Notes for G9001: Galaxy Formation

Greg Bryan

September 28, 2004

Part I Observational overview

These notes will first present a very broad-brush review of the observations relating to galaxy formation that we would hope to explain with a complete theory, and then review theoretical approaches and methods aiming at explaining those observations.

1 Galaxy Structure and Morphology

One of the most obvious facts about galaxies is that they come in a variety of sizes and luminosities. Perhaps more surprisingly is that, for a given size and luminosity, there are usually a number of different types. Remarkably, the inherent reason for this variety is still not well understood. It is one of the key requirements for a successful theory: that we can reproduce the rich diversity of galaxies.

1.1 Galaxy Classification [Binney & Merrifield 4.3]

There are many ways of classifying galaxies but by far the most well-known is the Hubble system, which has originally suggested by Hubble in his 1936 book, The Realm of the Nebulae. Galaxy classification is a topic all of it’s own (for example see Sidney van den Bergh’s Galaxy Morphology and Classification), so here we just remind readers of the basic classes and the difficulties encountered in classification by any system.

In Hubble’s tuning-fork diagram, on the left are the sequence of elliptical galaxies spanning the range from circular to highly elliptical systems. These are denoted $E_n$, where $n = 10(1 - b/a)$ is computed by measuring the axial ratio $b/a$ of the ellipse formed on the sky. Note that these are measured from the projection on the sky, so that the intrinsic three-dimensional distribution may be considerably different (in fact, in general elliptical galaxies are likely to be triaxial, and the axes can change as a function of radius).

The elliptical systems are often further sub-divided into dwarf types (dE), although it is not clear if the dwarfs are simply small elliptical or a class of their own. Dwarf spheroidals (dSph) are a closely related type, the primary difference being that they are even less luminous and more diffuse (they are an interesting targets for dynamical studies because of their very high dark matter content). With the advent of new technology that permits detailed studies, the smallest galaxies have become a topical area of research. For a summary of known dwarf galaxies in the local group, see the article by Mateo et al. in the 1998 Annual Reviews of Astronomy and Astrophysics (Vol 36, p. 435).

After the ellipticals, Hubble’s diagram splits in two, with two parallel sequences differentiated by the presence (top sequence) or absence (bottom) of a stellar bar in the center. The first galaxies are the so-called lenticular or spheroidal galaxies, designed S0 (or SB0 if there is a bar). These systems appear to be strongly oblate like a spiral system but without any clear spiral structure. Like ellipticals, they usually have little dust or gas content, although some systems can contain significant amounts of dust: the strength of the dust is sometimes signified by a subscript ranging from S0$_1$ (no dust) to S0$_3$ (prominent dust lane).

After the spheroidals come the spirals, designated Sa, Sb or Sc depending on three characteristics: the tightness of the spiral arms, the size of the central bulge and the definitions of the arms. Spirals with tight, well-defined arms and a strong central bulge are Sa, while Sc systems have loose, poorly defined spiral arms
and little or no central spheroidal component. Note that these three characteristics almost vary together, but exceptions exist which can make classification difficult. If the spiral has a bar, it is classified with SBa, SBb or SBc. Additions have been suggested to this system, including de Vaucouleurs addition of Sd and Sm spiral systems that include galaxies which are looser or have even more poorly defined arms than the Sc class.

Hubble thought that galaxies evolved along his sequence, going from elliptical ("early") galaxies to spiral ("late") galaxies. While this idea is no longer thought likely, the terminology has stuck. It is not uncommon to find a numerical T stage assigned to the Hubble type for the use in some forms of statistical analysis. This ranges from -5 (E0) to 0 (S0) to 6 (Sc) to 10 (Ir). Entirely outside of this sequence are the irregular galaxies Irr which lack symmetry or well-defined spiral arms, and the peculiar galaxies which may be undergoing a merger or interaction.

As we will see in more detail later, galaxies are not uniformly distributed in space. While isolated, field galaxies do exist, the majority of galaxies are found in groups or clusters. Very small groups, like the local group, consist of only a few large galactic systems. Large groups may have up to 100, while full-fledged clusters often have thousands of large galaxies. Galaxies of different types are not equally represented in different environments. In particular, ellipticals are much more likely to be found in dense cluster cores than in the field, while spirals show the opposite trend. This is most commonly called the morphology-density relation and is clearly telling us something about what sets galaxies morphology, but it is not clear whether spiral galaxies are transformed into ellipticals by the dense environments of clusters, or if such regions preferentially formed elliptical galaxies even before the cluster formed (clusters are thought, from both theoretical and observational evidence, to be younger than galaxies).

1.2 Galaxy Luminosity Function

We define \( \Phi(L) dL \) as the number density of galaxies with luminosities from \( L \) to \( L + dL \). The luminosity (energy per unit time) \( L \) could be in a particular wave-band or it could be integrated over all wavelengths (the bolometric luminosity). In either case, the luminosity function is usually normalized so that,

\[
\int_0^\infty \Phi(L) dL = n_{gal} \tag{1}
\]

where \( n_{gal} \) is the local density of all galaxies (we have not specified it but the luminosity function can and does vary with position). The luminosity function is a difficult thing to measure but it is clear that it drops quickly at large luminosity and has a large extension to low luminosities. A useful parameterization is the known commonly as the Schechter function:

\[
\Phi(L) = \left( \frac{\Phi^*}{L^*} \right) \left( \frac{L}{L^*} \right)^\alpha \exp\left( -\frac{L}{L^*} \right), \tag{2}
\]

which is simply a power-law of slope \( \alpha \) with an exponential cutoff. The cutoff occurs at \( L^* \) and as long as \( \alpha > -2 \) (which is observationally always the case), the total light integrated over all luminosities is finite:

\[
L_{total} = \int_0^{\infty} L \Phi(L) dL = \Phi^* L^* \Gamma(2 + \alpha) \tag{3}
\]

For the typical value of \( \alpha = -1 \), the Gamma function is unity and since \( \Phi^* \) is approximately the number density of galaxies with luminosity of \( L^* \), the total light is clearly contributed mostly by galaxies around \( L^* \). Our own Milky-Way galaxy has approximately this luminosity, and it defines the typical brightness of a large galaxy.

On the other hand, the total number of galaxies is not always convergent and indeed large numbers of low-luminosity galaxies are observed (although there does appear to also be a lower cutoff which is not modelled by the Schechter function). The luminosity function is difficult to determine observationally for a number of reasons. First, the low luminosity end is difficult to determine because faint galaxies can only be seen nearby. More worryingly, the surface brightness of many galaxies falls close to the limit which can easily be observed. Even intrinsically luminosity galaxies can appear dim if they are large and spread across a large region of the sky. Indeed, recent work has uncovered a whole population of low-surface brightness (LSB) galaxies. These LSB galaxies are typically only a few percent of the surface brightness of the background.
sky and so are difficult to see. However, they are clearly of great importance both because they may be as numerous as their high-surface brightness cousins, but also because many appear to be largely dark-matter dominated and so provide good cases to study the distribution of non-luminous matter in galaxies.

2 Elliptical galaxies

We first address elliptical galaxies in more detail. Elliptical galaxies are in some ways simpler because they generally do not have gas and dust, so are largely stellar systems. On the other hand, they do not have the simple orbits that characterize disk systems. We start with radial structure of elliptical galaxies, and then move on to global properties of the systems and their scaling relations.

2.1 Radial brightness profile [Binney & Merrifield 4.3]

While elliptical galaxies are generally not spherical, it is often useful to examine how their brightness varies as a function of distance from the center. This is done either by fitting an elliptical model or by fitting the major and minor axes separately. In either case, it is found that the surface brightness (i.e. energy per unit second per unit area per unit solid angle) measured in magnitude per square arcsec ($\mu$) varies with projected distance $R$ as $R^{1/4}$. If we write surface brightness in linear units (recall that $I \propto 10^{-0.4 \mu}$), then it is usually to express this relationship as:

$$I(R) = I_e \exp(-7.67[(R/R_e)^{1/4} - 1]).$$  \hspace{1cm} (4)

The numerical factors are chosen such that if 1/2 of the total light (for a spherical system) is emitted within $R_e$:

$$\int_0^{R_e} \pi R I(R) dR = L_e/2.$$  \hspace{1cm} (5)

For this reason, $R_e$ is a useful measure of a galaxies size and is called the effective radius. It is also useful to define the mean surface brightness within the effective radius: $L_e = \pi R_e^2 < I >$.  

While most elliptical galaxies follow this distribution, some giant elliptical galaxies show an excess of stars at large radii. These cD ellipticals are generally found at the center of large groups or clusters of galaxies. The excess is commonly interpreted as an overall envelope of stars that more properly belong to the clusters as a whole than to the cD galaxy. This intra-cluster light can contribute significantly (20% in some cases) of the observed starlight from the entire cluster.

2.2 The fundamental plane of elliptical galaxies [Binney & Merrifield 4.3]

Elliptical galaxies are unlikely to be significantly supported by bulk rotation, as is the case for spiral galaxies. However, the stellar velocities still give an indication of the gravitational mass. Usually, it is impossible to measure individual stellar velocites in elliptical galaxies, but the distribution of line-of-sight velocities can still be measured as a function of position across the galaxy. This is done by measuring the spread of stellar emission line caused by the Doppler shifting of all of the stars at that (projected) distance. The velocity dispersion at the center of an elliptical galaxy is denoted by $\sigma_0$ and is usually straightforward to measure.

It was noticed quickly that this quantity was strongly correlated with the galaxies luminosity:

$$L_e \sim \sigma_0^4$$  \hspace{1cm} (6)

This is known as the Faber-Jackson relation. Other relations between these quantities and $< I >$ or $R_e$ were found, but all of them (including the Faber-Jackson relation) can be thought of as projections of a single correlation between three quantities. This relation defines a two-dimensional plane within the three-dimensional space define by (say), $R_e$, $< I >$ and $\sigma_0$. One common parameterization of this is given by:

$$\log R_e = 0.36 \mu_e + 1.4 \log \sigma_0,$$  \hspace{1cm} (7)

where $\mu_e$ is the surface brightness in units of blue magnitudes per arcsec (recall $< I > \propto 10^{-0.4 \mu}$ , $R_e$ in units of kpc and $\sigma_0$ in km/s.