arXiv:1010.3022v1 [astro-ph.CO] 14 Oct 2010

Colour-magnitude diagrams, probabilistic synthesis models and the upper mass limit of the initial mass function

Miguel Cerviño[1], Enrique Pérez[1], Nestor Sánchez[1], Carlos Román-Zúñiga[1,2] and David Valls-Gabaud[3]

[1]*Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía, 18008 Granada, Spain*

[2]*Centro Astronómico Hispano Alemán, 18006 Granada, Spain*

[3]*GEPI, CNRS UMR 8111, Observatoire de Paris, 5 Place Jules Janssen, 92195 Meudon, France*

Abstract. We present the underlying relations between colour-magnitude diagrams (CMDs) and synthesis models through the use of stellar luminosity distribution functions. CMDs studies make a direct use of the stellar luminosity distribution function while, in general, synthesis models only use its mean value, even though high-order moments can also be obtained. We show that the mean, high-order moments and integrated luminosity distribution functions of stellar ensembles are related to the stellar luminosity distribution function, within the formalism of probabilistic synthesis models. More details have been yet presented in Cerviño & Luridiana (2006) and references therein. As a direct application of this formalism, we discuss two key issues. First, inferences on the upper mass limit of the initial mass function as a function of the total mass of clusters. Second, we apply extreme value theory to show that that the cluster mass obtained from normalising the IMF between m_{max} and m_{up} does not provide the cluster mass in the case where only one star in this mass range is present, as assumed in the IGIMF theory. It provides instead the cluster mass with a 60% probability to have a star with mass larger than m_{max}, and we argue that in light of this result the basic formulation ofthe IGIMF theory must be revised.

1. From stars to stellar ensembles and the mass-luminosity relation

Our basic knowledge of the Universe stems from the light received from observed sources. In a first-order approximation (neglecting interactions with the interstellar medium and non-stellar components), we can consider two types of sources: individual stars and stellar ensembles. Since in this case the emission of an stellar ensemble is just the sum of its individual components, we can refer to this emission as *integrated* light, coming from an unresolved system such as a distant cluster, or a pixel/slit/IFU in an image of a galaxy. The problem of inferring physical properties from the observed data can thus be reduced to the analysis of the observed light of individual stars in terms of theoretical stellar models, or, in the case of integrated light, to decompose the integrated light into its (stellar) components.

Obviously, the interpretation and physical inferences that can be obtained from the integrated light depends on our physical knowledge of the individual sources that

would be present in the ensemble. It implies to have theoretical models which cover the emission of all possible stars in the ensemble or, at least, to be able to model the emission of the most luminous ones, which are also the more massive in a first order approximation. The reason being that they dominate the integrated emission of stellar ensembles. However, the evolution and spectra of massive/luminous stars are far from being solved problems in astronomy, and constitute a very active research area. The impact of mass loss, rotation, magnetic fields, binary interactions, etc, can change the physical inferences obtained from the observed data (see the contribution by Massey 2010, in these proceedings).

The fact that integrated light is dominated by the most luminous stars could be an advantage for some studies dealing with age inferences. It is, however, a problem if we are interested in inferring the total mass as the statistics of stars present in the field (e.g. Salpeter 1955) show that low-luminosity, low-mass stars are the most numerous ones and dominate the total mass budget of any ensemble, while high-mass stars are the less numerous, with a small contribution to the total mass, but dominate the integrated light.

The problem we are addressing here, then, is the inference about the total mass of the ensemble given its total, integrated luminosity. Homology relations in zero-age main sequence stars allow to relate the current stellar mass $m(t)$ and its bolometric luminosity (which is extrapolated to the luminosity in any band, wavelength or time $\ell_\lambda(t)$ in synthesis codes, see Cerviño & Luridiana 2005, and references therein for a discussion) by a power law as

$$\ell_\lambda(t) \propto m(t)^\gamma \,, \tag{1}$$

where γ has typical value of around 3 for main-sequence stars.

In the case of *unresolved* stellar ensembles we are seeking a mass-luminosity relation with an even simpler functional form:

$$\mathcal{L}_\lambda(t)/\mathcal{M}(t) \propto \text{constant} \,, \tag{2}$$

where $\mathcal{L}_\lambda(t)$ the integrated luminosity in a given band/wavelength of an stellar ensemble whose total mass in stars is $\mathcal{M}(t)$. We stress the difference in the functional dependence of the luminosity in Eq. 1 and Eq. 2. To obtain these simple relations was actually the problem that Beatrice Tinsley aimed to solve (e.g. Tinsley 1980), and it leads to the development of stellar population synthesis models.

2.   Stellar evolution, CMDs and *probabilistic* synthesis models

Stellar evolution theory (in particular evolutionary tracks) predicts the evolution of stars with a given initial mass m as a function of time t measured from the so-called zero-age main sequence. We note this as, for example, the luminosity of a star as a function of time, *given* its mass: $\ell_{\text{track}}(t|m)$. During the main sequence phase, the mass-luminosity relation is more or less well described by Eq. 1, although such a relation does not apply for evolved phases. Provided a large grid of tracks and an interpolation algorithm we can transform the luminosity of a set of stars with given initial masses as a function of time, $\ell_{\text{track}}(t|m)$, into the evolution of stars at a *given* time as a function of their initial masses, $\ell_{\text{iso}}(m|t)$, that is, isochrones (such transformation is not always a trivial task, we refer to, e.g., Cerviño & Luridiana 2005, for details).

At first glance, Colour-Magnitude Diagrams (CMDs) are just a projection of such isochrones in a particular plane given by the choice of filters. Until recently, the comparison of the *shape* of the observed sequences in these diagrams with a set of isochrones provided highly valuable information about physical properties of the ensemble of stars, such as ages, distances and metallicities. The comparison of shapes, however, results in a number of degeneracies, such as the age-metallicity degeneracy in the sense that isochrones of different ages selected at different metallicities have the same *shape*. One has to keep in mind that, besides the observational errors, the location of stars in a CMD is a sequence of *mass*. When a CMD is not just considered as the 2-D locus of stars of a given age and metallicity, but as a density structure which takes into account *how many* stars are in each point of the diagram, the degeneracy can be lifted. The reason is that the density of stars *along* an isochrone is a very sensitive probe of both age and metallicity. So, taken into account the density, not only can the basic parameters of these single stellar populations be inferred (see Hernandez & Valls-Gabaud 2008, for in a classical Bayesian way), but also their star formation and chemical evolution histories (see, for a review, Valls-Gabaud 2011).

A similar formalism for interpreting CMDs taking into account the density is the representation of the stellar luminosity distribution function (the distribution of stars that share a common, given luminosity). Taking the distribution that provides how many stars are born with a given mass m at a given age t, that is, the stellar birth-rate $b(m, t|\mathcal{M})$, any stellar luminosity distribution function or CMD can be produced. The aim of CMD analyses is just the inverse problem, that is, to infer $b(m, t)$ from observations.

It is traditionally assumed (see, e.g., Tinsley 1980) that the stellar birth rate can be decomposed into two *separable* functions: the stellar initial mass function (IMF) $\phi(m)$, and the total amount of gas transformed into stars at a given time, i.e., the star formation rate history $\psi(\mathcal{M}, t) = d\mathcal{M}(t)/dt$:

$$b(m, t|\mathcal{M}) \, \mathrm{d}m \, \mathrm{d}t = \phi(m) \, \mathrm{d}m \times \psi(t|\mathcal{M}) \, \mathrm{d}t \; . \tag{3}$$

This decomposition of the stellar birth rate is the basis of most of the actual research in the studies of stellar ensembles: the use of isochrones make sense only under the assumption that the stellar masses produced in an star formation event do not depend on the amount of gas transformed into stars in the event itself nor in previous events. Formally, this decomposition with small modifications remains valid if an initial cluster mass function, ICMF $\varphi(\mathcal{M}_{\rm clus})$ is included (Weidner et al. 2010):

$$b(m, \mathcal{M}_{\rm clus}, t|\mathcal{M}) \, \mathrm{d}m \, \mathrm{d}\mathcal{M}_{\rm clus} \, \mathrm{d}t = \phi(m) \, \mathrm{d}m \times \varphi(\mathcal{M}_{\rm clus}) \, \mathrm{d}\mathcal{M}_{\rm clus} \times \psi(t|\mathcal{M}_{\rm clus}) \, \mathrm{d}t \; , \tag{4}$$

as far as an integrated galactic initial mass function, IGIMF, which includes the IMF and ICMF functions, can be separated from the star formation history, and the star formation history refers to the amount of gas transformed into clusters, $\mathcal{M}_{\rm clus}$, instead of the amount of gas transformed into stars, $\mathcal{M}$.

In the following we just consider the IMF instead the IGIMF to obtain the stellar luminosity function. Given the IMF and the isochrone at a given time we can obtain the distribution function that gives us the probability of obtain a star with a given luminosity. The implicit assumption is that the IMF is a probability distribution function which provides the probability to have a star with a given mass, but *not* an exact number of stars. The stellar luminosity function $\varphi_\ell \, \mathrm{d}\ell$ is, trivially, (Cerviño & Luridiana 2006)

$$\varphi_\ell = \phi(m) \times \left(\frac{d\ell(m)}{dm}\right)^{-1} , \tag{5}$$

The different inflection points in the mass-luminosity relation will yield features in the luminosity function which do not appear in the mass function. The *mean* of the distribution (i.e. the mean luminosity of clusters which contain just one star, $\langle\mathcal{L}_{1*}\rangle$) is

$$\langle\mathcal{L}_{1*}\rangle = \int_{m_{\text{low}}}^{m_{\text{up}}} \ell(m)\,\phi(m)\left(\frac{d\ell(m)}{dm}\right)^{-1}\,\frac{d\ell(m)}{dm}\,dm = \int_{m_{\text{low}}}^{m_{\text{up}}} \ell(m)\,\phi(m)\,dm , \tag{6}$$

which is the equation solved by synthesis codes.

Equation 6 shows two important issues that must be considered with caution. The first one is related with the derivative of ℓ with the initial mass m. Such derivative is not always defined and isochrones have discontinuities in post-main sequence evolutionary phases, the so called *fast evolutionary phases* where Eq. 1 does not apply. Actually such phases are better described in terms of the lifetime of the phase for a given mass m (typically the turn-off mass) rather than with initial masses. This issue leads to different algorithms to compute synthesis models, either using *fuel consumption theorem* or else isochrone synthesis (see Buzzoni 1989; Marigo & Girardi 2001, for more details). The important point here is that since it is the stellar luminosity distribution function the one that must be well sampled, size-of-sample effects may be still present in the sampling of fast evolutionary phases, *even with a well-sampled IMF*.

The second cautionary note refers to the IMF mass limits themselves. The formulation and the modelisation of stellar ensembles is valid as far as m_{up} (and formally m_{low}) have well-defined values. The mean value depends on both parameters and the functional form of the stellar luminosity distribution function.

When dealing with an ensemble of N stars the corresponding distribution which describes the possible integrated luminosities of the ensemble is N times the self-convolution of the stellar luminosity function (under random sampling of the stellar luminosity distribution function assumption). Then the mean total luminosity is (see Cerviño & Luridiana 2006, for details):

$$\langle\mathcal{L}_{N*}\rangle = N \times \langle\mathcal{L}_{1*}\rangle = N \times \langle\ell\rangle . \tag{7}$$

We stress again that, as far as (*i*) the stellar luminosity distribution function must be sampled and (*ii*) its high luminosity tail is related to the lifetime of particular evolutionary phases, a random sampling hypothesis (for the stellar luminosity function) is still valid even under a sorted sampled IMF hypothesis.

This formalism can equally be applied to the total mass of the ensemble: the total mass distribution of an ensemble of N stars is described by N self-convolutions of the IMF itself (Selman & Melnick 2008) and the mean value of the total mass of the ensemble is N times the mean mass of the IMF:

$$\langle\mathcal{M}_{N*}\rangle = N \times \langle\mathcal{M}_{1*}\rangle = N \times \langle m\rangle . \tag{8}$$

Remarquably, this implies that

$$\frac{\langle\mathcal{L}_{N*}\rangle}{\langle\mathcal{M}_{N*}\rangle} = \frac{N \times \langle\mathcal{L}_{1*}\rangle}{N \times \langle\mathcal{M}_{1*}\rangle} = \frac{\langle\mathcal{L}_{1*}\rangle}{\langle\mathcal{M}_{1*}\rangle} = \text{constant} \tag{9}$$

and hence validates the approach given by Eq. 2. Note that this relation provides a valid (averaged) total mass as far as the observed luminosity is a good proxy of the mean luminosity of stellar ensembles with *N* stars, and the age is known.

3. Size-of-sample issues

We have shown that synthesis models results are related to CMD analyses through a collapse of the underlying distribution that defines the CMD in its moments. Obviously, the mean value of the distribution is not the only possible outcome of synthesis codes, and high-order moments or even the distributions themselves can be obtained (Cerviño & Luridiana 2006). Working with distribution moments has the advantage of providing a first order approximation of the underling distribution. From the Central Limit theorem, we know that an infinite self-convolution of a distribution with finite limits tends to the Gaussian form. High-order moments allows us to define how many stars (and average mass) are needed to obtain Gaussian or quasi-Gaussian integrated luminosity distributions (see Cerviño & Luridiana 2006, for details). This is the basic requirement for χ^2 fitting algorithms[1]. In particular the second moment of the stellar luminosity distribution function (the variance) also allows one to evaluate the relative error in the use of an observation as a proxy of the mean of the underling distribution. However the situation becomes more complicated when a *small* number of stars defines the integrated luminosity[2].

In general three cases can be defined: (*a*) Systems/observations which contain a large number of stars yielding gaussian distributions of the theoretical integrated luminosity. In this case the observations can be used as proxy of mean values safely. These situations typically correspond to galaxies and systems with masses larger than 10^8 M$_\odot$ when all the light from the system is covered by the IFU/slit. (*b*) Systems/observations with a large number of stars to ensure gaussian-like distributions, but not large enough to provide a negligible variance. In these cases both the mean value and the variance of the underling distribution for each band/wavelength *must* be considered. This situation corresponds to systems typically more massive than 10^5 M$_\odot$ although the exact value depends on the considered band/wavelength. (*c*) *Extreme size-of-sample effects*: these are systems/observations with a number of stars so small that the distributions of the integrated luminosities are not Gaussians. In these cases, the analysis of observed data in terms of the mean and high-order moments of the distribution of the integrated luminosity is *not* suitable since they do not provide a proper information on the distribution (mean and mode differs, variance can be not translated in confidence intervals etc.). In this situation, it is more accurate to work directly with integrated luminosity distribution functions which can be obtained by self-convolution of the stellar luminosity distribution function (which has several problems at computational level except for a low number of stars) or by the analysis of Monte Carlo simulations (which are more economic at the computational level but which lead to some subtle problems, such as

---

[1] Although χ^2 seems the simplest tool to make inferences from observed data (the inverse problem) we emphasize that there are many different methodologies (as well as interpretations) on how to deal with inverse problems. For a simple introduction see Tarantola (2006) as well as other articles at `http://www.ipgp.fr/~tarantola`.

[2] Note that *small* is not just an intrinsic property of the studied system, but also an observational effect: an IFU observation intrinsically contains a smaller number of stars than a large aperture observation.

increasing the number of simulations when the number of stars in the simulated clusters decreases to provide a correct sampling of the underling distributions).

4.   Low Luminosity Limit (LLL), $\mathcal{M}_\mathrm{clus}$ and m_up

Unfortunately we have no estimation of the total mass of the system before an analysis is performed, and simple recipes are useful to know in which of the previous régimes the observational data lie. A simple method, called the Low Luminosity Limit (LLL), was proposed by Cerviño & Luridiana (2004). The method just compares the luminosity of the observed ensemble with the luminosity of the most luminous star assumed in the model, that is, the extreme value of the stellar luminosity distribution function (the LLL). It provides the lowest luminosity that an stellar ensemble should have without been possibly confused with an single star.

An additional simple test is just the comparison of the observed light from a stellar ensemble with the locus of the mean values obtained by synthesis models and the locus of individual stars in color-color diagrams (see Barker et al. 2008, as an example): clusters affected by strong sampling effects (where the mean values are not representative) cover the intermediate area between stars and the mean values obtained by synthesis models.

A similar situation is present while making inferences of total masses $\mathcal{M}_\mathrm{clus}$ for studies of the maximum mass m_max that a cluster with total mass $\mathcal{M}_\mathrm{clus}$ would contain with $m_\mathrm{max} - \mathcal{M}_\mathrm{clus}$ diagrams (Selman & Melnick 2008; Weidner et al. 2010, and references therein). The observed maximum stellar mass can be obtained from observations. However, $\mathcal{M}_\mathrm{clus}$ must be inferred under a situation where we know for sure that the IMF is not well sampled. In these cases, one has to use all observational constrains (such as the number of stars with different masses), run a large enough set of Monte Carlo simulations which fulfills the observational constrains, and obtain the distribution of possible total masses. Note that the usual method of correcting the total mass in the unobserved mass range with the mass obtained from a truncated IMF from m_low to the minimum observed mass $m_\mathrm{min}^\mathrm{obs}$ implicitly assumes that the IMF is fully sampled in the low mass interval (which is equivalent to say that the IMF is sorted sampled up to $m_\mathrm{min}^\mathrm{obs}$).

Finally, we address the problem of the inference of m_up from observations. The probability of an extreme value, say m_up, is precisely the subject of *Extreme Value Theory* (EVT). This branch of statistics deals with extreme deviations (maxima or minima) from the median and it has important consequences in every day life (such as economic crashes) and natural catastrophes produced by deviations in annual flood flows, precipitation maxima or earthquakes (also with human-life and economic implications). The classical reference in the subject is Gumbel (1958), although we also suggest reading Sornette (2004) (especially pp. 18-23), where the case $m_\mathrm{up} = \infty$ is discussed in detail, and where results obtained in astronomy (e.g. van Albada (1968); Oey & Clarke (2005); Pflamm-Altenburg & Kroupa (2008)) can be also be found. A detailed discussion is presented in a forthcoming paper.

Within this EVT framework, we analyse the basic assumption made in the IGIMF theory proposed by Weidner et al. (2010) and previous works. Let us consider a sample of N stars. An event with probability p occurs typically $N \times p$ times, hence if we *expect* just 1 star in the mass range $m_\mathrm{max} - m_\mathrm{up}$ in the cluster we have:

$$p = \int_{m_{max}}^{m_{up}} \phi(m)\,dm; \quad Np = 1; \quad p = \frac{1}{N} \qquad (10)$$

and we can obtain the total number of stars and make an estimation of the total mass of such a cluster, $\langle \mathcal{M}_{clus} \rangle = N \times \langle m \rangle$, (provided that we are in a régime where $\langle m \rangle$ can be safely used! but see Selman & Melnick 2008).

We also know from EVT the PDF of the *extreme* values for N stars, so, we can ask what $1/N$ really means. Following Sornette (2004) the estimation of m_{max} from $1/N$ is in fact the stellar mass that is not exceeded with a probability $p = 1/6 = 0.37$ for a cluster with N stars. It means that $(1 - p)$=63% of clusters with N stars (and hence $\langle \mathcal{M}_{clus} \rangle = N \times \langle m \rangle$) have a star with mass *larger* than m_{max}. In other words, the estimation of the relation between m_{max} and $\mathcal{M}_{clus}$ as assumed in the IGIMF theory may not be correct in 63% of real cases.

5. Conclusions

In this contribution we have shown that the physical inferences obtained from CMDs studies are related with the results of evolutionary synthesis codes by the stellar luminosity distribution function. Since CMDs basically provide information about the distribution itself and the synthesis models typically just provide the mean of such distribution, a direct conclusion is that a CMD provides *always* a more complete information about star formation than the use of mean values obtained by synthesis models.

The mean of the stellar luminosity distribution function is defined by its functional form as well as its limits (the upper one is non-trivially related to m_{up}). This mean as well as high-order moments are directly proportional to the mean integrated luminosity of ensembles of N stars. Such proportionality provides mass-luminosity relations for systems/observations with a large number of stars (such that the average total masses are larger than 10^8 M$_\odot$) where the observed luminosity can be safely used as a proxy of the theoretical mean integrated luminosity. However, the use of one observation as a proxy of the mean of the theoretical distribution of integrated luminosities must be used with caution for less massive systems: it must be used together with the variance for ensembles with mean mass values in the range 10^5 to 10^8 M$_\odot$, and simply *not used* for ensembles with mean total masses smaller than 10^5 M$_\odot$. For these small (by number) systems ones has to use the distribution of integrated luminosities rather the parametric descriptions of the distribution.

The reason of these limitations is the incomplete sampling (or size-of-sample effects) of the stellar luminosity distribution function in the ensemble. Since the stellar luminosity distribution function contains a (low-luminosity) component related to the IMF, and a (high luminosity) component related to the short lifetime of fast evolutionary phases, an intrinsic scatter due to random sampling is always present even in cases of a perfect or sorted sampling of the IMF. More details about these limitations can be found in Cerviño & Luridiana (2004, 2005, 2006) and references therein.

One of the difficulties related with size-of-sample effects is that the mean values of the distributions are not representative of the distributions themselves. Or, in a practical way, observational data cannot be compared with the mean values provided by synthesis model to make physical inferences. This situation is especially relevant in the computation of total stellar masses of under-sampled clusters like the ones inferred to relate the total mass of a cluster with the mass of its more massive star. In these cases,

one has to use all observational constraints (such as the number of stars with different masses), run a large enough set of Monte Carlo simulations which fulfills observational constrains, and obtain the distribution of possible total masses. Any other methodology (such as to complete the unobserved mass range up to a given mass with a mean value of a truncated IMF) introduce biased results (just like an artificial sorted sampling).

Finally, we show that the basic assumption implicit in the formalism of the IGIMF, that is, that when the integral of the IMF over some given interval ($m^*(\mathcal{M}_{\rm clus})$ to $m_{\rm up}$) is normalised to unity this provides the cluster mass $\mathcal{M}_{\rm clus}$ that contains just one star in that interval (and that this individual star has a mass equal to $m^*(\mathcal{M}_{\rm clus})$) is not correct. Extreme value theory shows that such an assumption just provides a $\mathcal{M}_{\rm clus}$ value which has a 63% of probability to have a star with a mass *larger* than $m^*(\mathcal{M}_{\rm clus})$. In this situation some of the basis of the formulation of the IGIMF theory must be revised.

References

Barker, S., de Grijs, R., & Cerviño, M. 2008, A&A, 484, 711
Buzzoni, A. 1989, Ap.J.S.S., 71, 817
Cerviño, M., & Luridiana, V. 2004, A&A, 413, 145
— 2005, in Resolved Stellar Populations, edited by . M. C. D. Valls-Gabaud (ASP Conf. Ser.), in press. `arXiv:astro-ph/0510411`
— 2006, A&A, 451, 475
Gumbel, E. J. . 1958, Statistics of Extremes (New York: Columbia University Press)
Hernandez, X., & Valls-Gabaud, D. 2008, MNRAS, 383, 1603
Marigo, P., & Girardi, L. 2001, A&A, 377, 132
Massey, P. 2010, in UP: Have Observations Revealed a Variable Upper End of the Initial Mass Function? (ASP Conf. Ser.), in press. `arXiv:astro-ph/1008.1014`
Oey, M., & Clarke, C. 2005, ApJ, 620, L43
Pflamm-Altenburg, J., & Kroupa, P. 2008, Nature, 455, 641
Salpeter, E. E. 1955, ApJ, 121, 161
Selman, F., & Melnick, J. 2008, ApJ, 689, 816
Sornette, D. 2004, Critical phenomena in Natural Sciences (Berlin: Springer series in Synergetics)
Tarantola, A. 2006, Nature Physics, 2, 492
Tinsley, B. 1980, Fundamentals of Cosmic Physics, 5, 287
Valls-Gabaud, D. 2011, in Local Group Cosmology, Proceedings of the XX Canary islands Winter School, edited by D. Martinez-Delgado (Cambridge: Cambridge University Press)
van Albada, T. S. 1968, Bulletin of the Astronomical Institutes of the Netherlands, 20, 57
Weidner, C., Kroupa, P., & Bonnell, I. A. D. 2010, M.N.R.A.S., 401, 275