Downloaded from rsta.royalsocietypublishing.org

An empirical approach to galaxy formation

Joseph Slik

Phil. Trans. R. Soc. Lond. A 2000 **358**, 2153-2158
doi: 10.1098/rsta.2000.0637

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. A* go to:
http://rsta.royalsocietypublishing.org/subscriptions

This journal is © 2000 The Royal Society

Downloaded from rsta.royalsocietypublishing.org

10.1098/rsta.2000.0637

THE ROYAL SOCIETY

An empirical approach to galaxy formation

By Joseph Silk

Department of Astrophysics, University of Oxford,
Keble Road, Oxford OX1 3RH, UK

I describe an approach to galaxy formation that commences with star-formation constraints for nearby galaxies and attempts to reconstruct their past in a statistical sense, by predicting the mean properties of the high-redshift Universe.

Keywords: galaxy formation; cosmology; dark matter

1. Introduction

The cold-dark-matter (CDM) description of the Universe is so successful as to practically constitute a paradigm for large-scale structure. The correlations of galaxies, the properties of galaxy clusters, the dark halos of galaxies, and even the structure of the intergalactic medium have been successfully reproduced in a universe with $\Omega_{\mathrm{CDM}} \approx 0.3$.

Where the theory has encountered a more difficult passage, however, has been in the area of galaxy formation. The CDM infrastructure has not accounted for the sizes of galactic discs, nor for the dark-matter conspiracy that restricts the mass of dark matter to be roughly comparable with that of luminous matter within about three disc-scale lengths, nor for the luminosity function and rotation curves of dwarf galaxies. Introduction of simple feedback prescriptions for star formation has helped account for the luminosity function of galaxies, but leaves the other problems unresolved. It is commonly argued that stellar feedback may account for disc-galaxy sizes, but no detailed model has been developed. It is therefore difficult to attach much credibility to the quantitative implications of the *ab initio* theory of galaxy formation for the high-redshift Universe. More parameters can no doubt be introduced to meet any particular observational challenge, but with the loss of much of the model's predictive power.

It is useful to develop an alternative approach to making predictions of galaxy properties at high redshift that is more empirical and embedded in the phenomenology of nearby galaxies. For disc galaxies, current star-formation rates are measured, and there is an empirical dependence on surface density of cold gas (HI and H_2). In the case of elliptical galaxies, the fossilized relic properties of past star formation include colours, metallicity line indices, and correlations such as the colour–magnitude relation and the fundamental plane. I shall argue that one can run these star-formation histories backwards in time and develop phenomenological predictions of early phases of galaxy evolution, enabling one, for example, to simulate images that will be taken with the Next Generation Space Telescope (NGST). Bulge formation provides more uncertainty, especially with regard to timing, but this can be modelled empirically. One can then examine such issues as the contribution of different galaxy types to

© 2000 The Royal Society

Downloaded from rsta.royalsocietypublishing.org

the star-formation rate history of the Universe and to the diffuse background light at various wavelengths, as well as to the chemical evolution of the Universe.

2. Empirical disc-galaxy formation

The advantages of an empirical approach towards forming galaxies are that theoretical motivation can be combined with phenomenology to provide predictions of forming galaxies at extremely high resolution, limited only by nearby observational constraints, and at extremely high redshift. The principal weakness is that a model for merging must be incorporated, but this is no different from the standard approach to merging and hierarchical structure formation.

The foundation of the empirical model commences with observations of nearby disc galaxies. The star-formation rate is proportional to gas surface density, and the empirical fit can actually be improved if a dependence on galactic rotation rate is explicitly incorporated. The data must directly yield the observed star-formation rate per unit area, $\dot{\mu}_* \propto \mu_{\text{gas}}^{1.4}$, which is consistent with the naive theoretical expectation $\dot{\mu}_* \propto \mu_{\text{gas}}\Omega$ (Kennicutt 1998). Theoretical models that suggest this approximate proportionality include spiral-density-wave-induced molecular-cloud collisions and the gravitational instabilities of a cold gaseous disc to massive molecular-cloud formation. I will describe below an explicit derivation that incorporates feedback via a multiphase interstellar medium. In all of these cases, molecular clouds are assumed to be the precursor phase to the bulk of galactic star formation.

Detailed fitting to the Milky Way, where one can incorporate chemical evolution as well as star-formation relics and radial profiles of gas and stars, yields a generalization of the simplest star-formation law:

$$\dot{\mu}_* = \epsilon \Omega \mu_{\text{gas}} f(Q_{\text{gas}}) + \text{infall}.$$

This can be taken as an empirical model. Metal-poor gas infall is ubiquitous in cosmological modelling of early galaxy evolution, and is explicitly required in order to reproduce galactic chemical evolution even when other parameters, such as time-variations of yields or of the initial mass function, are introduced. The star-formation efficiency ϵ represents the fraction of gas converted into stars per dynamical timescale.

The key to understanding star-formation efficiency lies in deriving the velocity dispersion of the interstellar gas. For the smaller clouds, supernova remnants dump energy and momentum into the interstellar gas that is balanced by energy dissipated in cloud–cloud collisions. The most massive clouds are accelerated via mutual tidal interactions that are driven by disc gravitational instabilities. There must be a steady state in which the clouds acquire a net momentum, since there are continual dynamical interactions between clouds of all masses, for example via processes of accretion, coagulation and fragmentation. One infers that

$$\epsilon = \left(\frac{v_{\text{rot}}}{v_{\text{SN}}}\right)\left(\frac{l_{\text{cl}}}{l_{\text{t}}}\right)^2 \frac{1}{2\pi Q_{\text{gas}}} \approx 0.03.$$

Here, v_{rot} is the disc maximum rotational velocity, v_{SN} is the specific momentum injected by supernovae, l_{cl} is the typical cloud radius, l_t is the cloud tidal radius, and Q_{gas} is the Toomre parameter for the gas component of the disc. Self-regulation of

Downloaded from rsta.royalsocietypublishing.org

star formation requires that $Q_{\text{gas}} \sim 1$, and is responsible for the low efficiency of star formation and for the longevity of disc-star formation over several Gyr. The reason ϵ is so small and star formation is inefficient is that the self-regulated disc cannot be very unstable, e.g. with $Q_{\text{gas}} \ll 1$.

The disc gas velocity dispersion is given by

$$\sigma_{\text{g}} = v_{\text{rot}}(\epsilon\eta\mu_{\text{gas}}/\mu_*)^{1/3} \approx 10\text{–}20\ \text{km s}^{-1},$$

where η is a second, local self-regulation parameter, of order unity and to be defined below, that is associated with cloud–cloud interactions and star formation. Hence, the physics of disc-star formation can be reduced to specifying two parameters, ϵ and η, which are self-regulated to be of order unity by the physics of global star formation.

The feedback processes involve cloud aggregation and star formation. As previously argued, the molecular clouds are accelerated tidally and by supernova momentum, lose energy via collisions, and accumulate mass. Once sufficiently massive, as determined by a magnetically limited Jeans criterion, the clouds collapse and form stars. Cloud growth is regulated by the cloud–cloud interactions. These can be naively modelled in terms of a macroscopic viscosity. The radial flow of gas is limited by the effective viscosity, which, therefore, controls disc and bulge formation. One defines the viscosity coefficient by

$$\nu_{\text{eff}} = \tfrac{1}{3}\sigma_{\text{g}}\lambda,$$

where λ is the cloud effective mean free path. If one makes the ansatz that the viscous drag time-scale is equal to the star-formation time-scale, as plausibly expected in the viscosity-controlled model of self-regulation, then an exponential density profile results for the forming disc (Lin & Pringle 1987). I define the feedback parameter η as the ratio of these two time-scales: $\eta \equiv t_{\text{sf}}/t_\nu$.

Let us now ask the question as to how feedback operates to maintain Q_{gas} to be of order unity. The answer must lie in the detailed structure of the interstellar gas, and, in particular, in the equipartition of energy and momentum between the hot phase, directly generated by star formation, and the cold phase, within which the stars form. A simple model of porosity for the multiphase disc gas allows one to explicitly simulate the feedback. The porosity can be defined as

$$P = (\text{4-volume of supernova remnant})(\text{supernova rate per unit volume}),$$

and is related to the hot-gas disc-volume fraction f by

$$P = -\log(1-f).$$

Simple modelling of spherically symmetric supernova remnants shows that P depends primarily on the turbulent pressure $p_{\text{g}} = \rho_{\text{g}}\sigma_{\text{g}}^2$, and, more specifically, that $P \propto p_{\text{g}}^{-1.36}\rho_{\text{g}}^{-0.11}\dot{\rho}_*$.

In the case where feedback is important $P \sim f \sim 1$, and the star-formation rate is effectively determined. One can effectively derive the previously assumed empirical star-formation law to be $\dot{\rho}_* \propto \rho_{\text{g}}^{1.5}\sigma_{\text{g}}^{2.7}$. Note that the constant of proportionality, in effect the star-formation efficiency, is determined by the feedback ansatz.

Downloaded from rsta.royalsocietypublishing.org

3. Application of the empirical approach

Armed with a prescription for disc formation, one can account for about half of the star formation in the Universe. Approximately 50% of star formation occurs in forming spheroids. The key difference of course is that disc-star formation is generally ongoing, with a characteristic time-scale of several Gyr, whereas star formation in spheroids mostly occurred at high redshift and over a characteristic duration of the order of 10^8 years. Spheroid formation is readily modelled empirically, even though one has no fundamental understanding of the nature of starbursts. Much of the star formation in forming spheroids is probably obscured: we infer this because the diffuse background light provides a census of all cosmic star formation, and about half of this, in terms of energy density, is produced at far-infrared wavelengths. We infer that many starbursts are dust shrouded, the dust absorbing and re-emitting most of the starlight in the far infrared. This interpretation has generally been confirmed by direct observations at submillimetre wavelengths, where galaxies are found at high redshift that are almost entirely obscured at visible wavelengths.

This empirical knowledge, combined with simple rules for star formation in discs as observed and in spheroids as inferred, enables us to construct an empirical model of the early Universe. One can test simple predictions of galaxy evolution. For example, the disc-star-formation model predicts that discs grow inside out. This is in accord with observations, which clearly demonstrate a preponderance of small objects at high redshift. A more subtle issue is the timing of spheroid formation. For example, do bulges form before discs, simultaneously with discs or even after discs? Theoretical motivation exists for each of these possibilities. The CDM model generically predicts that spheroids form first, and that discs form by late infall. However, uncertainty in the efficiency of early star formation, assumed to be highly efficient in spheroid formation, means that this expectation has been subject to revision. Moreover, dynamical models of thin discs suggest that secular instabilities can lead to bar formation. The bars are unstable to spheroid formation. Since about half of all observed discs show central bars, secular instabilities cannot be ignored.

These alternative scenarios for bulge formation can be compared by using the high-redshift Universe as a probe. Simulations of deep galaxy fields should be expected to show bluer bulges if bulges form more recently. Unfortunately, the data are too sparse to be able to make any strong statements (Bouwens & Silk 2000). Moreover, a trickle of gas infall and late star formation can modify the predicted colours.

Deep images with the NGST will combine morphological and colour information and eventually provide a probe of the bulge-formation epoch. Perhaps even more challenging is the formation of massive bulges and of elliptical galaxies. Here, mergers played an important role. There are examples of nearby ongoing mergers that suggest that such events may be prototypes of forming ellipticals, as indicated by the high rate and efficiency of star formation and the emergent light profile. However, one can easily imagine a situation in which the bulk of stars formed in a series of small mergers, as is indeed characteristic of how the dark halos form in hierarchical models of CDM, or a monolithic scenario, in which most of the star formation occurs in a final catastrophic major merger. If the latter hypothesis is correct, elliptical formation occurs either at very high redshift ($z > 5$) or is shrouded by dust. The dust is likely to be removed by outflows driven by the starbursts. Indeed, chemi-

Downloaded from rsta.royalsocietypublishing.org

cal evolution, both of the stellar population in ellipticals, as inferred from Mg/Fe enhancements, and of the intracluster gas, as inferred from the Fe abundance, may require substantial outflows of processed gas to have occurred in the past. Another possible manifestation of protospherical outflows may be the necessity to remove the early enrichment, primarily in C, associated with the intermediate-mass stars whose white-dwarf remnants may account for up to 25% of the mass of the inner (*ca.* 25 kpc galactocentric radius) Milky Way halo according to gravitational microlensing measurements and halo and white-dwarf searches.

Another example of the empirical approach to galaxy formation concerns the presence of galaxies at very high redshift. Several galaxies have been reported at redshifts in excess of 5. There are substantial populations of galaxies detected via Lyman-break searches at $z = 3$ and 4. Indeed, the number densities are consistent with the local Schechter function, subject only to passive density evolution. This result, especially the existence of even a few ultraluminous galaxies at $z > 5$, presents a challenge to CDM theories of structure formation. The problem of early formation is greatly eased in a low-Ω_{m} universe, which of course favours formation at $z \gtrsim \Omega_{\mathrm{m}}^{-1} - 1$. An interesting consequence is that the availability of enough comoving space at, say, $z > 5$ is the critical ingredient that overwhelms the uncertainties in galaxy-formation modelling. Hence, an empirical, and simple, approach to galaxy formation more than suffices to help discriminate between cosmological alternatives. For example, it is not possible to seriously maintain a model in which the high-redshift galaxies are dwarf mass systems whose star-formation rate is briefly enhanced by collisions and mergers within an overall massive halo. While this model can account for the observed clustering of the Lyman-break galaxies, it is difficult to reconcile with the combination of high spatial frequency (e.g. in excess of two per square arcminute at $z > 5$) and high star-formation rate. Indeed, if observations continue to find a comparable surface density of luminous galaxies beyond $z = 6$, only a model with hyperbolic space may be able to provide the required frequency of star-forming galaxies (Robinson & Silk 2000).

Two puzzling properties of discs can be tentatively explained via empirical modelling (Silk 2000). Viscous coupling to star formation, an ansatz in the empirical model (i.e. $\eta \approx 1$) that is plausible if not rigorous, has the effect of forming the disc stars before most of the baryonic angular momentum has been lost as a consequence of clumpy gas inflow. This means that disc sizes can be understood. Empirical modelling also yields a star-formation rate in discs that is globally proportional to v_{rot}^2. Combining this result with the usual hierarchical model estimate that $M_{\mathrm{d}} \propto v_{\mathrm{rot}}^3$, one infers a Tully–Fisher relation that steepens systematically between blue and near-infrared wavelengths, as observed. More impressively, perhaps, one can also obtain the observed Tully–Fisher relation normalization, hitherto unaccounted for, in terms of the predicted star-formation efficiency.

A final application of empirical modelling uses the colour–magnitude relation and fundamental plane for ellipticals in clusters at various redshifts. One can lift the age–metallicity degeneracy in colour for ellipticals by projecting back in redshift. Incorporation of simple chemical-evolution modelling shows that the observed dispersion in the colour–magnitude relation requires some combination of gas outflows and metallicity-dependent yields (Ferreras & Silk 2000). Since the modelling controls the stellar mass, one can then predict the evolution of the fundamental plane, with redshift, if dynamical evolution is neglected.

Downloaded from rsta.royalsocietypublishing.org

I thank Rychard Bouwens and Ignacio Ferreras for many discussions about galaxy formation during the course of preparing this review.

References

Bouwens, R. & Silk, J. 2000 *Astrophys. J.* (In the press.)
Ferreras, I. & Silk, J. 2000 *Astrophys. J.* **532**, 193.
Kennicutt, R. 1998 *Astrophys. J.* **498**, 541.
Lin, D. & Pringle, J. 1987 *Astrophys. J.* **320**, 87.
Robinson, J. & Silk, J. 2000 *Astrophys. J.* (In the press.)
Silk, J. 2000 *Mon. Not. R. Astr. Soc.* (Submitted.)