SMALL STELLAR SYSTEMS AND GALAXY CORES

THE FORMATION OF GLOBULAR CLUSTER SYSTEMS: HOW, WHEN, AND WHERE?

WILLIAM E. HARRIS
Dept of Physics & Astronomy, McMaster University
Hamilton ON L8S 4M1 Canada

Globular clusters, as fossil remnants of the protogalactic era, provide unique traces of the earliest events of galaxy formation. However, new observations – especially from HST – are showing that massive, globular-like star clusters belong not only to the pregalactic era but can form right up to the present day under the right circumstances. Appropriate interpretation may now let us learn *simultaneously* about the process of cluster formation as well as the nature of the gaseous fragments from which the galaxies were assembled.

Let us first briefly visit some of the relevant observations that have been accumulated for globular cluster systems (GCSs) in a wide range of galaxies, and which seem to constrain the formation parameters the most strongly. These include the cluster metallicity distributions, their specific frequencies (relative numbers), and their distribution by mass.

Metallicity distributions are now available for GCSs in galaxies of all types (though most often in giant ellipticals, which hold by far the largest cluster populations). Most such data come from metallicity-sensitive broadband photometric indices ($C - T_1, B - R$, etc.) that have been calibrated spectroscopically. These results show unambiguously that globular clusters can form at very much the same mass range over more than two orders of magnitude in heavy-element enrichment, from [Fe/H] $\simeq -2.3$ up to above-solar metallicity. In dwarf ellipticals the clusters are almost uniformly metal-poor (Harris 1991; Durrell 1995). The Milky Way clusters, with their well known bimodal metallicity distribution corresponding to halo and bulge clusters (Zinn 1985; Armandroff 1989; Minniti 1995) cover most of this [Fe/H] range, and the distribution in M31 is basically quite similar (e.g. Reed et al. 1994). In the giant ellipticals, the mean abundance shifts to higher levels, with mean values typically <[Fe/H]> $\gtrsim -1$ even far out into their halos. The most extreme case discovered so far is in the giant

R. Bender and R. L. Davies (eds.), New Light on Galaxy Evolution, 87–95.

cD NGC 3311 (Secker et al. 1995), in which a large fraction of the clusters are probably above solar metallicity and the mean is $<[\mathrm{Fe/H}]> \simeq -0.3$. The gE's can also have bimodal or multimodal [Fe/H] distributions that strongly suggest several distinct enrichment phases during formation (e.g. Ashman & Zepf 1992; Zepf et al. 1995a). But many differ widely among themselves, exhibiting a menagerie of unimodal, multimodal, narrow, or wide distributions, with or without radial metallicity gradients (see Ajhar et al. 1994). These large galaxies may be the end results of complex and different formation histories, and it seems that any successful theory for globular cluster formation must be virtually immune to the degree of chemical enrichment or the type of the host galaxy. Additional evidence for this conclusion comes from the observation that one of the young, massive star clusters in NGC 4038/39 has near-solar metallicity (Zepf et al. 1995b).

Specific frequency S_N is the number of globular clusters per unit galaxy luminosity (Harris 1991), and represents the *global cluster formation efficiency* over the lifetime of the galaxy. Figure 1 shows the measured S_N values for E galaxies of all sizes and locations. Clearly, E's and dE's display very much the same range of specific frequencies (except for the occasional cD galaxies with anomalously high S_N; see Harris et al. 1995). This information for the dE's is especially valuable because these small galaxies must have formed in quieter, more isolated environments than the giants. Yet they have been just as efficient, on average, at generating rather normal globular clusters. This suggests, perhaps, that collisions and mergers which build up large galaxies may *not* be especially favorable places for globular cluster formation.

The **Mass Distribution Function** (MDF) is turning out to be a major new clue to the process of globular cluster formation. The MDF is, surprisingly, the most nearly constant phenomenon in the observations (Harris 1991, 1995; Harris & Pudritz 1994 [HP94]). The GCSs in all galaxies have near-identical MDFs, with a number distribution by mass following a power-law form $dN/dM \sim M^{-\gamma}$ for $M \gtrsim 10^5 M_\odot$. In the giant ellipticals, we generally find $\gamma = 1.6 \pm 0.1$, while for spiral galaxies and dE's $\gamma \simeq 1.9$ (HP94; McLaughlin 1994; Durrell et al. 1995; Durrell 1995). An exciting new result which reinforces the idea that the MDF is mostly determined by formation, rather than later dynamical evolution, is the discovery by Whitmore & Schweizer (1995) that the hundreds of recently formed star clusters in the merging NGC 4038/39 galaxy pair follow a MDF of just the same shape ($\gamma = 1.78$), simply shifted to higher luminosity in accord with their much younger ages. In short, the shape of the MDF is *almost* independent of *almost* everything: (a) parent galaxy size and Hubble type, (b) parent galaxy location and environment, (c) metallicity, (d) total cluster population S_N, and (e) even the epoch of formation, if the new HST

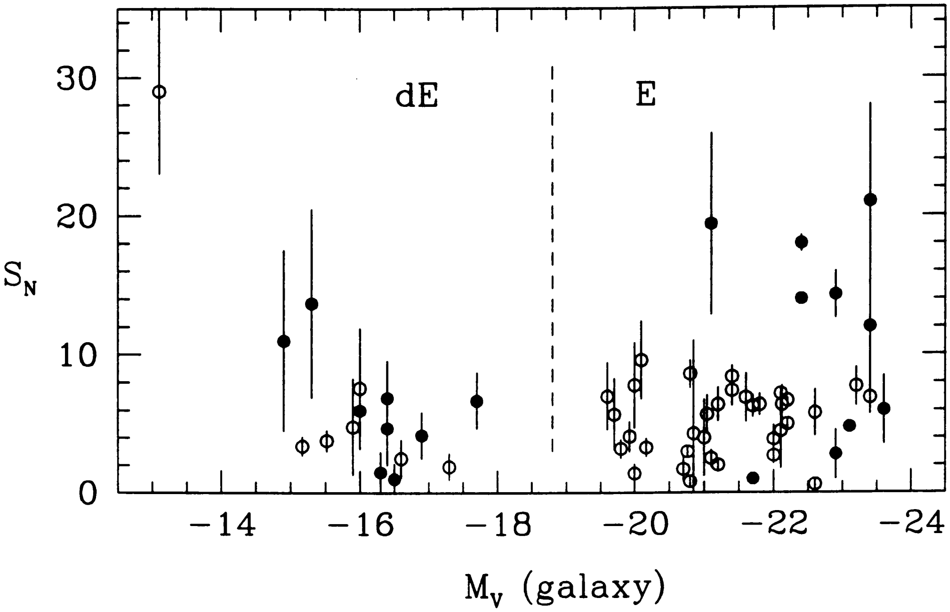


Figure 1. Specific frequencies for globular cluster systems in E and dE galaxies, where $S_N \equiv const \times N_{cl}/L_{V,gal}$. For the dE's, *filled symbols* denote nucleated dE's; while for the giant ellipticals, filled symbols denote central-giant cD galaxies. Most of the information from the dE's is from the new survey of Virgo dwarfs by Durrell (1995).

observations are to be believed.

Can all of this be put into a coherent framework? Our accumulated data add up to the view that globular cluster formation is an extremely robust process, and it seems time to put forward a new set of *observationally based* precepts which will allow us to construct a quantitative model. I suggest that we need the following three:

1. *Globular clusters in the early galaxies formed the same way we see star clusters form today:* that is, out of the rare, very dense gas cores that build up within giant molecular clouds (GMCs).

This view – that globular clusters do *not* come from some highly specialized formation mechanism but that they can form anywhere that gaseous raw material can collect into GMCs – is very close to the ideas that have frequently been expressed by Larson (e.g. 1990a,b), and is made explicit in HP94. As supporting evidence, we can note that virtually every part of the parameter space of cluster age, metallicity, and mass (τ, Z, M) is occupied by star clusters that can be found in *some* galaxy, and there is no compelling reason to distinguish any of these star clusters from one another.

Cluster formation within GMCs is a notably *inefficient process*: only a small fraction of the gas finds its way into the dense cores which eventually transform themselves into bound star clusters. Typically, in Galactic GMCs we observe that the median protocluster core mass is only $\sim 10^{-3}$ of the host GMC mass, and each GMC might form of order ~ 10 or fewer such cores (see HP94).

2. *The typical cluster mass goes up in direct proportion to the mass of the parent GMC.*

This statement implies that we need a *big* reservoir of gas to form star clusters as massive as globular clusters are. Consider, for example, that in the Orion cloud (a fairly typical, nearby GMC) we see a handful of clusters forming at the $10^2 - 10^3 M_\odot$ level. In the LMC (the 30 Doradus region), where more sizable amounts of gas have collected, we see a $\sim 10^4 M_\odot$ cluster forming. And in the centers of the big active galaxies where truly large amounts of gas have been brought together (NGC 1275, the Antennae, etc.), we now see clusters forming at the $10^5 M_\odot$ level and above. If we couple these observations with the 10^{-3} efficiency ratio mentioned above, then we would conclude that the formation of globular clusters of $10^5 - 10^6 M_\odot$ required host GMCs of $10^8 - 10^9 M_\odot$ – i.e., gas clouds that were far larger than any existing in the Milky Way today. HP94 define these as 'supergiant' molecular clouds (SGMCs).

Points (1) and (2) together suggest a rather straightforward answer to an old question: Why doesn't the Milky Way know how to make globular clusters any more? Our reply is that it *does* know how; but it simply does not have enough raw material left in the right form to build anything larger than the relatively tiny objects we call open clusters. The new HST observations demonstrate that any galaxy is capable of constructing massive clusters, as long as we can feed enough low-temperature gas into its potential well.

3. *The mass distribution function (MDF) of globular clusters is primarily a result of their formation process* and not dynamical evolution.

Numerous recent papers (see, e.g., Aguilar et al. 1988; Okazaki & Tosa 1995, for overviews) argue the importance of several dynamical effects on clusters (principally bulge and disk shocking, dynamical friction, and stellar evaporation). For low-mass clusters or ones very near the galactic nucleus, these destructive processes are unquestionably severe. But for most of the cluster mass range (i.e. $\gtrsim 10^5 M_\odot$) and for the majority of the halo ($R_{gc} \gtrsim$ 3 kpc), the MDFs described above are so similar in all galaxies that they seem unlikely to be the products of convergent evolution.

The first stages of a specific theory based on these three postulates are laid out in HP94 and in McLaughlin & Pudritz 1996 [MP96]. In this model, a given host GMC is assumed to contain a large supply of small cloudlets of mass m_0. These cloudlets collide and coalesce to build up larger clouds (and eventually protoclusters) over a range of masses. Once a protocluster gets above some critical point $m_\star$, it can begin star formation; soon after it does, it rapidly ejects its remaining gas by OB star winds and supernovae, thus partially repopulating the reservoir of cloudlets. A statistical equilibrium is thus set up, for which a rate equation can be written down and solved for the resulting distribution function $N(m)$. This agglomeration process is well known to yield a power-law MDF of the right approximate form (see, e.g., Field & Saslaw 1965; Kwan 1979, among others). MP96 generalize it by assuming that the *cloud lifetime against self-disruption* by star formation varies as $\tau_m \sim m^c$, for $m > m_\star$. Higher-mass clouds have lower mean densities (see HP94) and thus undergo relatively more disruptive star formation, so we expect $c < 0$; the actual model solutions (see below) give $c \simeq -0.6$. Now τ_m can be compared with the *collisional growth time*, which is simply $\tau_0 = m_0/\rho\sigma v$ for cloud velocities v, collision cross sections σ, and number densities ρ. The crucial ratio $\beta \equiv \tau_\star/\tau_0$ is what determines the steepness of the resulting mass distribution $N(m)$: the faster the growth time relative to the fiducial disruption time $\tau_\star$, the more clouds can build up to higher masses and the flatter the resulting exponent γ. At the very top end – near $5 \times 10^6 M_\odot$ if $m_\star \simeq 10^5 M_\odot$ – the buildup of still bigger protoclusters becomes so statistically improbable that the MDF finally cuts off quite abruptly.

This model has been used by MP96 to produce the first quantitative theory of the MDF for globular clusters. Model fits to the real galaxies with the best available data (M87, M31, and the Milky Way) are shown in Figure 2. The best-fit ratio β is noticeably higher for M87 than for the spiral galaxies, suggesting that the protocluster clouds built up much more rapidly in the giant proto-elliptical than in the proto-spiral halos.

A traditional problem with any cluster formation model has been to understand how the protoclusters can avoid rapid cooling through radiation from heavy elements or molecular hydrogen (e.g. Fall & Rees 1985; Murray & Lin 1992). That is, if the cloud is supported only by thermal pressure, then we expect $\tau_\star \sim \tau_0$ – essentially the free-fall timescale – and there would be no time for collisional growth to occur. However, GMCs and their embedded cores are strongly nonthermal, and appear to be supported primarily by turbulence and weak magnetic fields in the range of $10 - 100$ microgauss (e.g. Myers & Goodman 1988; Heiles et al. 1993). Thus the relevant quantity governing the cloud lifetime is the ambipolar diffusion time τ_{AD} for the magnetic field to leak out of the cloud, which is an order of magnitude larger than τ_{ff} (see HP94). This is the key factor which allows the clouds enough time to build up to high mass while still gaseous. For the Milky Way halo at $R_{gc} \simeq 8$ kpc, the expected growth time is only a few $\times 10^8$ yr (MP96), which sets a lower limit on the age dispersion among the globular clusters there.

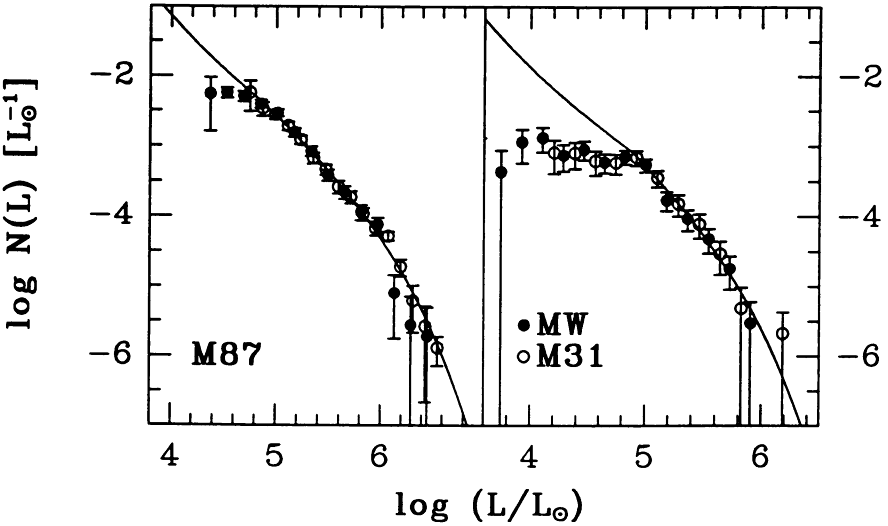


Figure 2. Observation vs. theory for the mass distribution functions of globular clusters in three galaxies, from McLaughlin & Pudritz 1996. Here $N(L)$ is the number of clusters per unit luminosity; the classic 'turnover' point (the maximum number of clusters per unit *magnitude*; see Harris 1991) is at $L_\star \sim 10^5 L_\odot$. The solid line in each panel is the best-fit model for each set of data (see text): for M87, the deduced ratio β (the ratio of cloud lifetime to collisional growth time) is $\simeq 50$, while for M31 and the Milky Way, $\beta \simeq 10$ (a steeper MDF). For $L < L_\star$ (the bottom 10% of the mass range), the collisional growth model predicts too many clusters, though these small objects may have been preferentially depleted by dynamical disruptive processes since their formation; see text.

For purposes of galaxy formation, the utility of this model is that we can use the known characteristics of globular clusters to gauge how large their parent clouds must have been, and thus to get a picture of the raw pieces that built galaxies. As is shown by HP94, the SGMC's need to be $10^8 - 10^9 M_\odot$ and up to $\sim$ 1 kpc in size. It is obvious that the SGMCs can easily be identified with the protogalactic 'fragments' of Searle & Zinn (1978), which they postulated in order to explain the metallicity distribution of the Milky Way halo clusters. Although we invoked the SGMCs for a completely different reason (as host environments for generating the cluster MDF), they have exactly the desired properties for reproducing several other observations too. For example, the MDF will be independent of metallicity as long as magnetic field is the principal support mechanism. The long cloud lifetimes $\tau_{AD} \sim \tau_\star$ also allow each cluster to be well mixed and thus chemically homogeneous by the time it forms stars. In addition, since *both* the cloud lifetime $\tau_\star$ and growth time τ_0 scale with the ambient gas density as $\rho^{-1/2}$ (the free-fall time; see MP96 for derivation), the critical ratio β is independent of galactocentric distance R_{gc}. Thus the MDF should also be independent of position in the halo. In other words, in the outer parts of the protogalaxy everything simply happens more slowly: the clouds grow more gradually because of the lower density, but their self-disruption lifetimes are longer in the same proportion, so the resulting MDF is the same (as long as they are not interrupted by external events such as mergers or tidal breakup).

In summary – to return to the three questions posed in the title – if this picture for cluster formation is valid then we can say that massive (globular) clusters will form anytime and anywhere a galaxy can accumulate gas clouds the size of SGMCs (Searle-Zinn fragments): in protohalos *and* protobulges of large galaxies; in dwarf galaxies; in mergers; and by late infall and accretion. But as a bonus, we have caught a better glimpse of the original building blocks of the galaxies themselves, inside of which the globular clusters assembled.

References

Ajhar, E.A., Blakeslee, J.P., & Tonry, J.L. 1994, *AJ*, **108**, 2087
Armandroff, T.E. 1989, *AJ*, **97**, 375
Ashman, K.M., & Zepf, S.E. 1992, *ApJ*, **384**, 50
Durrell, P.R. 1995, in preparation
Durrell, P.R., McLaughlin, D.E., Harris, W.E., & Hanes, D.A. 1995, *ApJ*, submitted
Fall, S.M., & Rees, M.J. 1985, *ApJ*, **298**, 18
Field, G.B., & Saslaw, W.C. 1965, *ApJ*, **142**, 568
Harris, W.E. 1991, *ARAA*, **29**, 543
Harris, W.E. 1995, in *Stellar Populations*, IAU Symposium No.164, in press
Harris, W.E., Pritchet, C.J., & McClure, R.D. 1994, *ApJ*, **441**, 120
Harris, W.E., & Pudritz, R.E. 1994, *ApJ*, **429**, 177
Heiles, C., Goodman, A.A., McKee, C.F., & Zweibel, E.G. 1993, in *Protostars and Planets*, ed. E.H.Levy & J.I.Lunine (Tucson: University of Arizona Press), p.279
Kwan, J. 1979, *ApJ*, **229**, 567
Larson, R.B. 1990a, *PASP*, **102**, 709
Larson, R.B. 1990b, in *Physical Processes in Fragmentation and Star Formation*, ed. R.Capuzzo-Dolcetta, C.Chiosi, & A.DiFazio (Dordrecht: Kluwer), p.389
McLaughlin, D.E. 1994, *PASP*, **106**, 47
McLaughlin, D.E., & Pudritz, R.E. 1996, *ApJ*, in press
Minniti, D. 1995, *AJ*, **109**, 1663
Murray, S.D., & Lin, D.N.C. 1992, *ApJ*, **400**, 265
Myers, P., & Goodman, A.A. 1988, *ApJ*, **329**, 392
Reed, L.G., Harris, G.L.H., & Harris, W.E. 1994, *AJ*, **107**, 555
Searle, L., & Zinn, R. 1978, *ApJ*, **225**, 357
Secker, J., Geisler, D., McLaughlin, D.E., & Harris, W.E. 1995, *AJ*, **109**, 1019
Whitmore, B.C., & Schweizer, F. 1995, *AJ*, **109**, 960
Zepf, S.E., Ashman, K.M., & Geisler, D. 1995a, *ApJ*, **443**, 570
Zepf, S.E., Carter, D., Sharples, R.M. & Ashman, K.M. 1995b, *ApJ*, **445**, L19
Zinn, R. 1985, *ApJ*, **293**, 424

DISCUSSION

ZEPF: (i) A comment about relating the color distributions of globular clusters to yesterday's discussion of the stellar populations of elliptical galaxies: The broad and often multi-peaked color histograms indicate complex formation histories, in agreement with those suggested yesterday. With deeper images going past the turnover, or spectroscopy, it may be possible to read the formation history of the galaxy in its globular clusters. (ii) How much mass is there in the difference between the predicted and observed LF for the Milky Way and M31 GCS's?

HARRIS: If our simple LF formation model is correct and there were originally many more low-mass clusters, most of them have now been dissolved into the field stellar population. But they would only add up to $\sim 10\%$ of the total mass of the GCLF, and much less than 1% of the whole halo mass. In other words, you can rip up huge numbers of these small clusters without affecting the halo at large.

MEURER: You emphasized cluster formation in merging systems. We have done UV imaging of several starburst galaxies and find that all have UV-luminous clusters. The fact that we see them in mergers may just be a secondary effect resulting from the fact that starbursts occur in merging systems. In addition, although cluster formation may be inefficient relative to the gas content, on average they make up $\sim 20\%$ of the UV flux of starbursts, so they make a significant contribution to the total star formation.

HARRIS: At early stages in a starburst it may look like the clusters or clumps are dominating the light output, but that may be because they form sooner than the field stars. In fact, only 1% or so of the total gas mass may be involved in the cluster phase. Thus what we really need to know for these active regions is, *after* the starburst has finished and everything has died down, what fraction of the mass ends up in bound clusters?

GERHARD: If clusters form from dense cloud cores, their process may require a minimum metallicity to cool the gas down to sufficiently low temperature and high density, something like $Z \simeq 0.01$ or so. For comparison, what is the minimum metallicity you observe in globular clusters?

HARRIS: The lowest observed metallicity for any globular cluster is very near $Z \simeq 0.005 Z_{\odot}$. However, there are various papers in the literature which suggest that nonequilibrium formation of H_2 can cool the clouds rapidly even at low metallicity. Once the magnetic field support has gone, there doesn't seem to be any barrier to rapid cloud collapse.

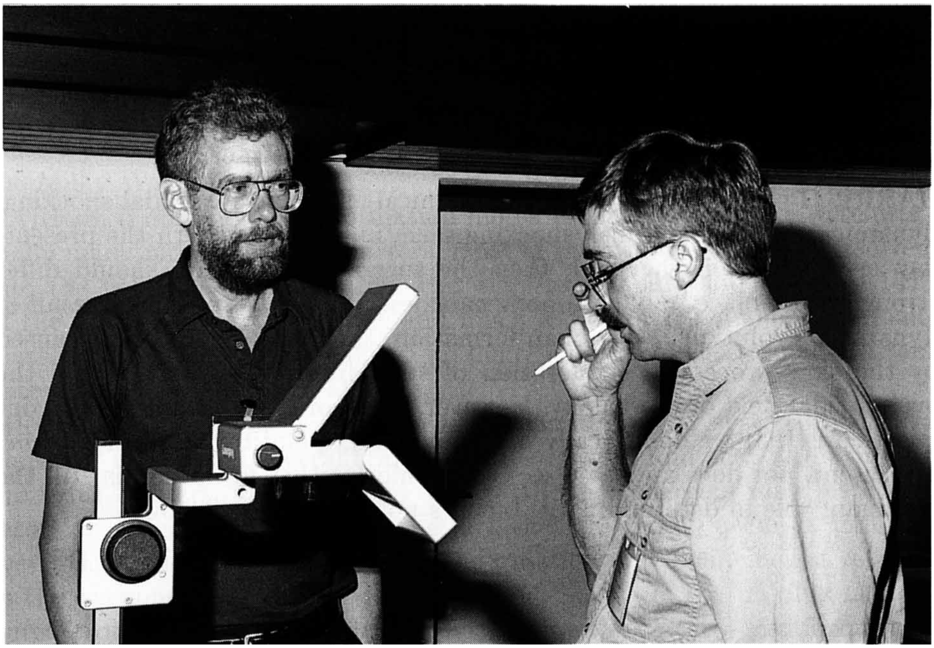
BRINKS: You might be pleased to hear that 'SGMCs' are now seen in *non*-destructively interacting galaxies, such as the 'ocular' galaxy IC2163 which is tidally disturbed by NGC 2207. Apparently, the interaction boosts the velocity dispersion in the neutral gas and $10^8 - 10^9 M_{\odot}$ clumps of neutral gas are formed (Elmegreen et al. 1995, in press).

MINNITI: Does the observed luminosity distribution depend on metallicity?

HARRIS: To first order it doesn't, as far as the available data show. To second order, there may be differences in (say) the turnover luminosity at the quarter-magnitude level in the sense that higher-metallicity systems have slightly fainter turnovers (Ashman et al. 1995, in press).

FABER: Your picture invokes a turnover caused by some kind of destruction effects. Those are likely to vary greatly from galaxy to galaxy and within a galaxy. How does $m_\star$ manage to be so constant?

HARRIS: I completely agree that dynamical effects such as tidal shocking, dynamical friction, and disk shocking seem unlikely to govern the present-day level of $m_\star$; if they did, then the turnover luminosity should differ strongly from place to place among galaxies. So if $m_\star$ actually *is* a result of dynamical evolution rather than formation, then it's probably just caused by the slow process of evaporation of stars from the cluster within the overall tidal field of the galaxy. However, the more correct answer to your question is that we don't really have any good, quantitative theory for $m_\star$ at present, nor does anyone else. Our opinion is that formation is still the dominant role in determining it, but it's not clear yet just how.



Bill Harris and Simon Lilly