Globular Cluster Systems in Normal Galaxies

0.1 Introduction

All but the smallest dwarf galaxies possess a family of globular clusters. They are among the oldest stellar systems in the Universe. As such these 'astronomical fossils' provide valuable clues about the formation and evolution of their host galaxy. They offer a unique probe of the metallicity and kinematics at large radii in galactic halos, including our own Milky Way galaxy.

For Local Group galaxies, spectroscopic studies of GCs are well advanced. However the bulk of our knowledge about extragalactic GCs comes from imaging. Imaging studies received a boost from wide field-of-view CCDs in the late 1980s and from the *Hubble Space Telescope* (HST) in the 1990s. It is now possible to study the magnitudes, colors and spatial distribution of GC systems at the distance of the Coma cluster and beyond. Imaging can be hundreds of times more efficient than spectroscopy for studying GC systems but it supplies little information about age, independent of metallicity, and nothing about kinematics or abundance anomalies.

Perhaps the most basic property of a GC system is the total number of GCs associated with a galaxy. This number varies from dwarf galaxies with a single GC to over 10,000 in the case of some cD galaxies. The number of GCs associated with a galaxy correlates roughly with the host galaxy luminosity (i.e. $N \sim L^2$). In order facilitate comparison between galaxies, the number of GCs is normalised to $M_V = -15$. This is termed the *specific frequency*,

$$S_N = N 10^{0.4(M_V+15)}$$

The S_N value is affected by environment, with cluster ellipticals having $S_N \sim 5$ and their field counterparts $S_N \sim 3$. Some of the most luminous galaxies, typically cD galaxies, have $S_N \sim 15$, i.e. several times more GCs per unit starlight than other early type galaxies. Therefore high S_N galaxies have either formed or acquired GCs more efficiency than they have field stars. It appears that high S_N galaxies have a high S_N value at all radii throughout the galaxy. Although suggestions to explain the high S_N phenomenon have been proposed, none are, as yet, universally accepted. Spiral galaxies have $S_N \sim 1$, but if we take account of the lower mass-to-light ratio of the young stellar populations in spiral disks then this rises to $S_N \sim 2$.

0.2 Local Group

0.2.1 Magellanic Clouds

Both Magellanic Clouds appear to contain young (< 1 Gyr), intermediate (~ 5 Gyr) and old (~ 15 Gyr) GCs. At a distance of only 50 kpc, this would make the Magellanic Clouds ideal laboratories to study GCs with a range of ages. However there are also some disadvantages of studying the Magellanic Cloud GC systems. First, assuming that GCs have been correctly associated with each Magellanic Cloud, the GC systems are small. The LMC has $\sim$ 13 *bona fide* (i.e. old) GCs and the SMC may only have one (NGC 221). Second, it is difficult to distinguish GCs from open clusters (the ratio of open to globular clusters is about 20:1). The LMC GCs have a similar mean magnitude and metallicity to the Milky Way halo ones. They differ in their shapes (slightly more extended and flattened) and kinematics (both young and old GCs reveal evidence for a bulk rotation about the LMC).

0.2.2 M31

The GC system of our sister galaxy, M31, reveals many similarities to our own, with the advantage of greater numbers (i.e. $N \sim 400$ compared to $N \sim 120$ for the Milky Way). For example, the GC luminosity function shows a similar peak and width. The metallicity distribution, although not as obvious, shows a similar metal–poor ([Fe/H] $\sim$ –1.5) and a metal–rich ([Fe/H] $\sim$ -0.5) dichotomy. Associating these with halo and disk populations via kinematics is complicated by the nearly face–on nature of M31. Nevertheless there is some evidence that the metal–rich GCs form a disk–like rotating system. As with the Milky Way and other galaxies, there is a deficit of GCs in the very inner galactic regions. An interesting difference in the M31 GCs is the presence of abundance anomalies. In particular, the CN and Ca H+K lines have enhanced line strengths relative to Milky Way GCs. It has been suggested that this difference may have a primordial origin.

0.2.3 M33 and other Local Group Galaxies

The late type spiral M33 contains $\sim$ 25 old GCs, and perhaps as many again which are young or intermediate aged. The magnitude distribution, mean metallicity and kinematics of the old GCs appear similar to the Milky Way halo system. Most other Local Group galaxies are dwarf spheriodals, many of which possess a handful of GCs. These GCs tend to be metal–poor and show similarities in their color–magnitude diagrams to the young halo GCs in our own Galaxy. Direct evidence for accretion of such GCs comes from the Sgr dwarf (and its 3–4 GCs) which is in the process of merging with our Galaxy. Most dwarf galaxies have relatively normal S_N values but the Fornax dwarf has $S_N \sim 20$ indicating that GC formation can be very efficient in some small galaxies.

0.3 Luminosity Function

The GC luminosity function (GCLF) is simply the number of GCs per luminosity interval. If the number is plotted against magnitude then it is well approximated by a Gaussian distribution, i.e

$$N = Aexp[-(M - M^0)^2/2\sigma^2]$$

where σ is the width of the Gaussian, M^0 is the turnover or peak magnitude and A is the number of GCs at the peak. An example of a GCLF for a giant elliptical is shown in Figure 1. What is remarkable about the GCLF is that the peak is nearly universal, i.e. all well–studied systems have a peak at $M_V^0 = -7.4 \pm 0.2$. For a V band mass-to-light ratio of 2 this corresponds to a mean mass of 1.6 $\times 10^5$ $M_\odot$. The small deviations from a universal M_V^0 value do exist and various causes have been suggested, namely luminosity, Hubble type, color (metallicity) or environmental effects. This issue has yet to be settled. However the mass function of GCs seems to be even more robust, and closely resembles that of Giant Molecular Clouds. Prompted by this observation, it has been suggested that GCs form in the cores of large GMCs.

If GCs do indeed have a nearly universal mean absolute magnitude, then this can be used as a standard candle for distance determination. This has many advantages over other distance determination methods: GCs are available in large numbers around all large galaxies, they do not require repeat measurements (e.g. the study of variable stars), and the method can work for a variety of Hubble types. Its main disadvantage is perhaps a lack of a well–established theory to explain the preferred mass of GCs. The *Hubble Space Telescope* has made a significant impact in this area with its ability to obtain high precision photometry several magnitudes fainter than the peak. The GCLF method now shows excellent agreement with surface brightness fluctuations and Cepheid variable determined distances, giving values of the Hubble constant around 75 km s^{-1} Mpc^{-1}.

0.4 Metallicity Distribution

In the last few years it has become clear that several, perhaps all, massive early type galaxies have two distinct populations of GCs. These subpopulations are apparent in high quality photometric observations as a bimodal color distribution. It is generally assumed that the color separation in the two peaks largely reflects differences in metallicity, and to a lesser extent age. The typical separation between the relatively metal–poor and metal–rich peaks is $\sim$1 dex in [Fe/H]. Breaking the age/metallicity degeneracy for GCs with spectroscopy and near–IR photometry is an active area of research. For the two cD galaxies studied to date (i.e. M87 and NGC 1399), their GC (metal–poor and metal–rich) populations appear to be quite old ($\sim 10 - 15$ Gyrs).

The relative numbers in each population are typically similar, although there is some evidence that the ratio of metal–poor to metal–rich GCs scales with S_N. In other words, high S_N galaxies have relatively more metal–poor GCs. An example of a bimodal metallicity distribution, for an elliptical with a low S_N value, is shown in Figure 2.

The presence of two GC populations appears to explain the long–standing observation that the mean metallicity of the GC system is $\sim$ 0.5 dex lower in [Fe/H] than the stars in massive ellipticals. This simply arises because the metal–rich GCs have a similar metallicity to the host galaxy stars while the metal–poor ones are $\sim$ 1 dex lower in [Fe/H]. Radial metallicity gradients for the entire GC system are due to the changing relative mix of the two populations with galactocentric radius (i.e. the metal–poor to metal–rich ratio increases with radius). Thus there may be no true radial metallicity gradient, suggesting a dissipationless formation process.

The existence of multimodal color distributions indicates that GCs formed in distinct star formation episodes from different metallicity gas, perhaps at different times. A simple monolithic collapse (expected to result in an unimodal color distribution) appears to be ruled out by such data. Interestingly, unimodal distributions may be present in low luminosity ellipticals although the smaller number of GCs present makes any bimodality more difficult to detect.

0.4.1 Metallicity–Luminosity Relation

The correlation between the mean metallicity of the GC system and the luminosity of the host galaxy has been subject to much debate. As the number of systems studied as increased, it now seems clear that some relationship exists from dwarf to cD galaxies. The type of relation, contribution of observational errors and second parameter effects are the subject of ongoing debate. For a linear fit, the slope of the GC system metallicity – galaxy luminosity relation is $Z \propto L^{0.4}$. This is similar to the slope of the galaxy metallicity – galaxy luminosity relation. Taken together these indicate that the GC system has had a similar chemical enrichment history as its host galaxy.

If we subdivide the GC system into the mean metallicity of the metal–poor and metal–rich peaks, then we find that it is the metal–rich peak which correlates with galaxy luminosity. This provides further evidence that the metal–rich GCs are closely coupled to the formation of the galaxy itself. The metal–poor peak shows little or no correlation.

0.5 Spatial Distribution

0.5.1 Radial Distribution

When plotted in the log, the radial surface density distribution for a GC system typically shows a flat 'core' region (i.e approaching zero slope) in the central few kpc and a roughly power–law like decline to larger galactocentric radii. The GC system surface density is often more extended than that of the galaxy starlight. Low S_N galaxies cover a range of power–law slopes but high S_N galaxies only have relatively flat slopes (i.e. an extended distribution). For cD galaxies there is a change in slope at the radius of the cD envelope indicating that the outer GCs and the stellar envelope have a similar origin.

0.5.2 Azimuthal Distribution

Less is known about the azimuthal distribution of GCs than their radial profile. However it appears that GCs generally track the position angle and ellipticity variations of the underlying galaxy starlight. In the future, a number of GC systems may have been studied in enough detail to make comparisons with the spatial properties of X–ray emitting gas. This will provide a unique probe of the dark matter distribution at large galactocentric radii.

0.6 Kinematic Properties

The kinematics of large samples of GCs have been determined for only a few galaxies to date. Although the combination of multiobject spectrographs on large telescopes is rapidly improving this situation. Studying the kinematics of the sub–populations promises to yield important clues as to their origin and ultimately place new constraints on the formation and dynamical evolution of their host galaxy.

An additional use of GC kinematic studies is their application to mass estimates of galaxies. They act like test particles in the gravitational potential and have the advantage of extending well beyond the optical limit of the galaxy. Masses are typically determined using either the projected–mass method or the virial theorem. For a system of N GCs the projected–mass is

$$M_{PM} = \frac{f_p}{G(N-\alpha)}(\Sigma_i^N V_i^2 r_i)$$

where f_p is a numerical factor which describes the type of GC orbits, α is a small correction factor, V_i is the velocity difference between the ith GC and the system mean velocity and r_i is the radial separation from the system centroid position. For small samples of GCs, the mass is likely to be overestimated. The mass from the virial theorem is

$$M_{VT} = \frac{3\pi N}{2G}\frac{\Sigma_i^N V_i^2}{\Sigma_{i<j} 1/r_{ij}}$$

where r_{ij} is the radial separation between the ith and jth GC.

Only a few galaxies have mass estimates from GC motions, although this is likely to be an active area in the coming years. Such estimates can be compared to those from X–ray satellites for massive ellipticals. However for small ellipticals, without extended hot gas halos, GC–based estimates may be the only feasible method for estimating mass at large galactocentric radii.

0.7 Formation and Destruction

The origin of GC systems is not currently known. Traditionally models fell into three main categories: primary (formation before the galaxy), secondary (formation of GCs and galaxy contemporaneously) and tertiary (formation after galaxies). The current formation ideas tend to be more hybrid in nature, motivated by the recent discoveries of two GC populations in massive ellipticals. Ideas include: creation of new GCs in the gaseous merger that forms the elliptical, *in situ* formation plus accretion from smaller galaxies, *in situ* formation in two distinct star formation phases. The reality may be a combination of these basic processes (i.e. *in situ* formation, tidal accretion and the creation of new GCs), operating at varying degrees in different systems.

Once formed, a GC can be destroyed by a variety of internal and external mechanisms. In the case of protoglobular clusters, they will disrupt themselves if supernovae inject more energy than is binding the cluster together. Low mass protoglobular clusters may evaporate or be tidally shocked. Either way low mas clusters must be destroyed if their observed power–law luminosity function is to eventually resemble the log–normal distribution of old GCs. Other long–term dynamical destruction effects include bulge and disk shocking. Disk shocking may destroy a GC as it passes through the increased density of a galactic disk. Bulge shocking is strongest in galaxies with high central densities, i.e. low luminosity ellipticals. Only recently have predictions for GC destruction rates in different types of galaxies become available.

0.8 Bibliography

The literature on GCs in external galaxies is growing rapidly. Below is a brief, and incomplete, selection.

Reviews of GC systems can be found in
Harris, W. E., *Ann. Rev. Astron. & Astrophys.* **29**, 543, (1991)
Ashman, K. M., & Zepf, S. E. *Globular Cluster Systems*, (Cambridge University press), (1998)

A review of GC luminosity functions as distance indicators is given by
Whitmore, B. C., in Livio, M., Donahue, M., & Panagia, N. eds *The Extragalactic Distance Scale*, (Cambridge University press), (1996)

The first photometric studies of GCs beyond the Local Group include
Racine, R., *Pub. Astron. Soc. Pac.* **80**, 326 (1968)
Hanes, D. A., *Mon. Not. Roy. astr. Soc.* **188**, 901 (1979)
Strom, S. E., *et al.* *Astrophys. J.* **245**, 416 (1981)

Photometric studies revealing two GC subpopulations include
Harris, W. E., *et al.* *Astron. J.* **104**, 613 (1992) for NGC 5128
Lee, M. G., & Geisler, D. *Astron. J.* **106**, 493 (1993) for M87
Secker, J., *et al.* *Astron. J.* **109**, 1019 (1995) for NGC 3311
Zepf, S. E., Ashman, K. M. *Astrophys. J.* **443**, 570 (1995) for NGC 3923
Geisler, D., Lee, M. G., & Kim, E. *Astron. J.* **111**, 1529 (1996) for NGC 4472
Forbes, D. A., Brodie, J. P., & Huchra, J. *Astron. J.* **113**, 887 (1997) for NGC 5846
Forbes, D. A., *et al.* *Mon. Not. Roy. astr. Soc.* **293**, 325 (1998) for NGC 1399 and 1404

Spectroscopic studies of extragalactic GCs include
Mould, J. R., Oke, J. B., & Nemec, J. M. *Astron. J.* **92**, 53 (1987)
Brodie, J. P., & Huchra, J. P. *Astrophys. J.* **379**, 157 (1991)
Grillmair, C. J., *et al.* *Astrophys. J.* **422**, L9 (1994)
Kissler–Patig, M., *et al.* *Astron. J.* **115**, 105 (1998)

Recent attempts to explain the presence of two GC subpopulations in massive ellipticals includes
Ashman, K. M., & Zepf, S. E. *Astrophys. J.* **384**, 50 (1992)
Forbes, D. A., Brodie, J. P., & Grillmair, C. J. *Astron. J.* **113**, 1652 (1997)
Cote, P., Marzke, R. O., & West, M. J. *Astrophys. J.* **501**, 554 (1998)

The formation of GCs from large giant molecular clouds is described in
Harris, W. E., & Pudritz, R. E. *Astrophys. J.* **429**, 177 (1994)
and subsequent papers.

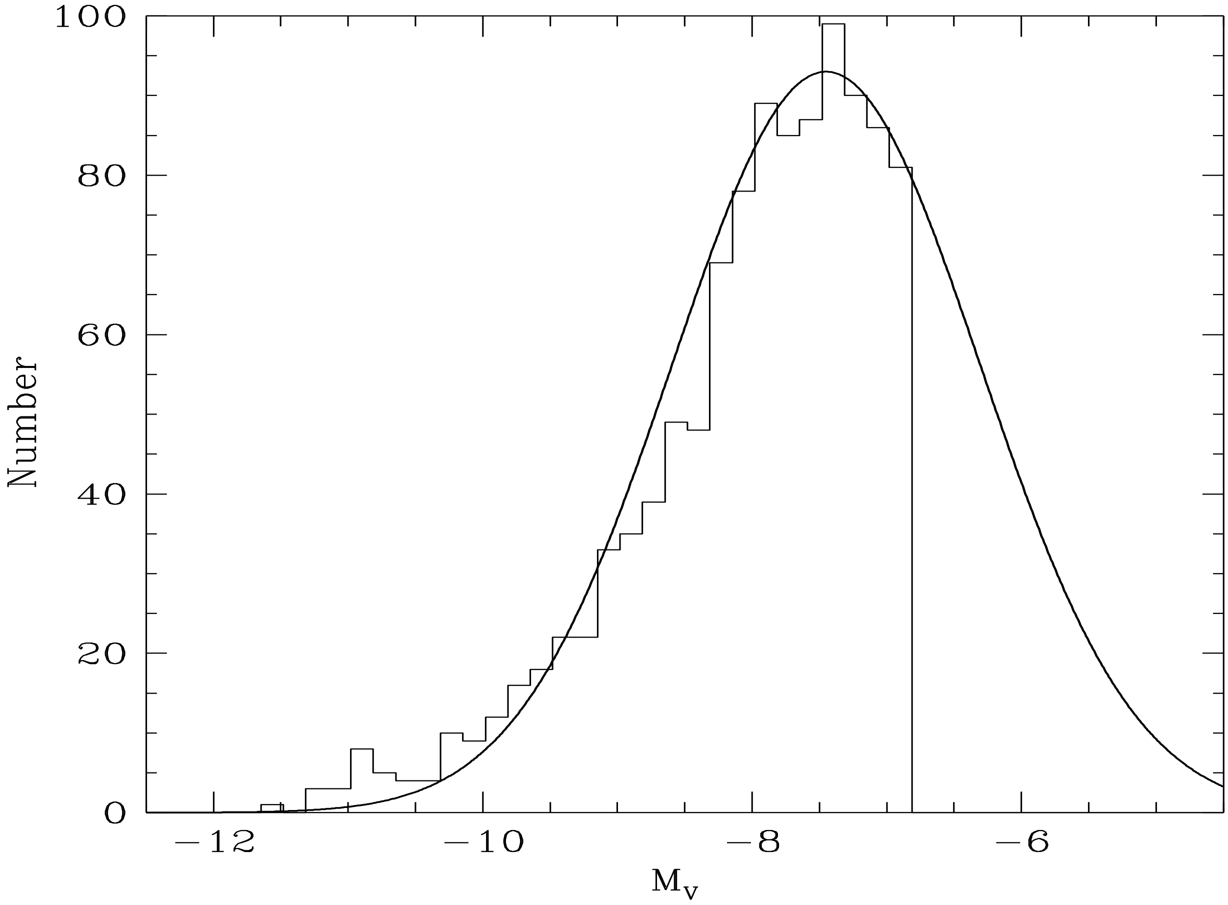

Figure 1: Globular cluster luminosity function for the giant elliptical galaxy NGC 5846. The histogram shows the binned globular cluster counts and the solid line shows the best fit Gaussian with a peak magnitude of $M_V^0 = -7.4$.

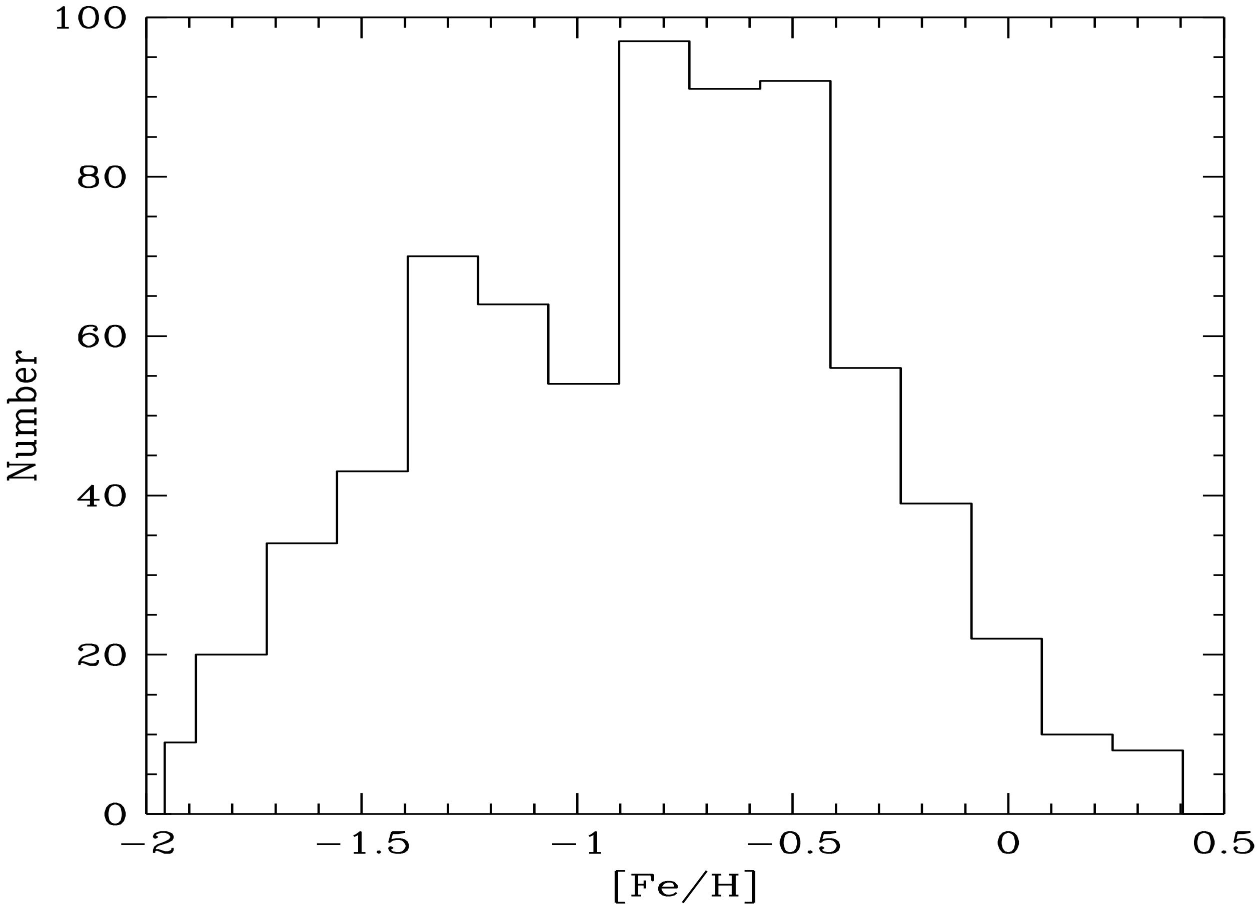

Figure 2: Globular cluster metallicity histogram for the giant elliptical galaxy NGC 5846. The histogram reveals clear evidence of a bimodal metallicity distribution indicating two distinct populations of globular clusters. Metallicities are estimated from photometry assuming a standard color-to-metallicity transformation.