Student Projects on *How our Milky Way was assembled*

1. Framework: Dissecting our Milky Way to understand how it was put together

Our Milky Way, by our position within it, is the only big galaxy that we can fully resolve into stars, and where we can measure the 3D positions (and 3D velocities) for a substantive fraction of all stars, which implies that we can construct their orbits. In the Milky Way, we can also measure the elemental abundances of vast number of individual stars; this 'degree of nuclear waste pollution' gives us a hint when and where the stars were born. For some stars, we can get directly their ages.

Taken together, we can in principle and soon in practice, get nearly complete information about the stellar body of our Galaxy. An that basis we can make very direct inferences about how our Milky Way was put together: e.g. are the old stars more in the inner part of the stellar disk, while the younger stars are at larger radii? If so, it would imply quite directly that our Milky Way's disk grew 'inside out'? Or: did most stars form on near-circular orbits in a thin disk, and then get increasingly perturbed (naively, more eccentric orbits and a thicker disk) as time goes on? Do the stars stay (even approximately) at the radii at which they were born? At the same time, the stars' motions can probe the gravitational potential in which they orbit, which could provide us e.g. with very direct constraints on the *local* amount of dark matter.

Through a sequence of large spectroscopic and imaging surveys (called, e.g. SDSS, RAVE and PanSTARRS1), which will culminate in the Gaia space mission to be launched next year, we are embarking on a Galactic 'Genom Project', establishing our Milky Way as a model organism for galaxies. Fortunately, our Milky Way is as typical as it gets in the present-day world of galaxies, implying that we should be able to generalize results that we find 'at home'.

2. Surveys: an exciting and intimidating flood of information

A variety of technological breakthroughs (and the willingness to spend money) have made it possible to get basic information on stars (positions, velocities, chem. abundances, ages) for vastly larger sample than in the past.

Photometric Surveys now cover the much of the sky, providing not only precise positions, but als 'proper motions' of the apparent sky position (to about 2 mas/yr). The precise colors provide an effective temperature determinations. For stars on the main sequence (most of them) the color determines the luminosity, which – in conjunction with the observed flux – provides an estimate of the distance. Between the positions, the distance estimates and the proper motions we have (albeit imprecise) constraints on 5 dimensions of the 6D position-velocity space. MPIA (HW Rix) is currently leading the Milky Way mapping of the most ambitious of these surveys, called PanSTARRS1.

Spectroscopic Surveys provide three important additional pieces of information: a line-of-sight (Doppler) velocity, an estimate of the stars absolute luminosity (through the surface gravity and its effective temperature), and the abundances of the 'heavy' elements. The SDSS and RAVE surveys have by now gotten over half a million good spectra. And the VLT whas just started a 300 night survey (HWR, co-I) to get spectra of unprecedented quality.

Gaia, ESA's next big space mission, will take this to a dramatically new level: it will provide motions for billions of stars and direct and precise distances through parallax measurements. The data will only become available in 2015+ , but now is the time to think through 'what exactly could we do, if we had the data?'

3. Projects on Milky Way Mapping:

In this context, there is a large range of research projects that can and should be done. They revolve around two practical questions: 1) how do I find sub-set of stars that are of particular interest; and 2) how and how well can I convert the observables into the physical quantities of interest? [e.g. 'age' is never a direct observable.]

What is the 'character' of such project, a few of them I'll spell out below? Basically those surveys produce vast 'catalogs' that have set of observables (and their errors) for long lists of stars. The task is then to use a combination of existing software to query those catalogs and then use simple programs (20 - 200 lines) to select interesting sub-sets of objects and calculate physical quantities for them. In practice this provides a straightforward introduction to coding/computing in a context where the results matter (because there are interesting and at the frontiers of research); languages are python (my preferred one), IDL, C or any language of your choosing. In the end, it boils down to reading in information, carrying out calculations on those data, and making cool (i.e. informative) plots. This can introduce you to two concepts that are of increasing importance in research: data mining and 'proper' treatment of 'errors' and probabilities; the latter means concepts like Monte-Carlo simulations, Markov-Chains and Bayesian statistics, all concepts that are surprisingly simple, very powerful and rarely taught before the masters level. All projects scale easily: there are simple initial goals, and *much* more to do (until it's time to write up). It is unlikely (though not impossible) that a Bachelor project all become a publication in itself, but unite likely that it will be part of a publication (with co-authorship). The work will be with HW Rix, but augmented direct collaboration with post-docs and PhD students in the group.

3.1. Project 1: For how many stars will GAIA get 'good ages'?

As stars are equilibrium systems, it is not easy to get their ages. E.g. stars like the Sun will sit on the main sequence for nearly 10 billion years, changing their properties only slightly. Yet, for reconstructing the formation history of our galaxy, age is as elementary a parameter as it gets. Fortunately, there are particular stellar phases, where the age can be determined well. This is illustrated in he attached Figure. It shows a simple stellar population in a star cluster (in the Magellanic Cloud in this case). I.e. all stars have the same age and chemical composition, but they differ in their mass. Of particular interest are the stars on the so-called sub-giant branch (see Figure caption). Why? We know that they have just turned off the main sequence and that they were MS stars of the same luminosity just a short while before. If we have (from GAIA) precise parallax distances to those stars, we have their precise luminosities. For a given (known) metallicity, we however know exactly the main sequence of such stars. So, **if** we know a star is on the sub-giant branch, and if we have its luminosity, we get its age.

Goal of the project is to estimate for how many stars and over what volume GAIA can do that? Steps: Figure out which fraction of stars of a population of age t is on the sub-giant branch (this is on the basis of existing 'isochrones'). How much does the sub-giant branch luminosity vary with the age of the stars? This tells us how a luminosity error translates into an age error. Figure out to what distance GAIA can get good enough parallax distances (this is a function of the distance and of the luminosity of the star). Turn all of this into a statement: We can expect to get ages to 10% for X stars in the age range Y within a volume Z. Then we can evaluate how useful the measurements will be do get the age structure of the Milky Way.

3.2. Project 2: How well can we identify 'red clump stars' in the PanSTARRS survey.

Red clump stars (see Figure) are very useful: they are bright and have a very well-defined absolute luminosity. 'Bright' means we can get good spectra to great distances; the well known absolute luminosity means we get a good (photometric) distance from their apparent magnitude. This makes them an excellent way to trace the Milky Way's rotation curve from their radial velocities.

The question (and the project) revelves around how (and how reliably) to find them: The first idea would be to pick them by color ($m_{555} - m_{814} = 0.95$). However, in the Milky Way we don't know the distance of the stars (before Gaia in 2015), so a star of $m_{555} - m_{814} = 0.95$ could in principle also main sequence stars that is 6 magnitude fainter in absolute luminosity. It would have to be 15-times closer to appear equally bright. Such very nearby stars should have far greater angular motion (proper motions). The idea is to find the stars of the right color, but with small angular motions (because they are more luminous and hence more distant). This approach has been tried in other context, but not in the context of the PanSTARRS1 data, a far larger survey than undertaken before. With it, we could identify enormous samples of such stars.

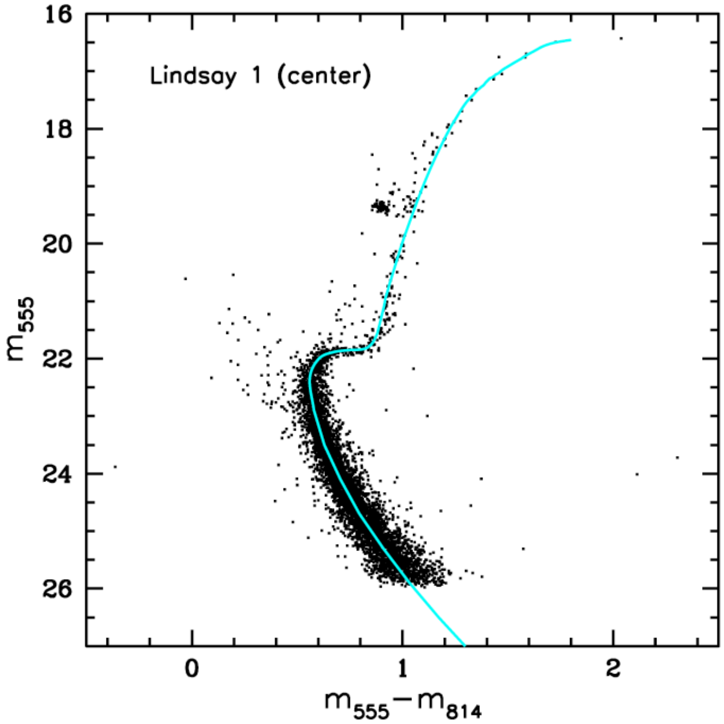

Fig. 1.— 'Color-magnitude diagram' of stars in a star cluster in the Magellanic cloud. All stars in the cluster have the same age and chemical composition, but they differ in their stellar masses. Color is on the X-axis, with redder/cooler stars on the right; flux (=luminosity at a given distance) is on the Y-axis, with more luminous stars at the top. Most stars are on the main sequence, fainter than ,$m_{555} > 22$; stars more luminous (and more massive) than the 'main sequence turn-off' have already evolved off the main sequence, first on the 'sub-giant branch' ($m_{555} = 21.5$ and $0.7 < m_{555} - m_{814} < 1$) and then up the giant branch (extending to $m_{555} = 17$ and $m_{555} - m_{814} = 1.4$, and then to the 'red clump' (the concentration of stars at $m_{555} = 19$ and $m_{555} - m_{814} = 0.95$). The cluster has an age of $\sim$ 1Gyr, an age at which the 'red clump' is prominent.