silk & Norman (1981), Faber (1982), Gunn (1982) and

H. van Woerden et al. (eds.), The Milky Way Galaxy, 603–610.
© 1985 by the IAU.

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

others, the dark matter is assumed to be pre-galactic and the spheroid and disc are assumed to have formed from residual gas that collapsed within this arena. The halo itself may have formed by hierarchical clustering or as part of a pancake, depending on whether the primordial spectrum of density perturbations had a cutoff below or above galactic scales.

II ROTATION OF THE PROTO-GALAXY

The radial extent of the collapse depends critically on the rotation of the proto-galaxy. This can be quantified in terms of the dimensionless spin parameter $\lambda \equiv J|E|^{\frac{1}{2}}G^{-1}M^{-5/2}$ where J, E and M are respectively the total angular momentum, energy and mass of the system and G is the gravitational constant. It is instructive first to consider the problem without a massive halo. In this case, a reasonable model for the proto-galaxy just prior to collapse is a sphere with uniform density and negligible kinetic energy. In terms of the maximum radius r_t at turnaround, the specific angular momentum is $J/M = (5GMr_t/3)^{\frac{1}{2}}\lambda$ and the free-fall time is $\tau_{ff} = \pi(r_t^3/8GM)^{\frac{1}{2}}$ for $\lambda << 1$. A self-gravitating exponential disc with a scale radius α^{-1} has $J/M = 1.11(GM/\alpha)^{\frac{1}{2}}$ and a maximum circular velocity $v_c = 0.65(G\alpha M)^{\frac{1}{2}}$ at the radius $2.2\alpha^{-1}$ (Freeman 1970). Thus, conservation of J/M during the collapse implies

$$\alpha r_t = 0.74\lambda^{-2} \qquad \text{and} \qquad \alpha v_c \tau_{ff} = 0.46\lambda^{-3} \tag{1}$$

These results are not affected by any internal redistribution of angular momentum that may occur during or after the formation of the disc.

The relation between the initial spin and radius of the proto-galaxy without a halo is plotted as the upper curve in Fig. 1 and labelled with several values of the free-fall time for α^{-1} = 4 kpc and v_c = 220 kms^{-1}. If the Galaxy formed by hierarchical clustering, it would have been endowed with some rotation by the tidal torques of neighbouring objects before it collapsed (Peebles 1969). The dashed lines in Fig. 1 show the 10, 50 and 90 percentile points in the distribution of spins generated by cosmological N-body simulations with Poisson initial conditions (Efstathiou & Jones 1979). For the median value, $\lambda \approx 0.065$, the initial radius of the proto-galaxy is $r_t \approx 700$ kpc and the corresponding free-fall time is $\tau_{ff} \approx 3 \times 10^{10}$yr, neither of which is acceptable. Unfortunately the distributions of spins for different initial conditions are not yet known with certainty and there are no reliable predictions for the pancake picture. As Fig. 1 shows, only a small range near $\lambda \approx 0.15$ is compatible with a spread of less than a few billion years in the ages of globular clusters.

The situation is changed dramatically by the presence of a massive halo. If this is modelled as a singular isothermal sphere out to some truncation radius r_t, then the specific angular momentum is $J/M = \sqrt{2}v_c r_t \lambda$ and the free-fall time is $\tau_{ff} = (\pi/2)^{\frac{1}{2}}(r_t/v_c)$ for $\lambda << 1$. As

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

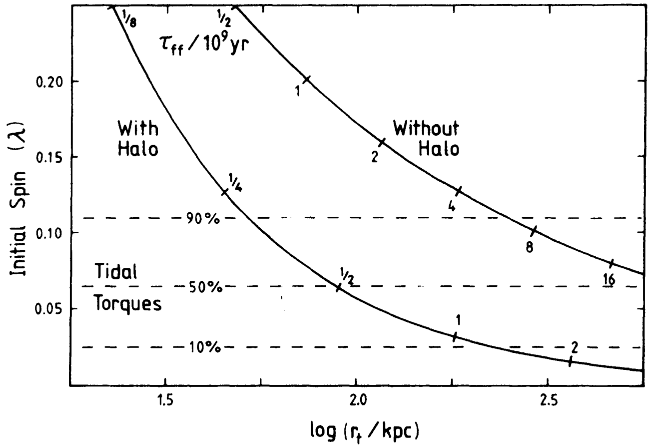

Fig. 1. Initial spin against initial radius for the proto-galaxy. The upper curve is from eqn. (1), the lower curve is from eqn. (2) and the numbers are free-fall times in units of 10^9 yr.

before, v_c denotes the velocity of test particles in circular orbits, which is generally smaller than the average rotation velocity of the dark matter. An exponential disc with negligible self-gravity in this halo has $J/M = 2v_c/\alpha$, which implies

$$\alpha r_t = \sqrt{2}\,\lambda^{-1} \qquad \text{and} \qquad \alpha v_c \tau_{ff} = \sqrt{\pi}\,\lambda^{-1} \qquad (2)$$

if both components have the same specific angular momenta. These results are plotted as the lower curve in Fig. 1 and they agree with more detailed calculations that include the gravitational field of the disc (Fall & Efstathiou 1980). For $\lambda \approx 0.065$, the initial radius is $r_t \approx 90$ kpc and the corresponding free-fall time is $\tau_{ff} \approx 5 \times 10^8$ yr, both of which are quite acceptable. In fact, they are similar to the values deduced by Eggen et al. (1962) from the motions of old stars in the solar neighbourhood.

III. FORMATION OF SUBSTRUCTURE IN THE SPHEROID

Some substructure may have been present before the collapse, especially if the Galaxy formed by hierarchical clustering, but much of it may have developed during the collapse. Gunn (1980) has pointed out that the minimum mass for gravitational instability behind isothermal shocks in the spheroid is comparable with the masses of present-day globular clusters. The following arguments suggest that such objects might form under fairly general conditions. Any irregularities in the flow are likely to heat some of the gas up to the virial temperature of the halo

$$T_h = \mu_h v_c^2/2k \approx 1.8 \times 10^6 \text{ K} \qquad (3)$$

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

where $\mu_h \approx 0.6$ amu is the mean molecular weight for an ionized mixture of hydrogen and helium and k is Boltzmann's constant. Since the dark matter is assumed to provide most of the gravitational acceleration, the free-fall time is $\tau_{ff} = (\pi/2)^{\frac{1}{2}}(r/v_c)$ from a radial position r in the halo. At a density ρ_h, the hot gas will radiate energy on a time-scale $\tau_{cool} = 3\mu_h kT_h/2\rho_h\Lambda(T_h)$, where Λ is the usual cooling function. It will remain hot if the condition $\tau_{ff} \lesssim \tau_{cool}$ is satisfied, which implies

$$\rho_h \lesssim 1.7\mu_h^{\frac{1}{2}}(kT_h)^{3/2}/r\Lambda(T_h) \approx 1.6 \times 10^{-3}(r/\mathrm{kpc})^{-1}M_\odot \mathrm{pc}^{-3} \quad (4)$$

for $\Lambda(T_h) \approx 2 \times 10^{-23}$ erg cm^3 s^{-1}.

Any gas that is more dense than the limit specified by eqn. (4) will cool rapidly and then be compressed to even higher densities by the hot gas. In this sense, the medium is thermally unstable and is likely to produce cold clouds at the temperature $T_c \approx 1.0 \times 10^4$K where the cooling function drops precipitously; they will be nearly neutral with a mean molecular weight $\mu_c \approx 1.2$ amu. In pressure balance, the densities in the hot and cold phases are related by

$$\rho_c/\rho_h = (\mu_c/\mu_h)(T_h/T_c) \approx 360, \quad (5)$$

irrespective of position in the halo. A compressed cloud is gravitationally unstable if its mass exceeds

$$M_{crit} = 1.18(\rho_h/\rho_c)^2(kT_h/\mu_h G)^{3/2}\rho_h^{-\frac{1}{2}} \gtrsim 3.2 \times 10^6(r/\mathrm{kpc})^{\frac{1}{2}}M_\odot, \quad (6)$$

where the inequality follows from eqns. (4)and (5). At the effective radius of the spheroid, $r \approx 3$ kpc, the internal density of the clouds is $\rho_c \lesssim 0.2\ M_\odot \mathrm{pc}^{-3}$ and their critical mass is $M_{crit} \gtrsim 5 \times 10^6\ M_\odot$. In comparison with globular clusters, the first prediction is low and the second is high, but they are both correct to within an order of magnitude. Moreover, the velocity dispersion in the cold gas, about 8 kms^{-1} in one dimension, is similar to that in globular clusters.

The free-fall time of a cold cloud, when approximated as a sphere of uniform density, is $\tau_{ff} = (3\pi/32G\rho_c)^{\frac{1}{2}}$ and its cooling time is $\tau_{cool} = 3\mu_c kT_c/2\rho_c\Lambda(T_c)$; thus

$$\tau_{cool}/\tau_{ff} = 2.8\mu_c kT_c(G/\rho_c)^{\frac{1}{2}}/\Lambda(T_c) \quad (7)$$

$$\gtrsim 3.2(r/\mathrm{kpc})^{\frac{1}{2}}[\Lambda(T_c)/10^{-28}\ \mathrm{erg\ cm^3 s^{-1}}]^{-1}.$$

At temperatures below 1×10^4K, the cooling function is $\Lambda(T_c) \lesssim 4 \times 10^{-28}$ erg cm^{3}s^{-1} for a fractional ionization $x = 0$ and a metallicity $Z \lesssim 0.01Z_\odot$ or for $x = 0.1$ and $Z \lesssim 0.001Z_\odot$ (Dalgarno & McCray 1972). When the metallicity is this low, τ_{cool} exceeds τ_{ff} and any clouds with masses greater than M_{crit} will contract quasistatically at roughly constant temperature. Since the Jeans mass within such a cloud varies as $\rho_c^{-\frac{1}{2}}$, it will fragment into a bound collection of smaller objects and perhaps ultimately into stars. If the condition $\tau_{cool}/\tau_{ff} \gtrsim 1$ is not

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

satisfied and if the clouds are not heated by some means, they will cool below 1 x 10^4K before they collapse. In this case, the resulting substructures will have higher densites and smaller masses than those given above. It appears likely, however, that radiation from the hot gas and star formation within the cold clouds will keep their temperatures near 1 x 10^4K even when the metallicity is high and the cooling times are short. This possibility will be discussed in detail elsewhere (Fall & Rees, in preparation).

IV. DISRUPTION OF SUBSTRUCTURE IN THE SPHEROID

Several processes could play a role in the disruption of substructure in the spheroid. The tidal field of the halo insures that the internal density of any stellar or gaseous object exceeds

$$\rho_{lim} = (3/4\pi G)[\omega^2 + (v_c/r)^2] \gtrsim 2.7(r/kpc)^{-2} M_\odot pc^{-3}, \quad (8)$$

where ω is the angular velocity and r is the radial position with respect to the Galactic centre. Even if a bound object is not tidally limited at the time of formation, it may become so near the pericentre of its orbit. If the mean density of the object is below ρ_{lim} it will lose some mass, and if the central density is below ρ_{lim} it will be completely destroyed. The debris of tidal disruption are therefore more likely to have radial orbits than the substructures that manage to survive this process. A comparison of eqns. (4) and (5) with eqn. (8) shows that the cold clouds produced by a thermal instability are particularly vulnerable and those with $r \lesssim 5$ kpc might even be prevented from collapsing if their central densities are low. As a result of tidal limitation, the clouds that do survive will have lower masses, higher mean densities and a greater resemblance to present-day globular clusters.

Another process that might disrupt some substructures while they are still mainly gaseous is star formation (Gunn 1980). The heat supplied by a few supernovae is enough to unbind a cloud with a mass of order $10^6 M_\odot$ and a velocity dispersion of order 10 kms^{-1}. If the metallicity is very low, this energy cannot be radiated efficiently at the densities of interest until the temperature reaches 1 x 10^4 K. Any clouds with much lower temperatures are therefore likely to become unbound whereas those near 1 x 10^4 K will almost certainly survive. However, the relevance of star formation is doubtful in the context of a thermal instability, because when the metallicity is low enough to permit the disruption of clouds with temperatures below 1 x 10^4 K, it is probably also low enough to prevent their formation. Any substructures that survive the relatively rapid effects of tidal limitation and star formation might be disrupted on much longer time-scales by several other processes (Fall & Rees 1977 and references therein). These include dynamical friction against the halo and the evaporation of stars by internal relaxation and external impulses. The effects that such processes have on present-day globular clusters are controversial, but they may have helped to narrow the range of surviving substructures.

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

The degree to which the spheroid consists of disrupted substructure might be revealed by any differences between globular clusters and field stars. The space distributions of both kinds of objects can be fitted by density profiles with similar effective radii, but obscuration prevents a detailed comparison in our Galaxy (de Vaucouleurs & Pence 1978). In some elliptical galaxies, however, the distribution of globular clusters is slightly more extended than the diffuse stellar component (Forte, Strom & Strom 1981). The velocity ellipsoid of metal-poor RR Lyrae variables in the solar neighbourhood, which may be typical of field stars in the spheroid, is elongated in the radial direction (Woolley 1978). In contrast, the velocity ellipsoid of globular clusters in our Galaxy appears to be nearly isotropic (Frenk & White 1980). These results are qualitatively consistent with the notion that substructures on radial orbits were preferentially disrupted and shed more of their mass than those on circular orbits. It would be interesting to refine this suggestion and make a quantitative comparison with the observations.

V CHEMICAL ENRICHMENT IN THE SPHEROID

The natural starting point for any discussion of chemical enrichment is the so-called simple model. It postulates that the initial composition of the gas is metal-free, the system is closed and well-mixed, the stellar mass function is constant and the recycling of enriched gas is instantaneous. In this case, the metallicity Z is related to the fraction of mass in gas μ by the familiar expression

$$Z = -y \, \ln \mu , \tag{9}$$

where y is the nuclear yield. If the enrichment is stopped suddenly by the complete removal of gas with metallicity Z_m, then the fraction of stars $f(Z)dZ$ with metallicities between Z and $Z + dZ$ is given by

$$f(Z) = y^{-1} \exp(-Z/y)/[1 - \exp(-Z_m/y)] \tag{10}$$

for $Z < Z_m$ and $f(Z) = 0$ for $Z > Z_m$. This distribution is entirely the result of nucleosynthesis at different times rather than different places, and it could not be expected to apply if the enrichment occurred in isolated substructures. The main justification for applying the simple model to the spheroid is that the gas and stars might behave as a co-moving system during a free-fall phase of the collapse.

The observed distribution of metallicites for globular clusters beyond the solar circle and high-velocity subdwarfs in the solar neighbourhood can be fitted by an exponential function (Hartwick 1976, Searle & Zinn 1978). There is some evidence for a deficiency of objects with $Z \lesssim 0.003Z_\odot$, but an estimate of their frequency depends on how the samples are selected and calibrated (Bond 1981, Hartwick 1983). Even so, this sets a firm upper limit on any pre-galactic enrichment by stars from the hypothetical population III. For $Z \gtrsim 0.003Z_\odot$, the observed

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

distribution is reproduced by eqn.(10) with the parameters $y \approx 0.02Z_\odot$ and $Z_m \approx 0.1Z_\odot$. The first is an order of magnitude smaller than the yield for the disc and the second implies, through eqn. (9), that nearly all of the gas was converted into stars during the formation of the spheroid. Unfortunately, the conditions that determine the stellar mass function in different environments are known too poorly to say whether the inferred value of y is a problem for the simple model. The depletion of gas in the spheroid is compatible with the formation of the disc by late infall, but it would be a problem if both components formed during a rapid collapse.

As an alternative to the simple model, Hartwick (1976) introduced a model in which the rate of gas removal is a constant c times the rate of star formation. In this case eqns.(9) and (10) are still valid, if μ is interpreted as the ratio of the mass in gas to the initial mass of the system and the nuclear yield y is replaced by the effective yield

$$y_{eff} = y(1 - R)/(1 - R + c), \tag{11}$$

where R is the fraction of mass returned to the gas by each generation of stars. With the conventional parameters for the disc, $y \approx 0.8Z_\odot$ and $R \approx 0.2$, the fit to f(Z) for the outer parts of the spheroid implies $c \approx 30$. Since the removal of this much gas from the proto-galaxy seems unlikely, Hartwick argues that it was transferred from the spheroid to the disc while they were in the process of formation. The fact that c is roughly equal to the ratio of the masses of the two components is consistent with this interpretation. Although the model is plausible, it does not explain why the rate that gas accumulates in the disc should be directly proportional to the rate that stars form in the spheroid. Clearly our understanding of chemical enrichment during the formation of the Galaxy is still rudimentary.

ACKNOWLEDGEMENTS

I thank Ray Carlberg, Ken Freeman, Cedric Lacey and Martin Rees for stimulating discussions of the material in this article.

REFERENCES

Bahcall, J.N. & Soneira, R.M.: 1980, Astrophys. J. Suppl. 44, 73.
Bond, H.E.: 1981, Astrophys. J. 248, 606.
Caldwell, J.A.R. & Ostriker, J.P.: 1981, Astrophys. J. 251, 61.
Dalgarno, A. & McCray, R.A.: 1972, Ann. Rev. Astron. Astrophys. 10, 375.
de Vaucouleurs, G. & Pence, W.D.: 1978, Astron. J. 83, 1163.
Efstathiou, G. & Jones, B.J.T.: 1979, Mon. Not. Roy. Astr. Soc. 186, 133.
Eggen, O.J., Lynden-Bell, D. & Sandage, A.R.: 1962, Astrophys. J. 136, 748.
Faber, S.M.: 1982, in "Astrophysical Cosmology", eds. Bruck, H.A., Coyne, G.V. & Longair, M.S., Pontificia Academia Scientiarum, p. 191.
Fall, S.M. & Efstathiou, G.: 1980, Mon. Not. Roy. Astr. Soc. 193, 189.
Fall, S.M. & Rees, M.J.: 1977, Mon. Not. Roy. Astr. Soc. 181, 37P.

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press

Forte, J.C., Strom, S.E. & Strom, K.M.: 1981, Astrophys. J. 245, L9.
Freeman, K.C.: 1970, Astrophys. J. 160, 811.
Frenk, C.S. & White, S.D.M.: 1980, Mon. Not. Roy. Astr. Soc. 193, 295.
Gott, J.R. & Thuan, T.X.: 1976, Astrophys. J. 204, 649.
Gunn, J.E.: 1980, in "Globular Clusters", eds. Hanes, D. & Madore, B., Cambridge University Press, p. 301.
Gunn, J.E.: 1982, in "Astrophysical Cosmology", eds. Bruck, H.A., Coyne, G.V. & Longair, M.S., Pontificia Academia Scientiarum, p. 233.
Hartwick, F.D.A.: 1976, Astrophys. J. 209, 418.
Hartwick, F.D.A.: 1983, Mem. Soc. Astr. Italiana 54(1), 51.
Larson, R.B.: 1976, Mon. Not. Roy. Astr. Soc. 176, 31.
Oort, J.H.: 1958, in "Stellar Populations", ed. O'Connell, D.J.K., North Holland, Amsterdam, p. 415.
Peebles, P.J.E.: 1969, Astrophys. J. 155, 393.
Sandage, A., Freeman, K.C. & Stokes, N.R.: 1970, Astrophys. J. 160, 831.
Searle, L. & Zinn, R.: 1978, Astrophys. J. 225, 357.
Silk, J. & Norman, C.: 1981, Astrophys. J. 247, 59.
White, S.D.M. & Rees, M.J.: 1978, Mon. Not. Roy. Astr. Soc. 183, 341.
Woolley, R.: 1978, Mon. Not. Roy. Astr. Soc. 184, 311.

DISCUSSION

F.H. Shu: In your discussion of the disruption of substructure in the Galaxy, it is not clear to me that the processes would all lead to complete destruction. Would there be any remnants of objects that might have resided outside the part of parameter space where we now find the globular clusters?

Fall: That is very hard to say. It could just be that they become dense and that we do not see them. The argument is very crude, only to orders of magnitude. It is not even clear, for example, that the evaporation of stars from a globular cluster actually leads to its total disintegration.

T.S. Jaakkola: How can the disk angular momentum be constant in time, as in your model? Should not there be friction between the disk and the massive corona? To me, the fast rotation of spirals indicates that rotation is not a relic effect, but rather, that galaxies are "rotationally active".

Fall: Suppose the halo of the Galaxy were strongly triaxial, then it probably would exert some torque on the disk material, and could act to spin it up or down. The assumption made here is that this is a small effect, hence the disk angular momentum is approximately constant.

https://doi.org/10.1017/S0074180900243234 Published online by Cambridge University Press