

Gas and Galaxy Formation

P. E. J. Nulsen

Department of Engineering Physics, University of Wollongong,
Wollongong, NSW 2522, Australia
pejn@uow.edu.au

Received 1998 November 2, accepted 1999 February 1

Abstract: The theory of galaxy formation is reviewed briefly. From the evidence of clusters today, the primordial gas fraction was 20% or more. Thus, while the Universe is dominated by dark matter, gas plays an appreciable role in galaxy formation. Collapses of dwarf protogalaxies produce predominantly cold gas. It is argued that, in such cold collapses, the collapsed gas is largely self-gravitating. As a result, gas processes play a critical role in determining the visible structure of galaxies.

Keywords: galaxies: formation

1 Introduction

We start with a brief and highly selective review of the theory of galaxy formation, with a focus on the importance of gas processes. It is then argued that most of the gas in dwarf protogalaxies will be self-gravitating, so that gas processes play a major role in determining the visible structure of dwarf galaxies.

Section 2 discusses the arguments for a dark matter dominated hierarchical collapse. While the dominant form of matter is dark, it is argued that the primordial gas fraction must be at least 20% to account for the high gas fractions in clusters. Section 3 outlines why fully numerical models for galaxy formation may not be the most reliable. It is argued in Section 4 that the early collapse of, at least, the larger protodwarf galaxies leads to major starbursts and outflows. Finally, in Section 5, it is shown that the disks of most galaxies are largely self-gravitating. In particular, this means that we should not expect gravitating matter in the region of the disk to be well-fitted by the universal dark matter profile of Navarro, Frenk & White (1997 — NFW), which models collisionless galaxy halos.

2 Standard View of Galaxy Formation

It is widely accepted that galaxies were formed in a hierarchical collapse, dominated by non-baryonic dark matter. We begin by considering the arguments that support this view.

The evidence that the Universe is dominated by dark matter has been well documented (e.g. Trimble 1987; Carr 1994). It can be summarised by the statement that, on large scales, almost all virialised objects have high mass-to-light ratios. Furthermore, in order to account for primordial nucleosynthesis, the baryonic contribution to the closure density must be relatively small (e.g. Walker et al. 1991). Using

their determination of the primordial deuterium abundance, Burles & Tytler (1998) found that the contribution of baryons to the density parameter is

$$\Omega_b \simeq 0.019h^{-2}, \quad (1)$$

where h is the Hubble constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Other determinations generally give a higher deuterium abundance, hence smaller values for Ω_b (e.g. Webb et al. 1997). The density parameter for all matter, Ω_0 , is still uncertain, but values about 0.3 are consistent with a range of recent cosmological data (e.g. Carlberg et al. 1997; Merchan et al. 1998; Cavaliere, Menci & Tozzi 1998; Donahue et al. 1998; Lineweaver 1998; but see Gross et al. 1998; Blanchard & Bartlett 1998). Comparing this to Ω_b , we see that the bulk of the dark matter must be non-baryonic.

The argument for a hierarchical collapse is well supported by both theory and observations. It would require a highly contrived distribution of density fluctuations to avoid a collapse hierarchy by forming all current structures in single collapses. The much simpler alternative is a hierarchical collapse. There is also considerable evidence for continuing hierarchical collapse, such as mergers between galaxies (Schweizer 1986) and between clusters of galaxies (Forman & Jones 1982).

While dark matter dominates, it is baryons that form all visible structure in the Universe. Clusters of galaxies are the largest virialised objects, so the baryon fractions in clusters are most likely to be representative of the Universe as a whole. Any warm dark matter (e.g. Gross et al. 1998) contributes more to the mass of the larger objects, so that the baryon fraction may decrease with mass. If so, the baryon fraction in clusters would be lower than average, but there are no obvious plausible

mechanisms to make the baryon fraction in clusters higher than average.

Most determinations of the gas fraction in clusters (e.g. White, Jones & Forman 1997) apply to a region within about 1 Mpc of the cluster centre, but it has been noted on many occasions that the gas fraction is a rising function of radius. In a recent determination, Ettori, Fabian & White (1998) found that the gas fraction within $0.85r_{200}$ of the centre of the Perseus cluster is about 30% (for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$; cluster gas fraction scales as $h^{-3/2}$). Here, r_{200} is the radius within which the mean density of the cluster is 200 times the background density. Although it is possible to contrive to make the numbers less certain e.g. Gunn & Thomas (1996), the mass of the X-ray emitting gas in clusters is quite well determined, so that the main source of uncertainty in the gas fraction is the total mass of the cluster. Although there has been some dispute about the accuracy of X-ray mass determinations (principally on small scales), they give results that agree with other methods on large scales (e.g. Evrard 1997). In order for such high gas fractions to be consistent with the baryon density limit from primordial nucleosynthesis (1), the density parameter Ω_0 must be low (White et al. 1993).

3 Numerical Simulations of Galaxy Formation

On the face of it, the combined numerical N -body and hydrodynamic simulations of galaxy formation implement physical collapse models most accurately (e.g. Thomas et al. 1998) and so ought to give the most reliable results. Unfortunately, they are hampered by a number of problems. First, there is our ignorance of what controls star formation. Since the stars endow galaxies with many of their observed properties, the process of star formation is critical to galaxy formation. Yet, we do not know what governs the rate of star formation or the initial mass function. This makes the handling of star formation equally uncertain in all models, so that numerical accuracy is not the advantage that it might be.

The other problems of the full numerical simulations are largely artifacts of limits on computing resources which limit the resolution of the simulations (e.g. Weinberg, Hernquist & Katz 1997). Ways have been found around most such problems, but they raise the issue that some problems have yet to be identified. An example of a problem that is not widely known is the excessive radiative cooling in shocks in numerical simulations (Maguire 1996). This arises because the numerically simulated shocks are orders of magnitude thicker than the real ones, keeping gas at intermediate temperatures, where the cooling rate is high, for much longer than in the real shocks. The effect is that some numerically simulated collapses produce considerably less hot

gas than they should. This effect is greatest when the cooling time of the shocked gas is comparable to the dynamical time, that is, in the collapse of 'normal' protogalaxies (Rees & Ostriker 1977).

A popular alternative to the full numerical models are the semi-analytical models (White & Frenk 1991; Kauffmann, White & Guiderdoni 1993; Baugh, Cole & Frenk 1996; Nulsen & Fabian 1997; Somerville & Primack 1998). In these models the collapse hierarchy is usually simulated using some form of Press-Schechter theory and the outcomes of individual collapses are determined from heuristic arguments. Because they require relatively little computation, the semi-analytical models make it easy to test large ranges of parameters (e.g. to test models of star formation). In principle, a galaxy formation model should account for everything we see, but in practice, the models try to account for the overall properties of galaxies. Semi-analytical models have had some success, but in many respects their treatment of collapses is based weakly on physical models. Furthermore, different collapse models can account for many of the same data. It is likely to be some time before we can say that we have a definitive model for the details of galaxy formation.

4 Dwarf (Cold) Galaxy Formation

Since dwarf galaxies are an important topic, the rest of my remarks are addressed to their formation. Although it needs modification to allow for dark matter, the argument of Rees & Ostriker (1977) is still essentially valid. It shows that shock heating during the collapse of small galaxies is transient at best. Thus, the collapse of a small protogalaxy results in cold gas within a dark matter halo. I will use the term 'dwarf' loosely to refer to any system resulting from such a cold collapse.

The collapse of the dark matter associated with the dwarf is dissipation-less and leads to a halo of the form proposed by Navarro, Frenk & White (1997). This may be modified by gas processes, as discussed below. Its only significant function for the current purpose is that it provides a potential that confines the cold gas, squeezing it to high density.

Consider the early collapse of a halo with a total mass of $10^{10} M_{10} M_{\odot}$. Using standard arguments from semi-analytical models for galaxy formation (e.g. Nulsen, Barcons & Fabian 1998), such a halo would have a velocity dispersion of about

$$\sigma \simeq 46 M_{10}^{\frac{1}{3}} t_9^{-\frac{1}{3}} \text{ km s}^{-1}, \quad (2)$$

if it collapsed at $10^9 t_9$ yr. If the gas fraction in the collapse is 0.2 (Section 2), then the protogalactic halo will contain about $2 \times 10^9 M_{\odot}$ of cold gas. While star formation in dwarfs, even those with substantial gas fractions, can be slow now, it is hard

to see how such a large mass of gas in such a deep potential well could be prevented from undergoing a massive star burst.

Since the last remark contradicts the evidence from local conditions, it worth considering more closely. For example, it is well known that according to the Jeans criterion, giant molecular clouds are unstable, yet they do not undergo rapid star formation. Apart from the fact that a very large mass of gas is trapped in a small region, why should we expect something different when galaxies are forming? The galaxies we observe at the present day are old, so we should expect them to have reached a relatively steady state, in which changes take place on a timescale defined by their age (since the last major disturbance). The only exceptions are systems where the time required to reach a steady state is longer than their ages, as may be the case for dwarf spheroidal galaxies (e.g. Gallagher & Wyse 1994). Thus, while it may be slightly surprising that giant molecular clouds do not collapse more quickly, it is no surprise that the overall star formation rate today is modest. There are numerous feedback processes that can help to maintain such conditions (e.g. Silk 1997). However, there is also ample evidence that disturbing a galaxy, e.g. by gravitational interaction with another galaxy, will significantly alter the star formation rate (e.g. Kennicutt et al. 1998). The conditions during collapse of a protogalaxy are about as far as it is possible to get from the near equilibrium that we see today. Some time is required (perhaps of the order of the time until the first supernovae, a few million years) for feedback processes to begin to limit the first wave of collapse that results in star formation. Thus, there is no inconsistency between the expectation of a massive star burst during the collapse of a protogalaxy, but only a relative trickle of star formation today.

Nulsen & Fabian (1997) argued that an initial burst of star formation proceeds on the dynamical timescale (or faster). Feedback from supernovae can limit the burst only if it is at least as fast, requiring the speed of the expanding blast (driven by overlapping supernova explosions) to be at least comparable to the escape speed from the potential well. As a result, supernovae can only limit the star burst by ejecting a substantial part of the remaining gas. Gas may be completely ejected from the halo, or it may re-accrete some time later. The ejected gas can account for the bulk of the damped Ly α absorption line systems in quasar spectra (Nulsen et al. 1998). Observations of the Hubble Deep Field show clear evidence for the massive star bursts required by this model (Sawicki & Yee 1998; Lilly et al. 1998). Note that such massive star bursts are only expected in those collapses where star formation is not impeded for some other reason (e.g. because the gas is stable according to the Toomre

criterion). This is why they tend to occur in early collapses of relatively massive systems.

5 Disk Formation in an NFW Halo

We want to relate the scale-length of an exponential disk to the properties of the halo in which it forms. The density distribution for the NFW potential (Navarro, Frenk & White 1997) is

$$\rho_{\text{halo}}(r) = \frac{\rho_0}{(1 + r/a)^2 r/a}, \quad (3)$$

where the halo scale-length a and the normalising density ρ_0 are constants. This density distribution is truncated at r_{200} and the NFW concentration parameter is defined as $c = r_{200}/a$.

The angular momentum of the halo is conveniently specified in terms of the spin parameter

$$\lambda_h = \frac{J_h |E|^{1/2}}{GM_h^{5/2}}, \quad (4)$$

where J_h is the angular momentum of the halo, E its total energy and M_h its mass. Tidal torques typically make the spin parameter about 0.05 (Steinmetz & Bartelmann 1995). We define a spin parameter for the disk λ_d to make the specific angular momentum of the disk the same as that of the halo when $\lambda_d = \lambda_h$, i.e. so that the specific angular momentum of the disk is

$$\frac{J_d}{M_d} = \frac{\lambda_d GM_h^{3/2}}{|E|^{1/2}}. \quad (5)$$

Any ejection is likely to selectively remove gas with higher than average specific angular momentum, so that we should expect the spin parameter of a disk that formed from gas to be comparable to or smaller than the spin parameter of the halo.

The angular momentum of an exponential disk is determined by its scale-length b and the mass distribution of the halo in which it resides, so that specifying λ_d determines the exponential scale-length in terms of halo properties. For $b \ll a$ (i.e. $\lambda_d \ll 1$), we have $b \propto \lambda_d^{2/3} a$. Figure 1 shows the ratio of disk scale-length to halo scale-length as a function of halo concentration parameter, for some reasonable values of the disk spin parameter. Physically reasonable values of these parameters give $b/a \simeq 0.1$.

The small value for the disk scale-length raises the issue of whether the mass of the disk dominates the mass of the halo in any part of the disk. Figure 2 shows the ratio of disk mass to halo mass within radius r as a function of the scaled radius r/b , for some typical values of b/a and a gas fraction of 20% (the halo concentration parameter is taken as 3.5 and the results are not sensitive to it). About 60% of the mass of the disk is contained within

$r = 2b$ and about 90% within $r = 4b$. We see from the figure that the disk dominates the halo, except in the outermost part of the disk. Lowering the gas fraction and/or increasing the angular momentum of the disk can modify the degree of dominance, but for reasonable values of these parameters, we should expect at least the inner half of the disk (by mass) to be self-gravitating.

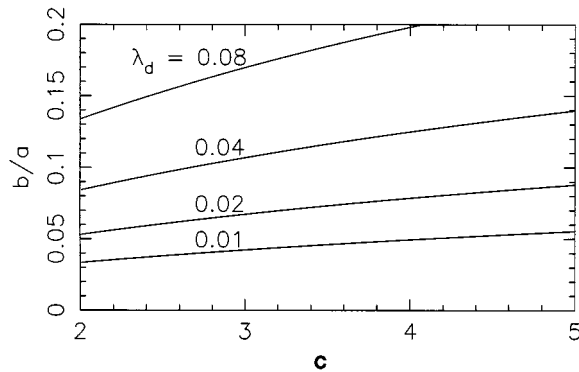


Figure 1—Disk scale-length versus halo concentration. The ratio of disk scale-length b to halo scale-length a is plotted as a function of halo concentration parameter c for several values of the disk spin parameter λ_d , shown above each curve.

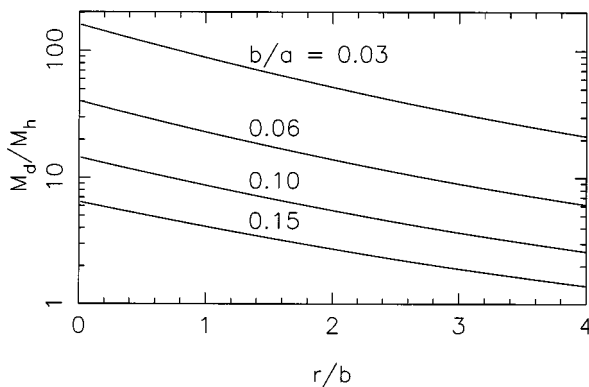


Figure 2—Disk mass over halo mass versus radius. This shows the ratio of the mass within r in an exponential disk to the mass within r in an NFW halo as a function of the scaled radius r/b , where b is the disk scale-length. Curves are shown for several values of b/a , where a is the scale-length for the NFW halo (equation 3). A gas fraction of 0.2 and a halo concentration parameter of 3.5 have been used.

Ejection and other processes can mean that the mass ending up in the disk is substantially smaller than the mass of gas that first collapses. Immediately after collapse the gas must be supported either by pressure or by rotation. In a cold collapse the support will usually be by rotation, so that the gas quickly collects in the region where the disk is to form (the exception being for small halos, where heating by photoionisation can keep the gas pressure supported). Based on the argument above, this gas will be largely self-gravitating. Star formation,

with the associated feedback, and other processes complicate this picture. Nevertheless, we can expect the gas and stars that form from it to be largely self-gravitating. This means that the gas cannot simply be regarded as collapsing in a dark matter halo for most cold (dwarf) collapses.

Self-gravitation can give the disk a substantially higher rotation velocity than the halo in which it forms. A rough estimate gives the boost to the rotation speed of the disk as about $v_{\text{rot,d}}/v_{\text{rot,h}} \simeq f_{\text{gas}}/\lambda_d$, for a gas fraction of f_{gas} . If the disk rotation speed is substantially higher than that of the halo, then the decline in the rotation curve at the edge of the disk may give the impression that there is no dark halo. Self-gravity will certainly modify the shape of the rotation curve near the centre of a galaxy (e.g. Gelato & Sommer-Larsen 1999).

The main issue here is that, if the gaseous component is largely self-gravitating, then gas processes must play a major role in determining the structure of the visible part of a galaxy. To date, treatments of gas processes, especially in cold systems, have been crude. Until we have a much better understanding of gas processes in dwarf galaxies we will not be able to construct a good theory of galaxy formation.

6 Conclusion

While the Universe is dominated by dark matter and this determines the large-scale structure of the dark halos that surround galaxies, the gas fraction is sufficiently large to play an important role in galaxy formation. In particular, the gas that is the progenitor of the disk is likely to be self-gravitating in many protogalaxies. This means that gas processes played a major role in determining the structure of the visible parts of those galaxies. Until we have a well developed theory for the behaviour of the gas, especially for star formation, we will not have a good understanding of the process of galaxy formation.

Acknowledgments

I thank the Astrophysics Group at the University of Melbourne for their hospitality during this meeting and the referee for constructive comments.

References

- Baugh, C. M., Cole, S., & Frenk, C. S. 1996, MNRAS, 283, 1361
- Blanchard, A., & Bartlett, J. G. 1998, A&A, 332, L49
- Burles, S., & Tytler, D. 1998, ApJ, 499, 699
- Carlberg, R. G., Morris, S. L., Yee, H. K. C., & Ellingson, E. 1997, ApJ, 479, L19
- Carr, B. 1994, ARA&A, 32, 531
- Cavaliere, A., Menci, N., & Tozzi, P. 1998, ApJ, 501, 493
- Donahue, M., Voit, G. M., Gioia, I., Luppino, G., Hughes, J. P., & Stocke, J. T. 1998, ApJ, 502, 550

- Ettori, S., Fabian, A. C., & White, D. A. 1998, *MNRAS*, 300, 837
- Evrard, A. E. 1997, *MNRAS*, 292, 289
- Forman, W., & Jones, C. 1982, *ARA&A*, 20, 547
- Gallagher, J. S., & Wyse, R. F. G. 1994, *PASP*, 106, 1225
- Gelato, S., & Sommer-Larsen, J. 1999, *MNRAS*, 303, 321
- Gross, M. A. K., Somerville, R. S., Primack, J. P., Borgani, S., & Girardi, M. 1998, in *Large Scale Structure: Tracks and Traces*, eds V. Mueller et al. (Singapore: World Scientific), p. 171
- Gunn, K. F., & Thomas, P. A. 1996, *MNRAS*, 281, 1133
- Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, *MNRAS*, 264, 201
- Kennicutt, R. C., Schweizer, F., Barnes, J. E., Friedli, D., Martinet, L., & Pfenniger, D. 1998, *Galaxies: Interactions and Induced Star Formation*, Saas-Fee Advanced Course 26 (Berlin: Springer)
- Lilly, S. J., Eales, S. A., Gear, W. K., Bond, J. R., Dunne, L., Hammer, F., Le Fèvre, O., & Crampton, D. 1998, *astro-ph* 9807261
- Lineweaver, C. H. 1998, *ApJ*, 505, L69
- Maguire, P. J. 1996, PhD Thesis, University of Wollongong
- Merchan, M. E., Abadi, M. G., Lambas, D. G., & Valato, C. 1998, *ApJ*, 497, 32
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
- Nulsen, P. E. J., & Fabian, A. C. 1997, *MNRAS*, 291, 425
- Nulsen, P. E. J., Barcons, X., & Fabian, A. C. 1998, *MNRAS*, 301, 168
- Rees, M. J., & Ostriker, J. P. 1977, *MNRAS*, 179, 541
- Sawicki, M., & Yee, H. K. C. 1998, *AJ*, 115, 1329
- Schweizer, F. 1986, *Science*, 231, 227
- Silk, J. 1997, *ApJ*, 481, 703
- Somerville, R. S., & Primack, J. R. 1998, *astro-ph* 9802268
- Steinmetz, M., & Bartelmann, M. 1995, *MNRAS*, 272, 570
- Thomas, P. A., Colberg, J. M., Couchman, H. M. P., Efsthathiou, G. P., Frenk, C. S., Jenkins, A. R., Nelson, A. H., Hutchings, R. M., Peacock, J. A., Pearce, F. R., White, S. D. M. 1998, *MNRAS*, 296, 1061
- Trimble, V. 1987, *ARA&A*, 25, 425
- Walker, T. P., Steigman, G., Kang, H.-S., Schramm, D. M., & Olive, K. A. 1991, *ApJ*, 376, 51
- Webb, J. K., Carswell, R. F., Lanzetta, K. M., Ferlet, R., Lemoine, M., Vidal-Madjar, A., & Bowen, D. W. 1997, *Nature*, 388, 250
- Weinberg, D. H., Hernquist, L., & Katz, N. 1997, *ApJ*, 477, 8
- White, D. A., Jones, C., & Forman, W. 1997, *MNRAS*, 292, 419
- White, S. D. M., & Frenk, C. S. 1991, *ApJ*, 379, 52
- White, S. D. M., Navarro, J. F., Evrard, A. E., & Frenk, C. S. 1993, *Nature*, 366, 429