THE FORMATION OF GIANT ELLIPTICAL GALAXIES AND THEIR GLOBULAR CLUSTER SYSTEMS

PATRICK CÔTÉ
Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council of Canada, 5071 West Saanich Road, Victoria, BC, V8X 4M6, Canada, and California Institute of Technology, Mail Stop 105-24, Pasadena, CA 91125; pc@astro.caltech.edu

RONALD O. MARZKE
Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council of Canada, 5071 West Saanich Road, Victoria, BC, V8X 4M6, Canada, and Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101; marzke@ociw.edu

AND

MICHAEL J. WEST
Department of Astronomy and Physics, Saint Mary's University, Halifax, NS, B3H 3C3, Canada; west@ap.stmarys.ca

Received 1997 August 20; accepted 1998 February 23

ABSTRACT

We examine the formation of giant elliptical galaxies using their globular cluster (GC) systems as probes of their evolutionary history. The bimodal distributions of GC metallicities in such galaxies are often cited as evidence for the formation of giant elliptical galaxies through mergers involving gas-rich spirals, with the metal-rich GCs forming during the merger process. We explore an alternative possibility: that these metal-rich clusters represent the galaxy's intrinsic GC population and the metal-poor component of the observed GC metallicity distribution arises from the capture of GCs from other galaxies, either through mergers or through tidal stripping. Starting with plausible assumptions for the luminosity function of galaxies in the host cluster and for the dependence of GC metallicity on parent galaxy luminosity, we show using Monte Carlo simulations that the growth of a preexisting seed galaxy through mergers and tidal stripping is accompanied naturally by the capture of metal-poor GCs whose chemical abundances are similar to those observed to surround giant ellipticals. We also investigate the spatial distribution of GCs in isolated galaxies of low and intermediate luminosity and conclude that, at the epoch of formation, the GC systems of such galaxies are likely to have been more spatially extended than their constituent stars. Thus, the capture of GCs through tidal stripping, unlike mergers, does not necessarily conserve GC specific frequency. Comparisons of model GC metallicity distributions and specific frequencies to those observed for the well-studied galaxies NGC 4472 (=M49) and NGC 4486 (=M87), the two brightest cluster members of the nearby Virgo cluster, show that it is possible to explain their bimodal GC metallicity distributions and discordant specific frequencies without resorting to the formation of new GCs in mergers or by invoking multiple bursts of GC formation. Finally, we discuss the possibility of using the ratio of the numbers of metal-poor to metal-rich GCs in giant elliptical galaxies as a diagnostic of their merger histories. We use this method to derive upper limits on the number of galaxies and total luminosity accreted to date by NGC 4472.

Subject headings: galaxies: clusters: general — galaxies: elliptical and lenticular, cD — galaxies: interactions — galaxies: star clusters

1. INTRODUCTION

An understanding of galaxy formation is one of the principal goals of modern astrophysics. An important first step in this campaign came with the demonstration by Toomre (1977) that mergers of late-type galaxies often produce end-products that bear a striking resemblance to giant ellipticals (gE's). In subsequent years, support has grown for the notion that mergers play an important, if not dominant, role in the formation of gE galaxies. Not only have investigations into the structure and kinematics of low-redshift gE's revealed fossil evidence of past mergers (Bender & Surma 1992; Tremblay & Merritt 1996; Faber et al. 1997), but theoretical models of galaxy formation invariably predict the growth of such galaxies through mergers (Kauffmann, White, & Guiderdoni 1993; Cole et al. 1994; Navarro, Frenk, & White 1995). Particularly close attention has been paid to cD galaxies and centrally dominant ellipticals in rich clusters since it is widely believed that the formation of these rare objects is fundamentally connected to the formation of the clusters themselves (see, e.g., Ostriker & Tremaine 1975; Hausman & Ostriker 1978; Merritt 1984, 1985; Mackie 1992; West 1994).

The globular cluster systems (GCSs) of gE galaxies have traditionally provided some of the most stringent constraints on models for the formation and evolution of these objects (see, e.g., Harris 1986). Since GCSs are the oldest and brightest objects that can be identified in gE galaxies, they are powerful tools for investigating the structure and chemical evolution of their host galaxies at the epoch of formation. Indeed, any model that seeks to explain the formation of gE's must be able to account for the number, chemical abundance, and spatial distribution of their associated systems of GCs.

One of the strongest arguments against the view that gE galaxies represent the remains of merged spirals (see, e.g., Toomre 1977) is that of van den Bergh (1982). Simply stated, ellipticals have GC specific frequencies (i.e., the
number of GCs per unit parent galaxy luminosity; see § 2.3) that are 2–3 times higher than those of spirals and irregulars (Harris 1991). Of course, these morphologically disparate galaxies exhibit very different clustering properties, so the extent to which their GCs have been influenced by environment is unclear (see West 1993). For instance, Forte, Martínez, & Muzzio (1982) suggested that the high specific frequency of NGC 4486, the centrally dominant galaxy in the Virgo cluster, might be explained by the capture of GCs through tidal stripping of less massive galaxies (cf. Muzzio 1987).

An additional constraint on galaxy formation models comes from the observed distribution of GC metallicities. Following initial suggestions of non-Gaussian metallicity distributions based on rather small samples of GCs (Ostrov, Geisler, & Forte 1993; Lee & Geisler 1993), a flurry of activity has now demonstrated beyond question that many, and perhaps most, luminous ellipticals show distinctly bimodal GC metallicity distributions (see, e.g., Whitmore et al. 1995; Elson & Santiago 1996). The number of gE’s having reliable GC metallicity distributions has now risen to roughly a dozen (see the recent compilation of Forbes, Brodie, & Grillmair 1997; hereafter FBG97). As these authors point out, the mean metallicity of the metal-rich component increases with the luminosity of the parent galaxy in a manner that is entirely consistent with that originally proposed by van den Bergh (1975). However, the position of the metal-poor peak shows no strong correlation with the properties of the host galaxy: it falls in the range $-1.5 \leq [\text{Fe/H}] \leq -0.7$ with little or no dependence on parent galaxy luminosity. The challenge for models of gE formation is therefore to explain not just the origin and size of these chemically distinct GC populations, but also to understand why the location of the metal-poor component peak depends so weakly on the luminosity of the parent galaxy.

One possible explanation for the high specific frequencies and bimodal GC metallicity distributions of gE’s is that the same mergers that give rise to the galaxies themselves also serve as catalysts for the formation of new GCs (Schweizer 1986; Ashman & Zepf 1992). In other words, the populations of metal-rich GCs associated with galaxies such as NGC 4486 (Whitmore et al. 1995; Elson & Santiago 1996) and NGC 4472 (Geisler, Lee, & Kim 1996) are interpreted as being the relics of merger-induced GC formation, whereas the metal-poor components are assumed to constitute the preexisting GCs. However, as Geisler et al. (1996) point out, both of these galaxies show bimodal GC metallicity distributions yet have specific frequencies that differ by nearly a factor of 3 (Harris 1986), so it seems unlikely that any single mechanism, such as merger-induced GC formation, can explain entirely the properties of their GCs. In light of these difficulties, FBG97 have argued that the chemically distinct GC populations in gE’s are indigenous to their parent galaxies and are formed in “two distinct phases of star formation from gas of differing metallicity.” Nevertheless, it is difficult to identify the mechanism responsible for the separate GC-formation episodes and for delaying the onset of the second burst.

In this paper, we show that bimodal GC metallicity distributions in gE’s are a natural consequence of the capture of GCs through mergers and tidal stripping of other galaxies. Our model differs from that of Ashman & Zepf (1992) in that we identify the metal-rich component of the GC metallicity distributions as the gE’s intrinsic GCS; mergers and tidal stripping will lead almost inevitably to a second peak in the range $-1.5 \leq [\text{Fe/H}] \leq -0.7$, with the exact position depending weakly on the shape of the luminosity function in the host cluster. We also conclude that, contrary to some earlier claims, the capture of GCs through tidal stripping of other cluster galaxies does not necessarily conserve specific frequency, particularly if the bulk of the tidally stripped GCs originated in low- and intermediate-luminosity galaxies. Finally, we use this model to explain the bimodal GC metallicity distributions and discordant specific frequencies of NGC 4472 and NGC 4486—the two brightest members of the Virgo cluster.

2. GC METALLICITY DISTRIBUTIONS IN gE GALAXIES: EVOLUTION DUE TO MERGERS AND TIDAL STRIPPING

We seek to model the GC metallicity distributions of gE galaxies, particularly those located in rich galaxy clusters. Simply stated, our goal is to predict the metallicity distribution of GCs acquired through mergers or through tidal stripping of other cluster galaxies. To do this, we must first consider three factors that together determine the total number of GCs of a given metallicity that are present in any ensemble of galaxies: (1) the number of GCs as a function of parent galaxy luminosity; (2) the relationship between GC metallicity and parent galaxy luminosity; and (3) the distribution of galaxy luminosities. Combining these three relations allows us to predict the metallicity distribution of GCs that may be captured by gE’s from other galaxies.

2.1. The “Zero-Age” $[\text{Fe/H}]-M_V$ Relation

There have been numerous compilations of mean GC metallicity, $[\text{Fe/H}]$, and parent galaxy absolute magnitude, $M_V$, in the recent literature. Although it is now established that the two parameters are related in a manner that is consistent with that originally proposed by van den Bergh (1975), the exact form of the correlation remains an open question. Harris (1991) and Durrell et al. (1997) have found a slope of $d[\text{Fe/H}]/dM_V = -0.17 \pm 0.04$, whereas Ashman & Bird (1993) have argued that genuine halo GCs have similar $[\text{Fe/H}]$, irrespective of parent galaxy luminosity. Part of this discrepancy arises from the sample definition, as some investigators have chosen to exclude the GCs of massive galaxies such as gE’s and luminous spirals from their analyses since, at the high luminosities that are characteristic of these objects, the observed $[\text{Fe/H}]-M_V$ relation shows large scatter.

Here, we attempt to go a step further and reconstruct the initial, “zero-age” $[\text{Fe/H}]-M_V$ relation over the entire range of spheroidal galaxy luminosities. As a result of the detailed observations of GC colors in gE’s, we now know that the scatter at the bright end of the present-day $[\text{Fe/H}]-M_V$ relation reflects the observed bimodality in the gE metallicity distributions. Using these detailed color distributions, FBG97 showed that it is the position of the metal-rich peak which correlates best with the gE luminosity. Part of this correlation suggests that the metal-rich GCs in these galaxies formed via the same mechanism as the GCs in galaxies of low- and intermediate-luminosity, and we are thus led to identify the metal-rich population as the gE’s primordial GCS. Our identification of the metal-rich GCs as the intrinsic component of the GCs in these galaxies may seem odd at first glance, particularly since GCs are believed to be among the first objects
to have formed in the early universe. However, analyses of absorption line indices in gE galaxies using stellar population models suggest that the gas from which these systems formed was likely to be preenriched, and that the galaxies themselves underwent rapid chemical enrichment by massive stars (see, e.g., de Freitas Pacheco 1996; Greggio 1997). Likewise, the extremely high star formation rates found by Pettini et al. (1997) for primeval galaxies at redshifts of $z \sim 3$ provide additional evidence for rapid chemical enrichment in such massive galaxies. Regardless of the exact mechanism responsible for the early chemical enrichment, Cohen, Blakeslee, & Ryzhov (1998) have recently provided compelling evidence that it must have proceeded on a very short timescale: Balmer-line indices for 150 GCs surrounding NGC 4486 indicate that the metal-rich GCs have ages of $\sim 13$ Gyr, indistinguishable from those of their metal-poor counterparts.

If the metal-rich GCs are the intrinsic population (i.e., those originally associated with the seed galaxy that we now observe as a gE), then the zero-age $[\text{Fe/H}]-M_\text{V}$ relation must be defined by the correlation between the seed galaxy luminosity and the position of the metal-rich peak of the GC metallicity distribution. The latter is an observed quantity; unfortunately, the seed galaxy luminosity is not. However, under the assumption that the total number of GCs in the galaxy population is conserved during the merger process, we can establish the range of possible seed galaxy luminosities from the observed distribution of GC metallicities. We explore two limiting cases: first, we posit that all of the metal-poor GCs were acquired through mergers. In this case, the observed fraction of GCs in the metal-rich peak is a direct measure of the luminosity of the initial seed (provided that the captured galaxies have roughly the same GC specific frequency; see § 2.3):

$$L_{\text{init}} = L_{\text{final}} \frac{N_{\text{mr}}}{N_{\text{mr}} + N_{\text{mp}}},$$  \hspace{1cm} (1)

or, equivalently,

$$M_\text{V} = M_\text{V}^\prime - 2.5 \log \left( \frac{N_{\text{mr}}}{N_{\text{mr}} + N_{\text{mp}}} \right),$$  \hspace{1cm} (2)

where $M_\text{V}^\prime$ is the final absolute magnitude of the gE galaxy, and $N_{\text{mr}}$ and $N_{\text{mp}}$ refer to the number of metal-rich and metal-poor GCs, respectively. Since all three of these present-day quantities are available from FBG97, we can compute an initial, premerger $[\text{Fe/H}]-M_\text{V}$ relation directly from these observations (assuming that the dynamical friction timescale in the merged system is long, and that the metal-rich and metal-poor GCs are equally susceptible to destruction; see § 4.3). Strictly speaking, our “zero-age” relation is more accurately described as the initial $[\text{Fe/H}]-M_\text{V}$ relation as it would appear today provided the subsequent evolution has been totally passive (i.e., we apply no magnitude correction to account for the fading stellar populations in the original galaxies, since the same correction would then need to be subtracted when comparing to nearby gE's).

In the opposite limit, we make the extreme assumption that the initial luminosity is the same as the final luminosity: i.e., that the metal-poor GCs are stripped from neighboring galaxies without any accompanying field stars. This is obviously a firm upper limit to the seed luminosity and a conservative bound at that; stripping will inevitably transfer at least some field stars to the seed galaxy.\(^4\)

The resulting range of possible $[\text{Fe/H}]-M_\text{V}$ relations is shown in Figure 1. For the gE's, we plot the 11 gE's listed in FBG97. For each galaxy, we adopt the mean metallicity for the metal-rich GC population reported in FBG97 and assume an uncertainty in $[\text{Fe/H}]$ of $\pm 0.2$ dex. The modest corrections derived from equation (2) (e.g., $M_\text{V} - M_\text{V}^\prime \approx 1$ mag) are shown as the horizontal arrows in Figure 1, where we have assumed an uncertainty of $\pm 0.3$ mag in $M_\text{V}^\prime$ for each galaxy.

Because dwarf galaxies are less likely to have merged with smaller systems, they provide a more direct view of the original $[\text{Fe/H}]-M_\text{V}$ relation. A search of the literature revealed nine dwarf elliptical galaxies for which there exist data on the chemical abundance of their GCs. For these galaxies, we have plotted the raw values of $[\text{Fe/H}]$ and $M_\text{V}$ in Figure 1. For each galaxy, we have taken the best available estimate for the metallicity of each individual GC. For a few of the galaxies, there exist multiple determinations of the metallicities of some GCs; in such cases, we give preference to spectroscopic measurements. We then calculate the mean GC metallicity $[\text{Fe/H}]$, the intrinsic metallicity dispersion $\sigma_i$, and their respective uncertainties for each galaxy.

We also include in our sample NGC 1380, the only lenticular (S0) galaxy in which accurate metallicities have been measured for a large number of GCs (Kissler-Patig et al. 1997). Like most gE galaxies, NGC 1380 shows a bimodal GC metallicity distribution. Following Kissler-Patig et al. (1997), we adopt $[\text{Fe/H}] = 0.15 \pm 0.5$, $N_{\text{mr}} = 382 \pm 20$, and $N_{\text{mp}} = 182 \pm 20$.

Our sample of 21 galaxies is listed in Table 1, whose columns record the name, morphological type, and absolute visual magnitude for the parent galaxy of each GCS, the total number of GCs having measured metallicities, and the corresponding mean GC metallicity and intrinsic dispersion; references are given in the final two columns. Since it is clear that the relationship between $[\text{Fe/H}]$ and $M_\text{V}$ is nonlinear, we parameterize the $[\text{Fe/H}]-M_\text{V}$ relation in the form

$$[\text{Fe/H}] = a_0 + a_1 M_\text{V} + a_2 M_\text{V}^2.$$  \hspace{1cm} (3)

We then find the constants $a_0$, $a_1$, and $a_2$ that minimize the goodness-of-fit statistic

$$\chi^2(a_0, a_1, a_2) = \sum_i^N \frac{([\text{Fe/H}]_{\text{obs}} - a_0 - a_1 M_\text{V}_{i,j} - a_2 M_\text{V}^2_{i,j})^2}{\epsilon_{[\text{Fe/H}]_{\text{obs}}}^2 + (a_1 + 2a_2 M_\text{V}_{i,j})^2}.$$  \hspace{1cm} (4)

where $\epsilon_{[\text{Fe/H}]_{\text{obs}}}$ is the uncertainty in the mean GC metallicity, and we have made use of the fact that

$$\frac{d[\text{Fe/H}]}{dM_\text{V}} = a_1 + 2a_2 M_\text{V}.$$  \hspace{1cm} (5)

Minimizing equation (4) yields $a_0 = 2.31$, $a_1 = 0.638$, and $a_2 = 0.0247$. The resulting $[\text{Fe/H}]-M_\text{V}$ is shown as the

---

\(^4\) One might also consider a third case: that field stars are actually accreted more efficiently than GCs. However, the fact that no GCs are observed to be less extended that the underlying stellar halo makes this possibility seem very remote.
Fig. 1.—Mean GC metallicity, $[\text{Fe/H}]$, plotted against absolute visual magnitude, $M_V$, of the parent galaxy. For the $gE$ and lenticular galaxies, we plot the mean metallicities of their metal-rich GCs against the absolute visual magnitudes given by eq. (2). The dotted line shows the second-order polynomial that best fits the data, taking into account the uncertainties in both observables. The uncorrected magnitudes for the giant galaxies are indicated by the arrows (see text for details). Some of the scatter among the corrected magnitudes for the giant galaxies undoubtedly reflects the presence of radial gradients in the ratio of metal-rich and metal-poor GCs (i.e., the GCs are not observed over the same metric radii in all of the galaxies). The dashed line shows the $[\text{Fe/H}]-M_V$ relation obtained using the uncorrected magnitudes.

dotted line in Figure 1; the dashed line indicates the relation found without the correction given by equation (2). The difference between the two fits is never large; we show in § 3 that our conclusions depend very little on these small variations in the $[\text{Fe/H}]-M_V$ relation.

Finally, we investigate the possibility that the intrinsic dispersion in GC metallicity also depends on the luminosity of the parent galaxy. Based on the nine galaxies in Table 1 that have reliable estimates for $\sigma$, we find a weighted mean dispersion of $\sigma = 0.28 \pm 0.05$ dex with no significant evidence for a correlation between $\sigma$ and $M_V$. The increased metallicity dispersion among luminous galaxies noted by previous investigators (see, e.g., Harris 1991) is likely to be a consequence of their bimodal GC metallicity distributions, which went unnoticed for many years because of the small sample sizes and the limited metallicity sensitivity of the color indices used in early photometric surveys (see the discussion in Geisler et al. 1996).

2.2. The Initial Luminosity Function of Cluster Galaxies

Galaxies in clusters span an enormous range in luminosity and mass. The present-day galaxy luminosity function (LF) is determined both by the cosmological conditions when galaxies began to form and by subsequent evolution of the galaxy population. In the case of an $\Omega = 1$, cold dark matter universe, the predicted mass spectrum is steep at all redshifts; at $z = 0$, $dN/dM \propto M^{-2}$ at the low-mass end. The corresponding LF follows from the cycle of gas cooling, star formation, and supernova feedback, none of which (with the possible exception of the first) are well understood. Nevertheless, solutions derived using quite disparate physical assumptions yield a remarkably consistent picture. Depending mostly on the strength of the feedback, the derived LFs are always steep at the faint end: $-14 \leq \log N/ d \log L \leq -2.0$ (see, e.g., Kauffmann, Guiderdoni, & White 1994; Cole et al. 1994; Baugh, Cole, & Frenk 1996).

Unfortunately, our theoretical understanding of galaxy formation remains incomplete, and direct extrapolation of the redshift-dependent LFs is still our best hope for reconstructing the initial LF. Of the many analytic expressions proposed to describe the observed galaxy LF, $\phi(L_V)$, the most widely used is the Schechter (1976) function:

$$\phi(L_V) dL_V = \phi^* (L_V/L_V^*)^\alpha \exp(-L_V/L_V^*) d(L_V/L_V^*),$$

where $\phi^*$ is a parameter related to the number of galaxies per unit volume, $L_V^*$ is a characteristic luminosity (measured
for our purposes in the $V$-band, and $a$ is a parameter that determines the relative number of faint and bright galaxies.

Enormous effort has gone into measuring the galaxy LF both in the field and in clusters (see, e.g., the reviews by Bingelli, Tammann, & Sandage 1988; Ellis 1997). Virtually all studies agree that at low redshift, the LF turns over sharply near $L^*_{P}$ and is essentially flat (in log $\phi$ -- log $L_{P}$) down to approximately 3 magnitudes below $L^*_{P}$. Fainter than this, however, there is considerable disagreement over the LF slope. For instance, Marzke et al. (1994) find evidence for a steep upturn of $a \approx -1.9$ at the faint end of the field LF, while Ellis et al. (1996) claim that the LF remains flat to at least 4.5 magnitudes fainter than $L^*_{P}$. Recently, Loveday (1997) has suggested that a better parameterization of the field LF is a double Schechter function with a very steep faint-end slope of $\zeta \approx -2.8$.

Several recent studies of the cluster LF have produced generally similar results. For example, in their study of the Coma cluster, Biviano et al. (1995) have shown that a two-component LF, consisting of a Gaussian at the bright end and a steep Schechter function at the faint end, closely resembles the observed LF. Smith, Driver, & Phillips (1997), Wilson et al. (1997), Lopez-Cruz et al. (1997), and Trentham (1997) find similar behavior in several rich clusters over the redshift range $0 \lesssim z \lesssim 0.2$. However, these authors disagree on the question of how cluster LFs have evolved with time; for instance, Smith et al. (1997) find little difference between clusters at $z = 0$ and $z = 0.2$, while Wilson et al. (1997) claim that the galaxy colors suggest the presence of a fading population of dwarf irregulars whose star formation was truncated soon after $z \approx 0.2$.

Given the uncertainty in both field and cluster LFs at early epochs, we are trying to explore a range of possibilities for the initial LF. We are particularly interested in the initial LF in clusters, since it is bright cluster ellipticals that we are trying to understand. Given the observations, it seems likely that the present-day LF in, for example, the Virgo cluster represents a lower bound on the $HST$ data of Whitmore et al. 1995 and adopt $[\text{Fe/H}] = 0.0$ as the location of the metal-rich peak; 53% of the GCs in this sample belong to the metal-rich component.

### Table 1

| Galaxy       | Type | $M_V$  | $N_{\text{GC}}$ | $[\text{Fe/H}]$ | $\sigma_i$ | References ($M_V$) | References ([Fe/H]) |
|--------------|------|--------|-----------------|-----------------|-----------|--------------------|--------------------|
| **Dwarf Ellipticals** |
| Fornax       | dE0  | $-12.7 \pm 0.4$ | 5 | $-1.77 \pm 0.12$ | $0.20 \pm 0.07$ | 1 | 2, 3 |
| Sagittarius  | dE   | $-13.2 \pm 0.5$ | 3 | $-1.79 \pm 0.13$ | $0.21 \pm 0.09$ | 4, 5 | 6 |
| M81 -- FSD1   | dE   | $-14.25 \pm 0.26$ | 1 | $-1.8 \pm 0.3$ | | 7 | 7 |
| NGC 147      | E5p  | $-15.45 \pm 0.23$ | 2 | $-2.06 \pm 0.33$ | $0.27 \pm 0.16$ | 8, 9 | 10 |
| NGC 185      | E3p  | $-15.47 \pm 0.23$ | 5 | $-1.61 \pm 0.23$ | $0.31 \pm 0.11$ | 8, 11 | 10 |
| VCC 1254     | dE0, N | $-16.65 \pm 0.20$ | 13 | $-1.51 \pm 0.13$ | $0.45 \pm 0.09$ | 12, 13 | 12 |
| NGC 205      | E5p  | $-16.61 \pm 0.32$ | 6 | $-1.48 \pm 0.09$ | $0.14 \pm 0.06$ | 8, 14 | 8 |
| VCC 1386     | dE3, N | $-16.99 \pm 0.21$ | 8 | $-1.43 \pm 0.10$ | $0.21 \pm 0.52$ | 12, 13 | 12 |
| NGC 3115 DW1 | dE1, N | $-17.62 \pm 0.14$ | 24 | $-1.25 \pm 0.13$ | $0.65 \pm 0.09$ | 15, 16 | 16 |
| **Lenticulars** |
| NGC 1380     | S0   | $-21.0 \pm 0.3$ | | $0.15 \pm 0.5$ | | 17 | 17 |
| **Giant Ellipticals** |
| NGC 1052     | E4   | $-20.25 \pm 0.3$ | | $-0.4 \pm 0.2$ | | 18 | 18 |
| NGC 1399     | E1p  | $-20.85 \pm 0.3$ | | $-0.57 \pm 0.2$ | | 18 | 18 |
| NGC 1404     | E1   | $-20.1 \pm 0.3$ | | $-0.2 \pm 0.2$ | | 18 | 18 |
| NGC 3311     | E+2  | $-21.4 \pm 0.3$ | | $0.15 \pm 0.2$ | | 18 | 18 |
| NGC 3923     | E4-5 | $-21.4 \pm 0.3$ | | $0.0 \pm 0.2$ | | 18 | 18 |
| NGC 4472     | E2   | $-21.6 \pm 0.3$ | | $0.05 \pm 0.2$ | | 18 | 18 |
| NGC 4486     | E+0-1p | $-21.8 \pm 0.3$ | | $0.0 \pm 0.2$ | | 18 | 18 |
| NGC 4494     | E1-2 | $-20.0 \pm 0.3$ | | $-0.3 \pm 0.2$ | | 18 | 18 |
| NGC 5128     | E0p  | $-21.0 \pm 0.3$ | | $-0.2 \pm 0.2$ | | 18 | 18 |
| NGC 5846     | E0-1 | $-22.3 \pm 0.3$ | | $-0.2 \pm 0.2$ | | 18 | 18 |
| IC 1459      | E3   | $-20.5 \pm 0.3$ | | $0.0 \pm 0.2$ | | 18 | 18 |

$^a$ Following Durrell et al. 1997, we omit Terzan 7 from the calculation of $[\text{Fe/H}]$ since its high metallicity (Da Costa & Armandroff 1995), younger age (Layden & Sarajedini 1997), and marginally discrepant velocity (Da Costa & Armandroff 1995) suggest that it may not be a true Sagittarius GC.

$^b$ Shows marginal evidence for a trimodal GC metallicity distribution et al. We combine their intermediate- and metal-rich populations and find that 55% of the GCs comprise a metal-rich component with peak at $[\text{Fe/H}] = -0.57$.

$^c$ We adopt the dispersion of the metal-rich component of the double Gaussian distribution that best fits the photometrically derived GC metallicity distribution of Gessler et al. 1996.

$^d$ Shows marginal evidence for a trimodal GC metallicity distribution (Forbes et al. 1997). We fit a double Gaussian distribution to the $HST$ data of Whitmore et al. 1995 and adopt $[\text{Fe/H}] = 0.0$ as the location of the metal-rich peak; 53% of the GCs in this sample belong to the metal-rich component.

### References
- (1) Mateo et al. 1991; (2) Buonanno et al. 1985; (3) Dubath, Meylan, & Mayor 1992; (4) Ibata, Gilmore, & Irwin 1995; (5) Mateo et al. 1996; (6) Da Costa & Armandroff 1995; (7) Caldwell et al. 1997; (8) de Vaucouleurs et al. 1991; (9) Han et al. 1997; (10) Da Costa & Mould 1988; (11) Lee, Freedman & Madore 1993; (12) Durrell et al. 1997; (13) Ferrarese et al. 1996; (14) Lee 1996; (15) Ciardullo, Jacoby, & Tonry 1993; (16) Durrell et al. 1996; (17) Kisler-Privat et al. 1997; (18) Forbes et al. 1997.
2.3. Choice of GC-Specific Frequency

During interactions with other galaxies, the number of GCs available for capture by a gE galaxy depends on the luminosity of each passing galaxy. To first order, the number of GCs, \( N_{GC} \), scales directly with absolute magnitude of the parent galaxy. A convenient means of expressing the number of GCs per unit galaxy luminosity is the specific frequency, \( S_n \), which was first proposed by Harris & van den Bergh (1981),

\[
S_n = N_{GC} 10^{0.4(M_V + 1.5)} .
\]

For dwarf ellipticals and gE’s in rich clusters, \( S_n \approx 5 \) with a dispersion of \( \sigma(S_n) \approx 1 \) (see, e.g., Harris 1991). In what follows, we assume specific frequency to be independent of \( M_V \) (see Harris 1991, and references therein) and adopt a specific frequency of \( S_n = 5 \pm 1 \) for the interacting galaxies. In § 3.1, we discuss the implications of this choice of \( S_n \).

3. MONTE CARLO SIMULATIONS OF GC METALLICITY DISTRIBUTIONS

Having specified the \([\text{Fe}/\text{H}] - M_V\) relation, the specific frequency \( S_n \), and the initial galaxy LF, it is possible to simulate the metallicity distribution of GCs captured by a gE galaxy. As discussed in § 2.1, we assume that the luminosity of the initial seed galaxy, \( L_{init} \), has grown by the amount \( L_{cap} \) as a result of the capture of smaller galaxies. Therefore, we calculate \( L_{cap} \) for each gE as

\[
L_{cap} = L_{final} \frac{N_{mp}}{N_{mr} + N_{mp}} ,
\]

and we select at random a galaxy that, based on the probability distribution given by equation (6), may have merged with the initial gE galaxy. Since, by definition, the gE galaxy cannot consume an object larger than itself, we adopt a maximum luminosity of \( L_{init} \) for equation (6). Likewise, we assume that the faint-end of the LF is truncated at \( L_{vir} = 10^{7} L_{vir} \). This is the luminosity of the Fornax dwarf galaxy, the faintest galaxy known to harbor its own GCS (Mateo et al. 1991; Harris 1991). We repeat this process, recording the combined luminosity of the captured galaxies,

\[
L_{cap} = \sum_{k=1}^{N_{cap}} L_k ,
\]

until the condition

\[
L_{cap} \geq L_{final} - L_{init}
\]

is satisfied. At this point, the simulations are terminated and the total number of captured galaxies, \( N_{cap} \), is recorded.

For each captured galaxy, we calculate the total number of associated GCs using equation (7), including a random dispersion of \( \sigma(S_n) = 1 \) for both the captured and seed galaxies. A metallicity is assigned randomly to each captured GC from an assumed Gaussian distribution with mean metallicity, \([\text{Fe}/\text{H}]\), determined from equation (3), and an intrinsic spread of \( \sigma_\text{fg} = 0.28 \) dex, irrespective of the host galaxy luminosity. We account for the scatter in the \([\text{Fe}/\text{H}] - M_V\) relation by including a random dispersion of 0.2 dex when assigning a mean GC metallicity to each galaxy (as discussed in § 2.1). We repeat this process one last time for the initial gE galaxy and combine the intrinsic and captured GCs to create a final simulated GC metallicity distribution.

For simplicity, we assume that all galaxies have equal merger probabilities; in reality, the likelihood of capture may be a function of galaxy mass. For instance, circular orbits through an isothermal dark halo decay on a timescale inversely proportional to the mass \( M \) (Ostriker & Tremaine 1975; Tremaine 1976). For a more general distribution of orbits, however, the situation is more complicated. Cole et al. (1994) have argued on the basis of the \( N\)-body/hydrodynamical simulations of Navarro et al. (1995) that the merger timescale, \( \tau \), shows a relatively weak mass dependence: \( \tau \propto M^{-0.25} \). Since the exact form of the dependence is still under debate, we have chosen not to include an explicit mass dependence in our models. We simply note that for a constant mass-to-light ratio, the merger probability derived by Cole et al. (1994) is proportional to \( L^{0.25} \). If this is indeed the correct dependence, then each of the LFs considered in the following analysis would need to be corrected by an amount \( \Delta a = -0.25 \) (i.e., an assumed LF slope of \( a = -1.5 \) would correspond to an actual slope of \( a = -1.75 \)).

As an illustration of this method, we now compare the simulated GC metallicity distributions to those observed for a well-studied gE galaxy: NGC 4472, the brightest member of the Virgo cluster.

3.1. The GC Metallicity Distribution of NGC 4472

The GC metallicity distribution of NGC 4472 represents the best existing data set for any gE galaxy in terms of calibration, metallicity sensitivity, and large sample size (Geisler et al. 1996). This last factor is usually the most severe limitation on other data sets, which typically consist of relatively small numbers of GCs. In such cases, the lower order moments of the GC metallicity distributions (such as the mean metallicities of the metal-rich and metal-poor peaks) are reasonably well constrained, but it is impossible to model the detailed shape of the distribution function. Recent summaries of bimodal GC metallicity distributions in gE galaxies can be found in Geisler et al. (1996) and FBG97.

As mentioned in § 2.2, choosing the values of \( a \) and \( L_f \) (or, alternatively, \( M_f \)) to be used in the simulations is complicated by the possibility that the original Virgo LF differs significantly from the present-day LF. Sandage, Binggeli, & Tammann (1985) measure \( a = -1.25 \) and \( M_f = 10.6 \) for the present-day composite LF in Virgo, although the slope of the LF at the faint end depends rather strongly on the sample definition. For instance, the best-fit LF for dE’s and gE’s alone has \( a = -1.45 \). We have therefore carried out simulations for a range of assumed LF slopes, bearing in mind that \( a = -1.25 \) is likely to represent a lower limit on the initial LF slope. The evolution of \( M_f \) is expected to proceed more slowly, so we have combined the present-day estimate of \( M_f = 10.6 \) with the Virgo distance modulus of \( \Delta M_B = 31.04 \) mag (Ferrarese et al. 1996), a Galactic reddening of \( E(B-V) = 0.10 \), and a mean galaxy color of \( B-V = 0.9 \) at \( M_f \) (see, e.g., Roberts & Haynes 1994) to find \( M_f = -21.7 \).

Figures 2, 3, and 4 compare the results of our simulations with the observed GC metallicity distribution for NGC 4472 (Geisler et al. 1996). As initial conditions for the simulations, we assume \( M_f = -21.6 \) and \( M_f = -22.7 \) for NGC 4472 (see below). We show four different simulations for each assumed LF slope, chosen to constitute a representative sample of simulated GC metallicity distributions. As a
Fig. 2.—Observed and simulated GC metallicity distributions for NGC 4472. The filled triangles indicate the observations of Geisler et al. (1996), while the open squares indicate simulated GC metallicity distributions for an assumed LF slope of $a = 1.2$. The simulated data have been binned in the same manner as the real data. The simulated GC metallicity distributions shown in the four different panels are a representative sample for this choice of LF slope. The best-fit double Gaussian is shown as the solid line, while the dotted lines indicate the two separate components. The total number of captured galaxies, $N_{\text{cap}}$, is recorded in the upper left corner of each panel, along with the goodness-of-fit statistic $\chi^2_{\text{sim}}$ (see text for details).

measure of the similarity between the observed and simulated distributions, we calculate the goodness-of-fit statistic

$$\chi^2_{\text{sim}} = \frac{1}{N_{\text{bin}}} \sum_{i=1}^{N_{\text{bin}}} \frac{(N_{\text{obs},i} - N_{\text{sim},i})^2}{N_{\text{obs},i} + N_{\text{sim},i}}$$

for each simulation. The corresponding values of $\chi^2_{\text{sim}}$ are recorded in Figure 2, 3, and 4.

If we assume that all of the metal-poor GCs in NGC 4472 were captured in mergers, then it is possible to calculate the total number of galaxies consumed by NGC 4472.\(^5\) Of course, this is possible simply because the capture of even a single GC in a merger naturally entails the capture of the entire host galaxy, something that is not true of tidal stripping. The total number of captured galaxies is recorded in each panel of Figures 2, 3, and 4. Clearly, the stochastic nature of the merger process strongly influences the final GC metallicity distribution, as a galaxy may increase its luminosity in any number of ways: by swallowing many small dwarfs or just a few large galaxies. The effect is, however, less pronounced for steep LFs since there are relatively fewer luminous galaxies, and hence the initial gE is unlikely to capture a comparably bright galaxy. In such cases, the final simulated GC metallicity distribution is almost always bimodal.

According to FBG97, the absolute visual magnitude of NGC 4472 is $M_V = -22.7$. Since the observations of Geisler et al. (1996) indicate that 35% of the GCs in NGC 4472 are metal-rich, our estimate for the absolute magnitude of the captured component in NGC 4472 is $M_V = -21.6$, which corresponds to $\sim 65\%$ of its present-day luminosity. The number of captured galaxies depends on both $S_n$ and the assumed LF: for LF slopes of $a = -1.2$,
Fig. 3.—Same as in Fig. 2 except for an LF slope of $a = -1.5$

-1.5, and -1.8, the median number of captured galaxies found in 1000 Monte Carlo simulations (see below) are $44^{+10}_{-16} (S_n)$, $148^{+30}_{-46} (S_n)$, and $508^{+108}_{-106} (S_n)$, respectively. Although the total number of captures depends very strongly on the assumed LF, the majority of the captured galaxies are in most cases low-luminosity systems, each of which contributes relatively few GCs. For our best-fit slopes of $a \sim -1.8$, the bulk of the metal-poor GCs originate in galaxies having absolute magnitudes in the range $-15 \lesssim M_V \lesssim -19$.

It is worth bearing in mind that such estimates for the number of captured galaxies are best interpreted as upper limits since tidal stripping may have produced some of the observed metal-poor GCs. However, in the absence of large variations in $S_n$ within a single galaxy cluster—or between clusters that are morphologically and dynamically similar—this technique may provide a useful method of comparing gE merger histories.

### 3.1.1. Frequency of Bimodality

It is a remarkable fact that the GC metallicity distributions of virtually all well-studied gE's to date show some evidence for bimodality. How frequently do our simulated GC metallicity distributions exhibit such bimodality? To answer this question, we have generated 1000 simulated GC metallicity distributions for NGC 4472, assuming LF slopes of $a = -1.2$, -1.5, and -1.8. As before, the simulations have been binned in exactly the same fashion as the observations of et al. For each simulation, we fit single- and double-Gaussian distributions to the resulting GC metallicity distribution and determine the reduced $\chi^2$ for both parameterizations. An $F$-ratio test is then used to determine if the improvement obtained by including a second component is warranted (Bevington 1969). In the three panels in the left-hand column of Figure 5, we show histograms of the $F$-ratio statistic for 1000 simulated distributions having LF slopes of $a = -1.2$, -1.5, and -1.8. As expected, bimodality becomes increasingly obvious as LF slope is increased; for $a = -1.8$, 94% of the simulated distributions show bimodality. However, even for rather flat LFs, the majority of the simulations are bimodal. For instance, for an LF slope of $a = -1.2$, 80% of the simulated distributions are bimodal.

$^6$ The number of degrees of freedom, $v$, is 10 for the single Gaussian and seven for the double Gaussian.
distributions show statistically significant evidence for bimodality. It is also worth pointing out that although the simulated GC metallicity distributions are, in the vast majority of cases, well described by double Gaussians, for rather steep LFs there are occasions when more than two peaks are evident in the simulations. 

Histograms for the goodness-of-fit statistic, $\chi^2$, are given in the right-hand panels of Figure 5. Since the absolute calibration of this statistic depends on the appropriateness of our assumption of Poisson uncertainties in the observed and simulated distributions, it is probably best to interpret the $\chi^2$ values in a relative sense.

### 3.1.2. Dependence on LF Slope

FBG97 have shown that, for the dozen or so gE’s that are known to have bimodal GC metallicity distributions, the location of the metal-poor peak does not appear to obey the relation between GC metallicity and host galaxy absolute magnitude defined by the GCSs of fainter galaxies and the metal-rich GCSs of gE galaxies. Instead, the metal-poor peaks for these galaxies fall in the range $-1.5 \leq [\text{Fe/H}] \leq -0.5$, with little or no correlation with $M_V$. A possible explanation for this curious result is given in Figure 6, which shows the mean location of the metal-poor peak as a function of LF slope. (The dependence on $M_V$ is much weaker.) The metallicity distribution of the captured GCs shows a maximum in the range $-1.5 \leq [\text{Fe/H}] \leq -0.7$, though at fixed slope there is a scatter of $\approx 0.15$ dex in the peak location, a consequence of the stochastic nature of galaxy mergers. Thus, the weak correlation between the location of this peak and the luminosity of the gE is a result of the fact that the mean metallicity of the metal-poor GCs is determined primarily by (1) the detailed merger histories of the individual gEs and (2) the initial LF of the host cluster (see §6).

To explore this issue more closely, we have measured the location of the metal-poor GC peak as a function of both $z$ and final gE magnitude, $M_V$. Figure 7 shows the data of

---

*Note that Lee & Geisler (1993) have suggested that the GC metallicity distribution in M87 shows evidence for three peaks based on ground-based photometry for $\approx 400$ GC candidates, although Whitmore et al. (1995) have concluded from Hubble Space Telescope (HST) photometry for more than 1000 GC candidates that the GC metallicity distribution is more likely bimodal in nature.*

---

**Fig. 4.**—Same as in Fig. 2 except for an LF slope of $a = -1.8$.
FIG. 5.—Left-hand column: Histograms of the logarithm of the ratio $\chi_{10}^2/\chi^2_7$ for 1000 simulated metallicity distributions. For each simulation, the reduced $\chi^2$ of the best-fit single Gaussian, $\chi_{10}^2$, is measured along with the reduced $\chi^2$ of the best-fit double Gaussian, $c$. The three different panels correspond to input LF slopes of $\alpha = -1.2, -1.5,$ and $-1.8$. The dashed line in each panel indicates for ratios greater than this, the inclusion of the three additional free parameters is justified with 95% confidence according to an $F$-ratio test. Right-hand column: Histograms of the logarithm of the goodness-of-fit statistic, $\chi^2_{sim}$, for the observed and simulated distributions for NGC 4472. In each case, the LF slope is the same as is indicated in the panel to its left. The median value of $\chi^2_{sim}$ is given in the upper left-hand corner of each panel, along with one-half the difference between the upper and lower quartiles.

FBG97 (open squares) along with the results of our Monte Carlo experiments. Although the shift in the peak metallicity with $M_V$ at fixed $z$ is modest (particularly for steep LFs) and the intrinsic scatter in [Fe/H] is $\approx 0.15$ dex, there is a slight tendency for the brighter gE’s to capture GC populations that are more metal-rich than those accreted by their fainter counterparts (since the brighter gE’s are able to accommodate the capture of more luminous intruder galaxies). For simplicity, we have assumed equal numbers of metal-rich and metal-poor GCs in the simulations; the observed galaxy-to-galaxy scatter in this ratio (see Table 1 of FBG97) is expected to weaken further the correlations shown in Figure 7, as are differences in the initial LFs.

4. A POSSIBLE ORIGIN FOR HIGH-$S_n$ GALAXIES: THE CASE OF NGC 4486

Up to now, we have made no distinction between GCs acquired through mergers or through tidal stripping since the exact mechanism by which they are captured will have little effect on the shape of the GC metallicity distribution (see § 2.1). However, since tidal stripping will remove only a fraction of the GCS of an individual galaxy, the number of galaxies that must be stripped in order to explain the observed number of metal-poor GCs will always exceed the number of galaxies that must be captured. Moreover, the GCSs of many galaxies are known to be more spatially extended than the underlying halo light (see, e.g., Harris 1991), so it is possible that the capture of GCs by tidal stripping will affect the specific frequency of the gE differently than will capture by mergers (i.e., by preferentially stripping high-$S_n$ material from the outer regions of passing galaxies). We now explore the consequences of tidal stripping on specific frequency.

4.1. The Initial GCSs of Low- and Intermediate-Luminosity Galaxies

The possibility that the high specific frequencies of some centrally dominant galaxies can be explained through tidal
stripping of GCs from other cluster members was first suggested by Forte et al. (1982). Several potential problems with this scenario were pointed out by van den Bergh (1984) and McLaughlin, Harris, & Hanes (1994), the most serious being the objections (1) that, since tidal stripping will remove halo stars and GCs in equal proportion, such a process cannot explain why the specific frequency of NGC 4486 is roughly 3 times greater than that of other Virgo gE’s; and (2) that the high specific frequency problem in NGC 4486 is not just confined to the galaxy’s envelope, since even at a galactocentric distance of 1–2, the specific frequency is \( S_a \approx 10 \). Given that tidal stripping has long been recognized as providing a natural means of producing the extended envelopes of cD galaxies (see, e.g., Richstone 1976), it is worth investigating this issue more closely, particularly in light of several recent detections of probable tidally stripped debris in galaxy clusters, including intergalactic stars (Ferguson, Tanvir, & von Hippel 1998) and intergalactic planetary nebulae (Theuns & Warren 1997; Ciardullo et al. 1998; Méndez et al. 1997). The results of the planetary nebula surveys are particularly interesting since these studies suggest that \( \sim 26\%–50\% \) of the total Virgo luminosity may reside in intergalactic stars; if so, and if this diffuse component has a “normal” GC specific frequency of \( S_a = 5 \), then it should be accompanied by a population of 35,000–75,000 tidally stripped GCs. For higher specific frequencies (see below), the total number of GCs will of course be larger still.
Apart from this observational evidence, there are theoretical reasons to believe that the tidal field of the cluster potential well acts to remove matter from galaxy halos on timescales comparable to the crossing time (Peebles 1970; Gunn 1977; Merritt 1985). Consequently, this process, along with collisional stripping produced during galax-galaxy encounters (which usually operates on much longer timescales; Merritt 1984), may be responsible for the different kinds of tidal debris mentioned above, particularly since low-and intermediate-luminosity galaxies may be strongly stripped by the mean cluster tidal field.

It is recognized that the GCSs of many galaxies are more extended than the underlying halo distributions (see, e.g., Harris 1991). However, many of the galaxies upon which this conclusion is based are located in rich clusters, so it is unclear how environment has influenced the presently observed distribution of GCs. In other words, are the GCSs more extended as a result of tidal stripping, or did they form that way? Since tidal limitation by the mean cluster field is likely to occur early on in the evolution of the cluster (Merritt 1984, 1985), we need to consider the relative spatial extents of the halo and GC populations of galaxies at the time of galaxy/cluster formation.

To do so, we must identify a sample of isolated galaxies having well-studied GCSs and accurate surface brightness profiles. We therefore follow Minniti, Meylan, & Kissler-Patig (1996) in defining a “master dE” galaxy using the Local Group dwarf galaxies Fornax, NGC 147, NGC 185, and NGC 205: To characterize the halo light in each galaxy, we adopt the exponential scale lengths and central surface brightnesses given by Caldwell et al. (1992). We then correct each galaxy to a luminosity-weighted mean scale length of $\bar{\alpha} = 375$ pc and add the individual surface brightness profiles to create a composite profile. In converting from angular to linear scales, we adopt the same values for the distance at which the stellar density becomes much smaller than it is in the neighborhood of the orbiting GC, we adopt the effective radius of each galaxy, given by $r_{\text{eff}} = 1.678 \alpha$, (Impy, Bothun, & Malin 1988). We take $v_{\text{typ}}$ to be the mean circular velocity, calculated as $2 \sqrt{GM}/r$, where $\alpha$ is the one-dimensional stellar velocity dispersion in each galaxy (Mateo 1994). Integration of the above equation gives the initial galactocentric distance, $R_i$, of each GC in terms of its current galactocentric distance, $R_f$, as

$$R_f^2 = 0.856 \frac{GM}{v_{\text{typ}}} \ln \frac{\alpha_i}{\alpha} + R_i^2,$$

where $t_i$ is the time since formation. In practice, we calculate $R_i$ for each of the 24 GCs in the master dE using the values of $v_{\text{typ}}$ and $\alpha$ appropriate for each dwarf, and combine the $V$ magnitude of each GC (see De Costa & Mould 1988, and references therein) with our adopted distance estimates to derive individual masses. In doing so, we assume a constant GC mass-to-light ratio of $M/L_V = 2M_\odot/L_V\odot$.

For $b_{\text{max}}$, which denotes the distance at which the stellar density becomes much smaller than it is in the neighborhood of the orbiting GC, we adopt the effective radius of each galaxy, given by $r_{\text{eff}} = 1.678 \alpha$, (Impy, Bothun, & Malin 1988). We take $v_{\text{typ}}$ to be the mean circular velocity, calculated as $2 \sqrt{GM}/r$, where $\alpha$ is the one-dimensional stellar velocity dispersion in each galaxy (Mateo 1994). Integration of the above equation gives the initial galactocentric distance, $R_i$, of each GC in terms of its current galactocentric distance, $R_f$, as

$$R_f^2 = 0.856 \frac{GM}{v_{\text{typ}}} \ln \frac{\alpha_i}{\alpha} + R_i^2,$$

where $t_i$ is the time since formation. In practice, we calculate $R_i$ for each of the 24 GCs in the master dE using the values of $v_{\text{typ}}$ and $\alpha$ appropriate for each dwarf, and combine the $V$ magnitude of each GC (see De Costa & Mould 1988, and references therein) with our adopted distance estimates to derive individual masses. In doing so, we assume a constant GC mass-to-light ratio of $M/L_V = 2M_\odot/L_V\odot$.

---

8 The proximity of NGC 205 to M31 suggests it may be tidally interacting with its parent galaxy. It is included in the definition of the master dE since in § 4.2 and 4.3 we show that past interactions of this sort are likely to have preferentially removed GCs rather than field stars. Including NGC 205 in the sample therefore leads to a conservative estimate of the initial specific frequency profile of the master dE galaxy.

9 Our adopted value of $\alpha = 38.5$ km s$^{-1}$ for NGC 205 represents the average of the mean “core” and “halo” dispersions quoted by Mateo (1994).

10 For GCs IV, VI, VII, and VIII in NGC 185 (Ford, Jacoby, & Jenner 1978), there are no published magnitudes or colors. We have estimated approximate magnitudes for these objects by performing aperture photometry on a V-band image taken with the KPNO 4.0 m telescope, which was kindly provided by Peter Stetson.
Figure 9.—Specific frequency profiles for the GCS of the master dE galaxy. The triangles indicate the present-day specific frequencies exterior to the marked positions. The squares indicate the same profile using the initial GCS distribution (i.e., after correcting the position of each GC for the effects of dynamical friction). If placed in the core of the Virgo cluster, a galaxy of this mass would be tidally stripped to a radius of ~1 kpc. The tidal debris would have $S_n \gtrsim 10$, similar to that observed for the halo of NGC 4486 (McLaughlin et al. 1994). The initial and final GC positions are shown in the top panel by the filled and open symbols, respectively.

4.3. The Radial Profile of Tidally Stripped GCs

Is the spatial distribution of GCs surrounding NGC 4486 consistent with the existence of two chemically distinct GC populations, one that formed along with the body of the galaxy itself, and a second that was tidally stripped from other cluster galaxies? In Figure 10 we show the surface density profile of NGC 4486 GCs taken from Harris (1986). Though not as deep as recent CCD surveys of the NGC 4486 GCS (see, e.g., McLaughlin et al. 1994; Whitmore et al. 1995; Elson & Santiago 1996), this study offers much greater areal coverage and represents the best existing survey of the spatial structure of the GCS at large distances from the galaxy’s center. Note that we have plotted the total counts of unresolved sources in order to investigate the possibility that the GCS surrounding NGC 4486 has a spatial extent that exceeds that of existing wide-field surveys. GC candidates are indicated by the open circles. The filled squares show the arbitrarily shifted surface brightness profile for the galaxy itself. The dashed line represents the $R^{1/4}$ law that best fits the halo light interior to 4.5, the radius that marks the onset of NGC 4486’s cD envelope. If we assume that the population of metal-rich GCs originally
followed the same $R^{1/4}$ profile as the underlying halo light, we can then use the de Vaucouleurs & Nieto (1978) relation to define the radial profiles of both intrinsic and captured GCs, provided we know their relative surface densities at a single position.\footnote{\footnotesize{11} Unless the metal-rich GCs in NGC 4486 are \textit{much} more spatially extended than the underlying halo light, this assumption will have a negligible effect on the derived radial profile of metal-poor GCs in the outermost regions of the galaxy. For instance, if we assume that the metal-rich GCs are twice as extended as the halo stars, the residual GC surface densities are reduced by \leq 6% beyond 15'.} Such an estimate is possible using the deep \textit{Hubble Space Telescope (HST)} survey of Elson & Santiago (1996), who obtained $V$I photometry of 220 GCs in a field located $\approx 2.5$ from the center of NGC 4486.\footnote{\footnotesize{12} We use the Elson \& Santiago (1996) GC photometry instead of that of Whitmore \textit{et al.} (1995) in order to minimize concerns that the metal-rich and metal-poor GC populations have suffered different amounts of dynamical erosion in the innermost regions of the galaxy (i.e., the Elson \& Santiago field is located $\leq 2.5$ from the center of NGC 4486, whereas the field of Whitmore \textit{et al.} extends to within 14' of the galaxy's center).} According to Elson \& Santiago (1996), roughly 44% of the GCs in this field belong to the metal-rich component. Since the Harris (1986) star counts imply a total GC surface density of 26 arcmin$^{-2}$ at $R = 2.5$, we take the surface density profile of metal-rich GCs in NGC 4486 to be log $\Sigma = 1.452 - 3.33[(R/1.6)^{1/4} - 1]$. This profile is shown by the dotted line in Figure 10. The residuals between the measured GC profile of Harris (1986)—which is composed of both metal-poor and metal-rich clusters—and that expected for the metal-rich GCs are plotted as filled triangles.

How does the radial profile of these “residual” GCs compare to that of the dark matter in the Virgo cluster core? To answer this question, we make use of the dynamical study of Cohen \& Ryzhov (1997), who used radial velocities for more than 200 GCs to measure the cumulative mass distribution within $\sim 40$ kpc of NGC 4486. A least-squares fit to the data in their Table 8 yields a relation of the form

$$\log M(r) = a + b \log r$$

(17)

for the total mass, $M(r)$, within a deprojected radius $r$. Assuming a spherically symmetric mass distribution, this implies a density profile, $\rho(r)$, of the form (cf. eq. [16])

$$\rho(r) = \frac{ab}{4\pi} r^{b-3},$$

(18)

where $a = 11.08 \pm 0.08$ and $b = 1.70 \pm 0.15$ (1 $\sigma$ uncertainties). The surface density, $\Sigma(R)$, of a population of objects that follows the above mass distribution is therefore

$$\Sigma(R) \propto R^{-0.30 \pm 0.15}.$$  

(19)

In Figure 10 we plot this radial profile as the solid curve. The associated 1 $\sigma$ confidence limits are indicated by the upper and lower dotted curves.

Despite the facts that the uncertainties in the derived mass distribution are rather large and that the radial profile of the “residual” GCs was calculated under the assumption of negligible background contamination, the agreement between the two profiles is striking. Unfortunately, the photographic survey of Harris (1986) was carried out in a single bandpass, so no information is available on the colors of the GCs at large radii. Clearly, it is important to obtain deeper, multicolor CCD images of the Virgo cluster core having a comparable, or greater, field size. Such observations are now feasible using wide-field mosaic CCD cameras, which are available on many 4 m-class telescopes and would provide a straightforward test of the prediction that appreciable numbers of metal-poor, tidally stripped GCs reside at large distances from NGC 4486. For instance, at a distance of $R = 40'$ we expect a surface density of $\sim 5$ GCs arcmin$^{-2}$. At such distances, captured GCs are expected to greatly outnumber intrinsic GCs, so that the metallicity distribution of the outermost GCs is predicted to be metal-poor and essentially unimodal.

Conversely, the intrinsic population of GCs should dominate near the very center of the galaxy. Based on Figure 10 and on the GC surface densities for the inner regions of NGC 4486 reported by McLaughlin (1995), we estimate that the metal-rich GCs should outnumber the metal-poor ones inside $R \sim 30'$, with the exact ratio depending on the exponent governing the dependence of mass density on radius. Note that only those GCs initially within $R \sim 0.1$ are expected to have had their orbits decay to the galaxy center through dynamical friction (eq. [14]; Lauer \& Kormendy 1986). It is also worth pointing out that equation (19) predicts a surface density for the metal-poor GCs of $15^{+5}_{-4}$ arcmin$^{-2}$ at a galactocentric distance of $R = 15'$; since the total GC surface density at this radius is approximately 45 GCs arcmin$^{-2}$, the implied surface density of intrinsic GCs is $\sim 30$ arcmin$^{-2}$. Thus, the measured specific...
frequency of $S_a = 10 \pm 1$ at this distance (McLaughlin et al. 1994) would correspond to $S_a = 6.5^{+1.2}_{-0.9}$ for the intrinsic GCS, which is more in accord with that found for other Virgo gE galaxies.

5. A COMPARISON TO MERGER-INDUCED GC FORMATION MODELS

There have been several recent reviews of models for the formation of gE's and their associated GCSs (see, e.g., Harris, Pritchet, & McClure 1995; FBG97; Blakeslee et al. 1997). One model that has received particularly close attention—not only in the context of the so-called “high-$S_a$ problem” but also in terms of bimodal GC metallicity distributions—is that of Ashman & Zepf (1992), which contends that numerous GCs are formed during mergers of gas-rich galaxies. The model proposed here also involves mergers but makes very different predictions for the formation and evolution of gE galaxies and their systems of GCs. Therefore, we include below a brief comparison of the present model and that of Ashman & Zepf (1992), focusing on the well-studied pair of Virgo gE galaxies NGC 4472 and NGC 4486.

Perhaps the strongest evidence against the notion that the formation of new GCs in mergers can explain both the bimodal GC metallicity distributions and extraordinarily large GCSs of some gE galaxies is that of Zepf which contends that numerous GCs are formed during mergers of gas-rich galaxies. The model proposed here also involves mergers but makes very different predictions for the formation and evolution of gE galaxies and their systems of GCs. Therefore, we include below a brief comparison of the present model and that of Ashman & Zepf (1992), focusing on the well-studied pair of Virgo gE galaxies NGC 4472 and NGC 4486.

For gE's having specific frequencies greater than $\sim S_a$, we predict that the “excess” GCs should be metal-poor. This prediction is contrary to that of Ashman & Zepf (1992), who argue that the excess GCs should be metal-rich, having formed from chemically enriched gas. Figure 3 of FBG97 demonstrates that, among the dozen or so gE galaxies that are known to have bimodal GC metallicity distributions, the metal-poor GCs dominate the overall GC population in high-$S_a$ galaxies. In addition, an age-metallicity relation is expected for GCs in the Ashman & Zepf (1992) model, in the sense that the metal-rich GCs should be younger than their metal-poor counterparts. Such a trend is not observed for the NGC 4486 GCS (Cohen et al. 1998), nor is one expected within the framework of the model proposed here.

6. ADDITIONAL PREDICTIONS

In addition to our predictions that there should exist appreciable numbers of metal-poor GCs at large distances from NGC 4486, and that the ratio of metal-poor to metal-rich GCs should increase steadily with distance from the galaxy's center (see § 4.3), there are a number of other observations that may be used to test our model for the origin of the metal-poor GCs associated with gE galaxies. First, we expect no dwarf GCs to show bimodal GC metallicity distributions since these galaxies are unlikely to have complex merger histories. In addition, since the metal-rich GCs associated with gE galaxies are assumed to have formed in situ, during a “standard” dissipational collapse scenario, they may show a radial gradient in metallicity and some net rotation, depending on the amount of dissipation in the collapse of the protogalactic gas cloud (see, e.g., Majewski 1993). Based on Figure 6, we also predict a correlation between the location of the metal-poor GC peak and the slope of the initial galaxy LF, in the sense that steeper LFs produce lower metallicities for the captured GCs. However, testing this claim will require large sample sizes since the expected trend will be diluted by differences in the merging histories of individual galaxies. In a companion paper (Marzke, Côté, & West 1998), we discuss the formation of cD galaxies by tidal stripping and mergers and show that a similar trend should exist between initial LF slope and cD envelope color.

The metal-poor GCs surrounding NGC 4486 are predicted to have been tidally stripped from other cluster galaxies by the overall Virgo cluster potential. Figure 10 suggests that their spatial distribution is consistent with the idea that they are tidal debris and that they follow the same surface density profile as does the mass in the Virgo core. Thus, the metal-poor GCs in NGC 4486 are expected to show a larger velocity dispersion than the metal-rich GCs. Radial velocities for the most distant GCs (i.e., those beyond $R \sim 25'$) should have a dispersion comparable to those for the Virgo cluster galaxies. Recently, Kissler-Patig (1997) has presented evidence for such an effect in the Fornax cluster: the distant, and predominantly metal-poor, GCs surrounding the cD galaxy NGC 1399 have a velocity dispersion that is indistinguishable from that of the innermost Fornax galaxies.

Finally, since tidal stripping of GCs is dominated by the overall cluster potential and not by that of the gE galaxy itself, tidally stripped GCs should accumulate in the potential well of any massive cluster, regardless of the whether or not the potential well coincides with the location of a bright cluster galaxy (Merritt 1984; White 1987; West et al. 1995).

7. SUMMARY

Many, perhaps most, gE galaxies have GCSs that show bimodal metallicity distributions. We argue that such bimodal distribution may arise from the capture of other GCs, either through mergers or tidal stripping. For reasonable choices for the LF of galaxies in the host cluster and for the dependence of GC metallicity on parent galaxy luminosity, such captured GCs define a metal-poor population that has a maximum in the range $-1.5 \leq [\text{Fe/H}] \leq -0.7$. Since the precise location of this maximum depends primarily on the LF of the surrounding galaxy cluster, mergers and tidal stripping provide a natural explanation for not only the origin of the metal-poor peak, but also its poor correlation with the luminosity of the gE (FBG97). Our principal conclusion is that it is possible to explain bimodal GC metallicity distributions without resorting to the formation of new GCs in mergers or by invoking multiple bursts of GC formation.

A comparison between the simulated and the observed GC metallicity distributions for the well-studied galaxy
NGC 4472 shows excellent agreement, particularly for steeper LFs. We argue that the weaker cluster tidal field near NGC 4472 compared to NGC 4486 is responsible for the disparity in the sizes of the GCSs of these two galaxies since tidal stripping may, in some cases, lead to substantial increases in specific frequency. This is particularly true if the metal-poor GCs in gE’s were captured through tidal stripping of low- and intermediate-luminosity galaxies since the GCSs of such galaxies are likely to have been more extended than their constituent stars at the epoch of formation. It would seem the time is ripe to reexamine the dynamical evolution of GCSs using improved numerical simulations (i.e., existing N-body experiments suffer from limited resolution and restricted dynamic range; Muzzio 1987). With larger simulations and more sophisticated algorithms, it will be possible to investigate the time evolution of the size, spatial extent and metallicity distribution of the GCSs associated with bright cluster ellipticals.

We describe a new method for placing limits on the fraction of luminous matter that has been captured by individual gE galaxies. An application of the method to NGC 4472 suggests that as much as ~65% of its present-day luminosity could have been acquired in this way. This technique may provide a useful tool for studying the merger histories of gE galaxies.

The authors thank John Blakeslee, Greg Bothun, Doug Johnstone, Jim Hesser, Dean McLaughlin, David Merritt, Sterl Phinney, Hans-Walter Rix, Scott Tremaine, Sidney van den Bergh, and an anonymous referee for helpful comments that improved the paper. P. C. gratefully acknowledges support provided by the Sherman M. Fairchild Foundation. Additional support for this work was provided to R. O. M. by NASA through grant No. HF-0.096.01-97A from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. M. J. W. acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institute for Theoretical Astrophysics.

APPENDIX

TIDAL RADII IN AN NFW POTENTIAL

NFW have suggested that the dark matter distribution in galaxy clusters is given by

$$\rho(r) = \frac{\rho_{\text{crit}} \delta_c}{(r/r_s)(1 + r/r_s)^2}, \quad (A1)$$

where $\rho_{\text{crit}}$ is the critical density, $\delta_c$ is the dark halo’s characteristic overdensity, and $r_s$ is its scale radius.

The tidal radius, $r_t$, of a galaxy moving in a cluster potential $\phi(r)$ is given by (Merritt 1984)

$$r_t = \frac{\alpha \beta \sigma_v}{\sqrt{2}} \left[ \frac{3}{r} \frac{d\phi(r)}{dr} - 4\pi G \rho(r) \right]^{-1/2}, \quad (A2)$$

where $\alpha$ and $\beta$ are parameters related to the internal mass distribution of the test galaxy. Using Poisson’s equation,

$$\frac{1}{r^2} \frac{d}{dr} \left[ r^2 \frac{d\phi(r)}{dr} \right] = 4\pi G \rho(r), \quad (A3)$$

it is possible to calculate the potential gradient

$$\frac{d\phi}{dr} = \frac{4\pi G \rho_{\text{crit}} \delta_c r_s^3}{r^2} \left[ \ln \left(1 + r/r_s\right) + \frac{1}{(1 + r/r_s)} - 1 \right] \quad (A4)$$

corresponding to this density profile. Thus, equation [A2] implies a tidal radius of

$$r_t = \frac{\alpha \beta \sigma_v}{H_0 \sqrt{3 \delta_c}} \left[ \frac{3r_s^2}{r^2} \ln \left(1 + r/r_s\right) + \frac{1}{(1 + r/r_s)} - 1 \right]^{-1} \left[ \frac{1}{(r/r_s)(1 + r/r_s)^2} \right]^{-1}, \quad (A5)$$

where we have made use of the fact that $\rho_{\text{crit}} = 3H_0^2/8\pi G$.

The central velocity dispersion of ~600 km s$^{-1}$ for early-type galaxies in the Virgo cluster corresponds to a halo circular velocity of ~850 km s$^{-1}$. This most closely resembles model 16 of which has a scale radius of $r_s = 190$ kpc and a characteristic overdensity of $\delta_c = 28.200$. Thus, the tidal radius in the vicinity of NGC 4486’s cD envelope is found to be $r_t \sim 1$ kpc, assuming $H_0 = 75$ km s$^{-1}$ Mpc$^{-1}$, $\sigma_v = 30$ km s$^{-1}$, and $\alpha \approx \beta \approx 1$. The agreement with that found from equation (15) is not surprising since NFW have shown that the density profile given by equation (A1) differs significantly from isothermal only at large and small radii.

REFERENCES

Ashman, K. A., & Bird, C. M. 1993, AJ, 106, 2281
Ashman, K. A., & Zepf, S. E. 1992, ApJ, 384, 50
Baugh, C. M., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 27
Bender, R., & Surma, P. 1992, A&A, 258, 250
Bevington, P. R. 1969, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw Hill)

Binggeli, B., Tammann, G. A., & Sandage, A. 1987, AJ, 94, 251
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Biviano, A., Durret, F., Gerbal, D., Le Fevre, O., Lobo, C., Mazure, A., & Slezak, E. 1995, A&A, 297, 610
