Measuring Age and Distance of Globular Clusters

Our galaxy, the *Milky Way,* is surrounded by perhaps 200 hundred *Globular Clusters*[1] *(*we can see about 150 and we expect another 30-50 to be hidden from us by Milky Way dust). They are bound to the Milky Way by gravity and so in orbit about it, but although they must sometimes pass through the galactic disk they do not, on the whole, appear to exchange stars with the Milky Way.

From Earth they do not appear to be evenly distributed on the sky: there are distinctly more in the direction of to constellation Sagittarius, and this was the first intimation that the Sun was positioned well away from the centre of the Milky Way. It was suggested that they were evenly distributed throughout a sphere centred on our galaxy, whose centre would then be somewhere in the direction of Sagittarius. This, in fact led to the first reasonable estimate of the size of the Milky Way and the relative position of our Sun, by Harlow Shapley in the early part of the 20th Century. His first estimate was not particularly accurate because, as we shall see, measuring the distances to globular clusters was a considerable challenge. We now think that the diameter of this these sphere is about 100,000 lighter years, with the Sun about 30,000 light years form the galactic centre. It turns out that the closest clusters we see are only about 10,000 light years away, and the furthest almost ten times that distance.

A typical cluster may be few hundred light years in diameter, and since a cluster is likely to be tens of thousands of light years away, we can think of all the stars in the cluster as all being at effectively the same distance. That means, very importantly, that the *relative* difference in the observed brightness of stars in the cluster is also a measure of the *intrinsic* difference in their luminosity. (A star that looks twice as bright as another in the same cluster will really be twice as luminous.)

We need a distance measurement to work out the absolute luminosity of stars in the cluster. This type of measurement is usually a difficult and time consuming business with globular clusters - if possible at all. It turns out, however, that we can do some important astrophysics just by using the relative brightnesses. It will also turn out that if we can measure the distance to just one globular cluster we can get a reasonable estimate of the distance to many other clusters - even if they are orbiting other galaxies.

Each globular cluster looks rather similar to the image on the right of the object known as "**Messier 80**" (or more normally as just **M80**[2] - the 80th object in Messier's catalogue of "bright nebulae") - showing a more or less spherical distribution of stars tightly concentrated towards the centre. A typical cluster probably has 100,000 stars, but there is a big range of sizes and a large cluster, such as NGC104 (also known as *47 Tucanae*) may have more than a million stars. Stars in a typical globular clusters are actually physically very close to each other: if our Solar System were in the cluster above, other stars would regularly pass through the edges of

---

[1] https://en.wikipedia.org/wiki/Globular_cluster

[2] See Appendix C for some notes on conventions for identifying astronomical objects.

the planetary system. (This is one reason why we think that planetary systems are probably fairly rare in globular clusters - they would just get stirred up too much.)

In this image we can clearly see that the stars have different brightnesses and different colours (with some of the brightest stars looking distinctly reddish). There are many more dim stars than bright stars, which is very much in accord with our understanding of the way stars are born. (We have good reasons for thinking that the clouds that form stars tend to break up in ways that are likely to form many more small stars than large stars.)

We also have good reasons for believing that, in most cases, all the stars in a cluster were born at roughly the same time, in a single burst of star formation that used all the available gas. (There is certainly no gas left today.) We do, however, have to beware of exceptions: a few clusters associated with the Large Magellanic Cloud (a small nearby galaxy gravitationally bound to our own) seem to have had more than one phase of star formation, which is believed to be because they could pull gas out of the Cloud during close passages.

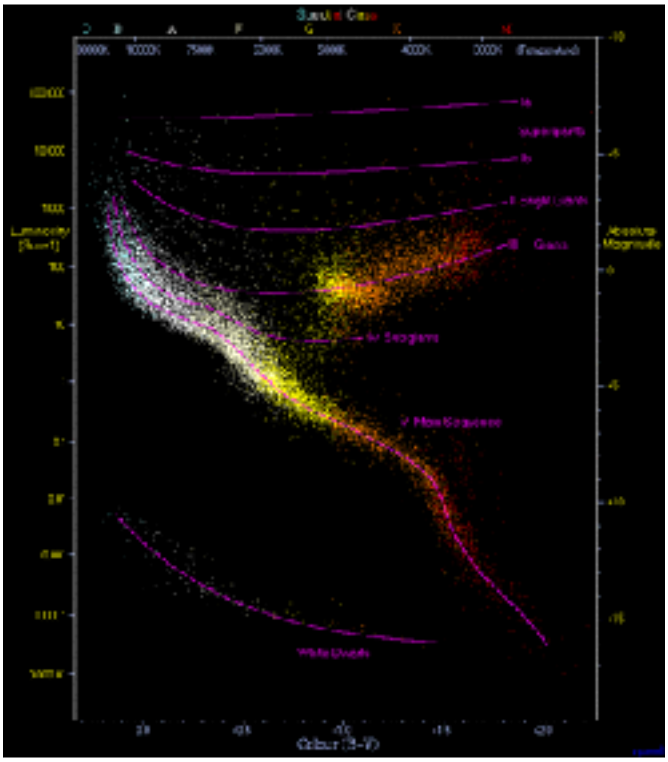
All this makes globular clusters extremely interesting for many astrophysical reasons, but principally because they are a snapshot of a group of stars of different masses born at about the same time. What we see today is the way stars of different masses have evolved differently over the subsequent billions of years.

Better still, although most globular clusters are billions of years old they did not all form at the same time as the Milky Way. Some clusters seem to be older than others. That means that we have several snapshots of stellar evolution which we can arrange in a time sequence. We think that the age of the oldest globular clusters puts a lower limit on the age of the Milky Way - the clusters could not have formed around the Milky Way unless there *was* a Milky Way. We will see in later how astronomers estimate cluster ages.

These type of observations and brilliant theoretical work by Sir Arthur Eddington started to reveal a good deal about ways stars evolve in time, especially after it was realised that the fusion of hydrogen into helium is the fundamental energy source. Much later, in the 1950s, it became possible to use computers to calculate stellar evolution in fine detail and in particular to make reliable estimates of the amount of time it takes a star to use up its fuel. We now believe we have a very good understanding of the lifecycle of typical stars over a wide range of masses and over most of their lifetimes. (There are, however, processes at the very start and very end of the stellar lifetime that get very complicated and still need more work.)

This is a brief summary of relevant points:

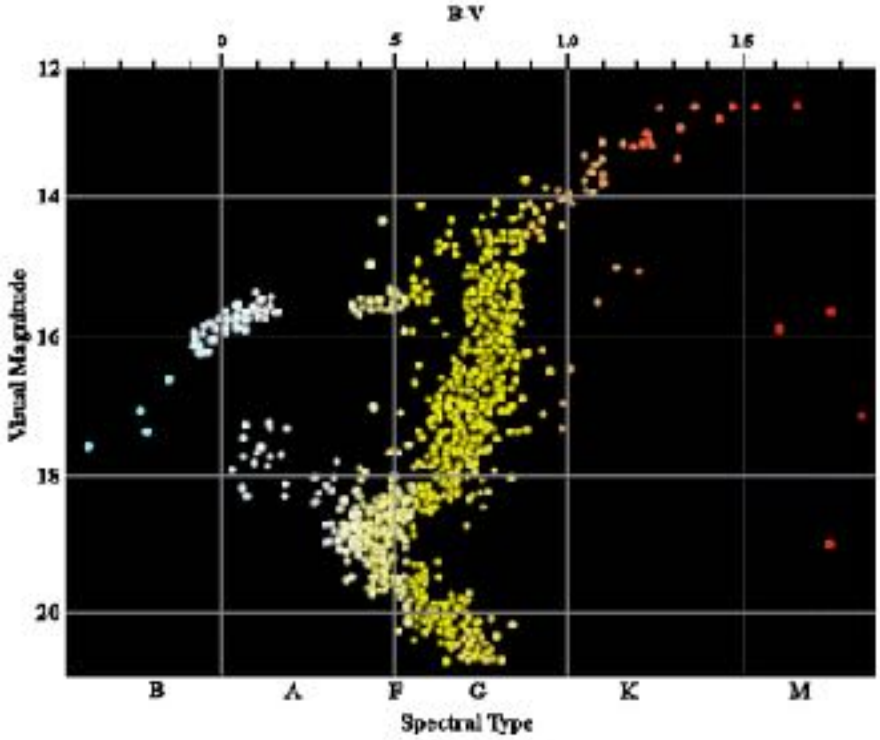
- Stars form relatively quickly compared to their overall lifetimes (taking, say, 100,000-1,000,000) years.
- They fairly quickly settle onto the **Main Sequence,** a term that comes from a plot known as **Hertzsprung-Russell (HR)** diagram of the of a star's absolute luminosity against its surface temperature (which we can estimate accurately from its colour). (See the example diagram for Milky Way stars on the right.) This is probably the most important graph in astrophysics and it shows most stars on a narrow diagonal band in which high luminosity stars are blue and low luminosity stars are red - but also some interesting "branches", such as the "giants". Such a very obvious pattern tells any scientist that they have been handed a major clue to the underlying physical processes.
- In fact we now know from relatively straightforward hand calculations -



undergraduate level astrophysics - that stellar luminosity must be roughly proportional to mass cubed - more or less as shown in the graph. We can also be confident that most stars spend nearly all of their lifetime on the Main Sequence, looking almost unchanged, until their fuel starts to run out. (Complex computer calculations are, however, need to accurately reproduce all the wiggles on the main sequence and also explain the **red giant** branch stretching up to the top right.)
- We also know from the same theory that the surface radii of stars of different masses in their main-sequence lifetimes does not vary by very much. (More mass means bigger gravity which pulls everything together more strongly.) That means that at higher masses, more light has to escape from a not very much larger surface, and that means that the temperature of the surface must be much higher. The strong correlations of both luminosity and temperature with initial mass explain the general trend of the HR diagram main sequence. (Explaining the **giant branch** and the **white dwarfs** is much more complicated.)
- The amount of fuel available to a star is proportional to its mass, so it is relatively easy to estimate that:
  - A star of about the mass and luminosity of the Sun will stay on the main sequence for roughly 10 billion years. (Computer calculations refine this estimate and show that it gets slightly brighter during this time, but not by very much.)
  - A star of 100 times the mass of the Sun (about as large as they come) is a million times as luminous, but it only has 100 times as much nuclear fuel. It will, therefore, burn through its fuel reserves in only about a million years. (More detailed computer calculations can produce more accurate estimates.)
  - A star that is less massive than the Sun (say 1/10th the mass) will have a lifetime that is much longer than that of the Sun and in fact such a star could have been born early in the history of the Universe and will have hardly changed at all today. Furthermore, it will continue to look much the same way for perhaps another 100 billion years.

The normal HR diagram for stars in the galactic plane looks like the image on the previous page, and shows the result of a situation where stars of all masses are forming all the time - and also dying. The stars that are not on the main sequence are mainly in the process of dying.

We can explain the all the features of the HR diagram because one of the things that happens to a most stars that are close to exhausting its fuel is that they turns into a *Red Giants* becoming both more luminous and much redder at the same time (it does this by expanding its surface by a vast amount). Although it is generating much more light, its surface is now so big that it is much cooler. Most eventually shed this extended outer envelope leaving only the small but very hot core and then fall down into the *White Dwarf* region of the HR diagrm. This lifecycle is typical of stars up to about 5-8 times the mass of our Sun. However, larger stars (from the upper left of the main sequence) explode as a supernova and completely disappear.



Hence, if we look at the HR diagram for a globular cluster that is perhaps 5-10 billion years old (the image to the right is typical) we would not expect to see hot blue stars - they have already used up their fuel and exploded. Stars of somewhat lower mass will have become red giants, but stars of about the Sun's mass, or less, may still be burning happily away on the main sequence. In fact, if we look for the main sequence "turn-off" - the distinct kink where the main sequence is truncated and there is a turn towards the upper right - we have a method of estimating of the age of the globular cluster: it corresponds to the age of the stars that would have just exhausted the hydrogen in their cores. All the other features on this diagram - kinks, curves and gaps - that arise from very the complex nuclear physics that comes into play when the hydrogen starts to run out, are explainable with modern computer modelling.

It is worth remarking that the bluish stars that at first sight may seem to be a continuation of the main sequence after a gap, at the left (visual magnitude about 15, B-V ~ 0), are NOT main sequence stars. They are stars that have already climbed from the main sequence up the giant branch and then at some point started burning helium in their core, which flips them to the left before they move back to the right when the helium is gone. This is known as the "horizontal branch" because these stars can change colour (and surface temperature) without changing their luminosity. They do this by shrinking their outer envelopes while getting hotter, before expanding while getting cooler. (The astrophysics of late stellar evolution are *very* complicated!) The horizontal branch is interesting because it should be possible to use it as a "standard candle": stars on the horizontal branch should have very nearly the same absolute magnitude (that is the same luminosity - see Appendix A). So, if we measure their relative magnitudes (how bright they look from Earth) we might be able to use then to estimate distance. This is not very easy because stars spend a very small fraction of their total life time on the horizontal branch so in order to see a significant number of stars marking this line you need to measure the brightness and colour of a very large number of stars in a globular cluster (tens to hundreds of thousands). This is only possible for the more distant clusters (in which we are particularly interested, because it is hard to measure their distance by other methods) with instruments like the Hubble Space Telescope (HST) and sophisticated software to automatically scan images and make hundreds of thousands of measurement.

Note the labelling of the axes on the globular cluster diagram. The horizontal axis is labelled firstly with the "B-V" colour index at the top. This compares the amount of light seen through a telescope in "Blue" and "Visible" (roughly green) filters and the size of this index correlates directly with temperature (see the galactic HR diagram above, in which temperature is also shown). The bottom of the plot also shows "Spectral Type", which is a label invented by astronomers early in the 20th Century to classify stars and arrange them in a sequence, labelled A-M, using only clearly observable characteristics in the star's spectrum, such as the appearance or disappearance of lines from particular elements. As the astrophysics became better understood it was later realised that some of the classification features they were using were less physically important than a relatively straightforward estimate surface temperature (which you can get just by looking at the colour). Astronomers are notorious for hanging on to outdated labels - especially if the older literature is full of them - and so still talk about "B" type or "G" type stars and so on. (You have to put up with this if you want to do astronomy.) There are acronyms to help newbies remember the temperature order, such as: "***O****nly* ***B****ad* ***A****stronomers* ***F****orget* ***G****enerally* ***K****nown* ***M****nemonics*". (The one I was taught was perhaps more instantly memorable - but rather less politically correct for modern tastes.)

Importantly, we can measure the B-V index relatively easily with just two photographs of the sky through two different coloured filters, and from those we can get a reasonably good indication of stellar surface temperature. This works because the variation of light intensity with colour from most stars is very close to being a so-called "black-body" spectrum, which means that its shape is determined *only* by temperature. (If, however, we want to distinguish between Main Sequence stars and those off the MS but with similar colour we need to record a full spectrum. There is, after all, some rationale behind still using spectral types.) I have explained more about the way astronomers measure the brightness of stars in Appendix A and the way they measure of colour in Appendix B.

On the vertical axis we plot visual magnitude, which is a logarithmic measure of the actual brightness we see through the telescope from Earth. There are a number of reasons why a logarithmic measure of light is useful in astronomy (see Appendix A), but it is particularly convenient when observing globular clusters, where all the stars are pretty much the same distance away. We know that a star that is, say, twice as luminous as another in the same cluster will always look twice as bright, whatever the distance from which they are observed, and so on a magnitude/colour diagram there will always be the save vertical distance between them, measured in magnitudes. This means that the *shape* of the diagram will always look the same from any distance, and would the same as that we might plot if we knew the real distance to the cluster and were able to plot a proper HR diagram with absolute luminosities (plotted with a logarithmic scale - or *absolute* magnitudes - see Appendix A) on the vertical axis.

If we now assume that the physics of stars in the cluster is the same as it is in the Milky Way, then we can also assume that the cluster Main Sequence should be the same correlation between B-V and intrinsic luminosity. If we know how bright something looks and also its intrinsic luminosity we should be able to work out its distance. This is discussed in Appendix D.

In practice, modern astronomers may prefer NOT to work with the B and V filters, and for globular cluster work may prefer a “visual” and a “red” filter combination. The longer wavelength light is less affected by dust in the plane of the Milky Way so we can see further. This does not greatly affect the shape of the HR diagram - though the numbers on the axes will be a little different. The main sequence is still a diagonal line, and for a particular cluster age the kink where stars start to evolve into red giants will occur at the same absolute visual magnitude.

Where can we get Colour/Magnitude Data?

There are easy ways and hard ways.

As a student I had a summer job at the Royal Observatory, Herstmonceux Castle working as a research assistant to a professional astronomer. I spent most of the hot summer of 1973 sitting in a cool[3], dark basement room carefully measuring the brightness of stars as they appeared on photographic plates (still then traditional emulsions!) using rather complex and sensitive optical "photometry[4]" equipment. It was not easy, because although great care had been taken with the original telescope observations, each photographic plate had a slightly different light sensitivity because it was made slightly differently or had been developed in a slightly different way, all of which affected the response of the emulsion to the light. Furthermore, the light response of photographic emulsions is intrinsically non-linear (which means that for very low levels of light the blackening of the photographic plate changes only slowly, then as light levels increase it blackens more easily, but finally starts to "saturate" at high light levels - that is, respond more slowly again). I spent a lot of time doing careful calibrations against stars of well characterised brightness so we could turn "blackness" into a light measurement. Someone had to do it! Students were cheap and it was good that they learned about the hard work that went into making the most basic observations reliable and reproducible.

Well, we do not have any time for all of that and, fortunately, most astrophotography is now done digitally, which has two big advantages: firstly the response of the "charged-coupled devices" (CCDs) in modern cameras is very nearly linear - we effectively count the number of photons arriving and within reason it does not matter whether they arrive slowly or quickly; secondly all the data is immediately available in digital form, so we no longer have to handle delicate photographic plates and optical photometers. All the calibration still has to go on, but it can mostly be done in software. A lot of this calibration is now highly automated, so by the time an astronomer gets his hands on the data from modern telescopes it has already been fully calibrated[5]. The final effect is that we get the image or the spectrum as an array of numbers which show precisely how much light energy is reaching each pixel.

Now, with suitable (and readily available) software, we can just draw a box around a chosen star and ask the program to add up all the light within the box. This is then converted into an *apparent* or *visual* magnitude - a well defined measure the the amount of light reaching us.

We repeat this for the same star on plates exposed through different coloured filters to get the colour index (used as the horizontal axis of the HR plot). One still has to be very careful that the star you have chosen is not actually two stars that are so close together that the telescope can not easily resolve them apart. This type of "confusion" is always possible when dealing with globular clusters because even with the best telescopes, because cluster stars can be so concentrated that some overlap of images is inescapable. We deal with this by only measuring

---

[3] The basement was maintained at a rather cool 18°C because the carefully calibrated optical equipment needed to stay at a constant temperature to stay carefully calibrated. I had to go into work carrying a wooly jumper when the temperature outside was hitting 30°C.

[4] Photometry is the process of measuring the amount of light registered. When we say that a star "has photometry" it means that someone has already made careful measurements, using in several wavebands. Photometry is the bread-and-butter of professional astronomy.

[5] The downside of the modern, very efficient and reliable method of making astronomical observations using professional technical staff at telescopes is that astronomical researchers no longer get to fly to exotic mountain top locations in Hawaii, Chile, the Canary Isles or Australia and work through the night in very cold telescope domes sitting on the end of a telescope, trying to load and unload delicate glass plates with frozen fingers while trying to stay awake. A trip to Chile was a temptation once held out to me when I was looking for a research studentship many years ago. In the event, I choose a research group that gave me a less exotic, but much more convenient five mile ride from Cambridge out to the radio telescopes at Lord's Bridge, where I could do all my observing work between breakfast and lunch - and I only had to cycle it at weekends.

stars in the outer regions of the cluster were there is no overlap. Even so, it is sometimes not possible to resolve very close binaries - of which there are significant numbers - into separate stars. Professional astronomers worry a lot such matters and have worked out ways of recognising this type of problem.

Even though this process is educational and is very much faster than the methods I had to use, it is still rather time-consuming if you need to produce plots containing hundreds or thousands or even tens of thousands or hundreds of thousands of stars. (Astrophysics undergraduates are still sometimes required to go through the basic methods so they understand the way it works.) Hence, mostly, these days astronomers generally use even more sophisticated software that *automatically* identifies all the stars on a plate, spots potential confusions, draws its own boxes and finally produces a *source catalogue* with the position and the measured brightness of each star. This is known as *source extraction*.

It is not quite as easy as it sounds, because the software has to have lots of built-in rules, for example to recognise and avoid measuring overlapping stars, and experienced users have to understand how these work in order to set the right sensitivities for the particular work they are doing - and they must carefully examine results to spot when things are not working properly. Astronomy research students have three years to master all the details, but we need to be even faster if we are going to get down to astrophysics.

Therefore, if we are *really* lucky, we might find that someone has already done the cataloguing for the globular clusters in which we are interested *and* published all the photometry data on the Internet.

There are three very good sources of data on the Internet:
- The Sloan Digital Sky Survey (SDSS) carried out by a 2.5m diameter at a fully automatic ground-based telescope in Arizona.
- The GAIA satellite, which is specifically designed to accurately measure the positions and brightness of billions of stars visible from Earth orbit.
- The Hubble Space Telescope (HST).

SDSS: The big advantage of the SDSS is that it covers a great deal of the sky. Every part of the sky visible from Arizona was scanned whether any astronomers were particularly interested in that specific area of the sky or not. All the published datasets are fully calibrated, and for each area there are five different images recorded through five filters of different colours from ultraviolet to infrared. Although the 2.5 meter diameter telescope used is relatively small by professional standards it is certainly adequate to resolve hundred and sometimes thousands of individual stars on images of globular clusters in which we may be interested.

We are lucky with the SDSS, in that for a small sample of globular clusters the plates have already been scanned and photometric magnitudes for a large set of cluster stars have been catalogued and made available on the Internet. See, for example, this page https://classic.sdss.org/dr7/products/value_added/anjohnson08_clusterphotometry.html.[6]

The SDSS data appears to confined to only about a dozen of the clusters around the Milky Way. If we want to look at a larger sample then we would have to download images of the unscanned clusters and revert to estimating colour using the star-by-star draw-a-box-and-measure method. This is an option if you really want to understand the whole experimental chain.

GAIA: It seems likely that data from GAIA will eventually become the definitive information source on stellar photometry, because it was specifically designed for this purpose. Indeed, its catalogues already contain data on *billions* of stars and is already possible to find that many globular clusters have been automatically scanned and tens of thousands of stellar brightnesses measured. For some of the closer clusters, parallaxes (i.e. distance measurements) of cluster stars are also available. I shall, however, not suggest that we follow the GAIA route further

---

[6] See Appendix C for a short note on the cataloging of astronomical objects and their common identification by widely used catalogue ids.

because getting hold of the exactly right data seems to involves more technical steps than we might want to get involved with for a relatively short project. (It is not immediately easy to say "*I want data on all stars associated with globular cluster NGC104*". You have to set up database searches that say "*I want data on all stars within a certain angular separation of this point on the sky*." None of this is particularly hard, and professional researchers know that they need to understand how to search the GAIA catalogues. For us, it is just an additional learning curve that we might be able to avoid.

Hubble Space Telescope: As explained earlier, globular clusters are important for understanding astrophysics, so a fair amount of HST time has been used to look at large sample of clusters, and many of these studies have involved making individual colour/magnitude catalogues for named clusters. Furthermore, because of the HST's superb resolution and sensitivity, we will be able to see many individual stars even in the most distant clusters, so these catalogues are often very extensive - in some cases containing 100,000 stars or more. Even better, as a matter of policy, the raw data from *all* Hubble research programmes are eventually published on the Internet through the "Hubble Legacy Archive"[7] which turns out to contain a great deal of photometry data for a large number of globular clusters. At the moment this is the *best* quality data available and it is also the most *easily* available, so it makes sense to use it for the basis of a project. The next section explains how we might get hold of this data.

---

[7] http://hla.stsci.edu is the "Welcome" page for Hubble data, which is searchable using standard catalogue ids for astronomical objects (e.g. "M80"). There are different ways of finding all data collected as part of one Hubble research proposal via the **Mikulski Archive** at https://archive.stsci.edu/index.html. Anyone at all can download this data - though the search tools were design by and for professional astronomers, so you will need a bit more specific guidance to help get the data you need onto your own computers.

Procedure for Getting Hubble Data on Globular Clusters

Hubble observations are all associated with a *proposal* for a scientific program. If the proposal is considered good then time is allocated on the telescope. The importance of globular clusters for astrophysical studies means that a substantial amount of time has been given to observations.

The "Proposal Identifier" of the relevant research programme on globular clusters is 10775[8]. We can use this identifier to search the Hubble archive for all related data.

1. Go to the web site http://hla.stsci.edu This will take you to the *Hubble Legacy Archive* "Welcome" page.
2. Click the green box saying "Enter Site Here".
3. Click on "advanced search" (next to the "search" and "reset" buttons).
4. Look for the box labelled "Proposal ID" and enter "10775" and press your "Enter" key. Wait while the system searches. It should return a long list (over multiple pages) of all data associated with this research programme. There are many globular clusters listed.
5. We can limit our search to a shorter list by placing the name of a known globular cluster in the box next to the "Search" button. I choose to search on "NGC104" (which I typed in without the quotation marks and with no spaces between NGC and 104). You will also see it referred to "47 Tucanae". Is is relatively near, intrinsically large, and therefore very bright on the sky.
6. We are interested in observations that have data for both of the filters F814W and F606W. Each observation with the HST is listed on a separate line.
    - Be sure that the "**Inventory**" tab is selected for the data listing. (It is usually selected by default.)
    - These will be identified with the string "F814W/F606W" in the column labelled "**Spectral_Elt**". (You may have to use the "Next" button a few times, or use the "**Show…results per page"** to select a larger number of lines to be displayed on a page.)
    - The observations that you will be able to use must have a hypertext links in the "**DAOCat"** and "**SEXCat**" columns. (This means that the data contains "catalogues" automatically compiled from a "source extraction" - i.e. a program that look for and measures stars. including photometry and positions.) The DAOCat catalogue contains data for point-like objects (e.g. stars). This is the catalogue in which we are interested. (SEXCat contains information on "extended" objects, such as galaxy images.)
    - I first selected the line for observations with the WFPC2 data (which has fewer points than the line for ACS/WFC data and is easier to plot with Excel). *The distance estimation will, however work better with ACS/WFC data, and a different plotting tool.*
    - You do not need the FITS files for this exercise. These contain the images captured by Hubble from which the photometry data has been extracted. They are potentially very large and also need special software to view as images.
    - For a selected line that contains the hypertext links detailed above, click the links with the little shopping basket (i.e. the DAOCAT links).
    - You can click on multiple files to download the collection together. (Try a few at first, then see if your computer connection can manage larger bundles.)
    - To download the data click the shopping basket tab at the top of the list, selected "zipped" options, and click "Fetch HLA Data".
    - The data should appear in your "Download" directory as a zipped file containing all the data items that you have requested.
7. Uncompress the data. It is a directory named: "HLADATA-<unique-request-number>" containing subdirectories each of which has requested data from one set of Hubble observations (e.g. a single line in the search results page). You will see that the directory and file names are build from sub-strings: 'hst_10775_<VisitNum>_<Detector>…..' where the <..> sub-strings correspond to column values in the HLA database. The relationship between the name (e.g. NGC104) in the "Target" column and the value in "VisitNum" column (e.g. 60 in the dataset name "hst_10775_**60**_acs_wfc_multiwave_daophot_trm.cat" is the only way you have of knowing that a dataset refers to a particular astronomical object. For convenience, to help with graph labelling, I have constructed a table in Appendix E showing this relationship.

---

[8] http://archive.stsci.edu/proposal_search.php?mission=hst&id=10775

```
# All refereed publications based on data obtained from the HLA must carry the following footnote:
#
#     Based on observations made with the NASA/ESA Hubble Space Telescope,
#     and obtained from the Hubble Legacy Archive, which is a collaboration
#     between the Space Telescope Science Institute (STScI/NASA), the Space
#     Telescope European Coordinating Facility (ST-ECF/ESA) and the Canadian
#     Astronomy Data Centre (CADC/NRC/CSA).
#
# One copy of each paper resulting from data obtained from the HLA should be sent to the STScI
# ------------------------------------------------------------------------------------------------------------
# Proposal ID    :      10775
# Image File Name: hst_10775_60_wfpc2_total_wf_drz.fits
# Target Name    : NGC104
# Date Observed  : 2006-03-13
# Time Observed  : 01:42:15
# Instrument     : WFPC2
# Detector       : WFPC2
# Target RA      :    6.021667
# Target DEC     :   -72.080833
# Orientation    :       0.000
# Aperture RA    :  -999.000000
# Aperture DEC   :  -999.000000
# Aperture PA    :     -999.000
# Exposure Start : 53807.070000
# Exposure Time  :      326.000
# CCD Gain       :     -999.000
# Filter 1       : detection
# Filter 2       :
# ------------------------------------------------------------------------------------------------------------
# Data Release Version :           (DAOphot Catalog)
# Aperture Radius     :0.30
# Nobj                :      4579
#
#  Object    Pixel Position      ICRS Coordinates            f606w     f814w    f606w    f814w
#     ID       X          Y              RA             DEC          | MagAP2----------| TotMag----------| CI--------------| Flags-----------
       1  1551.324   226.321  5.8098579 -72.1866695  21.213  20.543  21.043  20.359   0.864   0.905      0       0
       2  1519.427   229.627  5.8127544 -72.1865782  18.873  18.662  18.703  18.478   0.867   0.908      0       0
       3  1568.051   234.416  5.8083395 -72.1864443  21.438  20.648  21.268  20.464   0.939   1.065      0       0
       5  1503.282   250.457  5.8142215 -72.1859999  17.879  17.688  17.709  17.504   0.834   0.905     16       0
       7  1590.901   251.796  5.8062659 -72.1859610  22.100  21.340  21.930  21.156   1.292   1.265      0       0
       8  1597.274   251.541  5.8056872 -72.1859680  18.702  18.519  18.532  18.335   0.849   0.908     16       0
       9  1511.206   257.267  5.8135024 -72.1858106  20.395  19.898  20.225  19.714   0.891   0.990      0       0
      11   1572.373    267.874  5.8079493 -72.1855148  22.276  21.347  22.106  21.163   0.760   0.862     16       0
      16   1628.061    281.713  5.8028940 -72.1851292  20.778  20.233  20.608  20.049   0.861   0.939      0       0
```

Figure 1: the initial part of the DAOPhot catalogue file.

8. Each of these directories will contain the files you have selected for this line (probably only the DAOPHOT file if you have followed my recommendation). *N.B. The ASCII file may appear on your Windows computer with the "executable" flag set so it may look like a program to your virus checker and be blocked.* You can unset the executable flag with Windows "File Properties".
9. Load the file into a text editor. I used NOTEPAD. The file should look like Figure 1.
10. The best graphs - especially when you have lots of data points - are produced with dedicated graph plotting applications, such as GNUPlot (which is a favourite with physical scientists). However, you may only have Microsoft Excel available to you and we have to go through a bit of a rigmarole to make this work. (Later I will explain how Python can produce better graphs.)
    - Cut away the top section of the file - all the lines beginning with '#'.
    - Rename and save the file with a file extension of '.txt'.
    - Now import the file into MS Excel. It is NOT in an Excel compatible format so you will have to learn how to use the Excel data import tool. Excel likes to have "comma-separated-values" to delineate the different columns. We have spaces between the columns, so we have to find a way to tell Excel to use spaces to recognise columns.
    - The Excel data import user interface is somewhat different in each version of Excel and your version is almost certainly different to mine (which is old), so I do not intend to go through the details. This is a useful skill to master, so I leave it to you to read the "Help" pages or ask your ICT teacher.
    - In outline, in my version, I open Excel and on the Data tab I select "From Text". Then I have to specify that my file is "Delimited" and select "Spaces" to force the data to be retrieved into Excel columns.
11. Now you need to select the data you need. These are the columns that were labelled f606w, f814w/MagAP2. (These are the 6th and 7th columns counting from the left.) They represent the measured amount of light, in magnitudes through the filters centred on 606 and 814 nanometers when using a standard size of aperture placed over the star being measured. The

disadvantage of this measurement is that it includes a bit of background light around the star. In some BUT NOT ALL catalogue files columns 8 and 9 contain slightly smaller value corrected downwards to take account of this (as in Figure 1) giving a slightly better estimate of total magnitude. However, there are many catalogue files that just contain the values 999 in columns 8,9. For consistency I suggest we just ignore these slight better estimates.

- Generate another spreadsheet column with the difference f606w-f814w.

12. Now plot the visual magnitude (from column f606w) as the vertical axis against this difference as the horizontal axis. Excel knows this as an "X-Y" scatter plot. (You do not use "lines".)
    - Remember that higher visual magnitude numbers mean less light so the scale of the vertical axis should be reversed (high values at the bottom). You may also need to restrict the range of the axis to avoid the occasional bad data item (which will have the value "999.0". Excel can do all this if the select an axis format option for the vertical axis. In my version you need to click on the axis and select the "Numbers" tab.
    - You are trying to plot thousands of values. (Some files contain in excess of 100,000 values.) Excel will takes it time to draw them all. By default it will also use a fairly large plotting symbol and these are likely to overlap each other turning parts of your graph mainly black. Find out how to tell Excel to use the smallest possible dot for plotting. (Click on one of the plot symbols and follow the "Marker" tab to find various options for changing the symbol used and its size.)
    - For the photometry catalogue files that contain very large numbers of data points, even with the smallest marker available, the graph may still look largely black in a broad band around the main sequence.
    - Excel can certainly do plots for the smaller files, but it is also certainly a clumsy way to plot large numbers of points on an X-Y diagram. We really need a better way and I hope that we can find a way to use a Python program - which I can supply - which should run on your school computer network. (See Figure 3 below for an alternative type of graph.)
13. You now have a magnitude vs colour-index diagram. It is not exactly the same as a standard HR diagram, because we do not have absolute magnitudes and the coloured filters used are "Visble" and "InfraRed" rather than "Blue" and "Visible". It still shows all the same features so we can, however, use it for similar purposes.
    - My diagram (Figure 2) shows the main sequence running bottom left to top right, with a turn-off towards the "giant branch" (running toward the top right) at f606w = 18 and (f606w-f814w) ~ 0.2.
14. NGC104 is a relatively close globular cluster with an accurately measured distance of 4000 +/- 350 parsecs (about 13,000 light years). We can use it as a calibration standard to work out distances to other clusters. (Though we also need to check for signs of reddening by dust.)
15. You should now go through the same process for other globular clusters you can find.
    - Having obtained our colour magnitude plot for a different cluster, we now need to ask what magnitude offset needs to added to each stellar magnitude to make the main sequence fit as closely as possible over the NGC104 plot. (Note that the *gradient* of the main sequence is not always the same. Why is this?)
    - We leave it to you to devise the precise way this could be convincingly displayed. (It could involve printing the graphs on transparent film, for example, and shifting them up and down to get an good overlay. You can do the same electronically if you wish.)
    - Look up the definition of magnitudes to see how you can convert a magnitude difference to a distance multiplication factor.
16. Compare your estimates to published estimates. Your results do NOT take account of the dimming/reddening effect of interstellar dust, which is sometimes substantial - especially in the visible part of the spectrum. Do the plots show any signs that reddening is affecting the results? (Can you do a distance estimate based on aligning the horizontal branches where they are visible?)

Figure 2 below is a "scatter" plot containing about 4500 points, and even when the markers are at set to the minimum size they overlap and it is hard to pick out the exact line of the main sequence. However, unless we plot several thousand stars we would not pick up the "giant" branch pointing to the top right, containing the rarer, luminous short-lived red giant stars which are the eventual fate of most main sequence stars. Because this red giant phase of life does not last long this area of the plot is poorly populated.

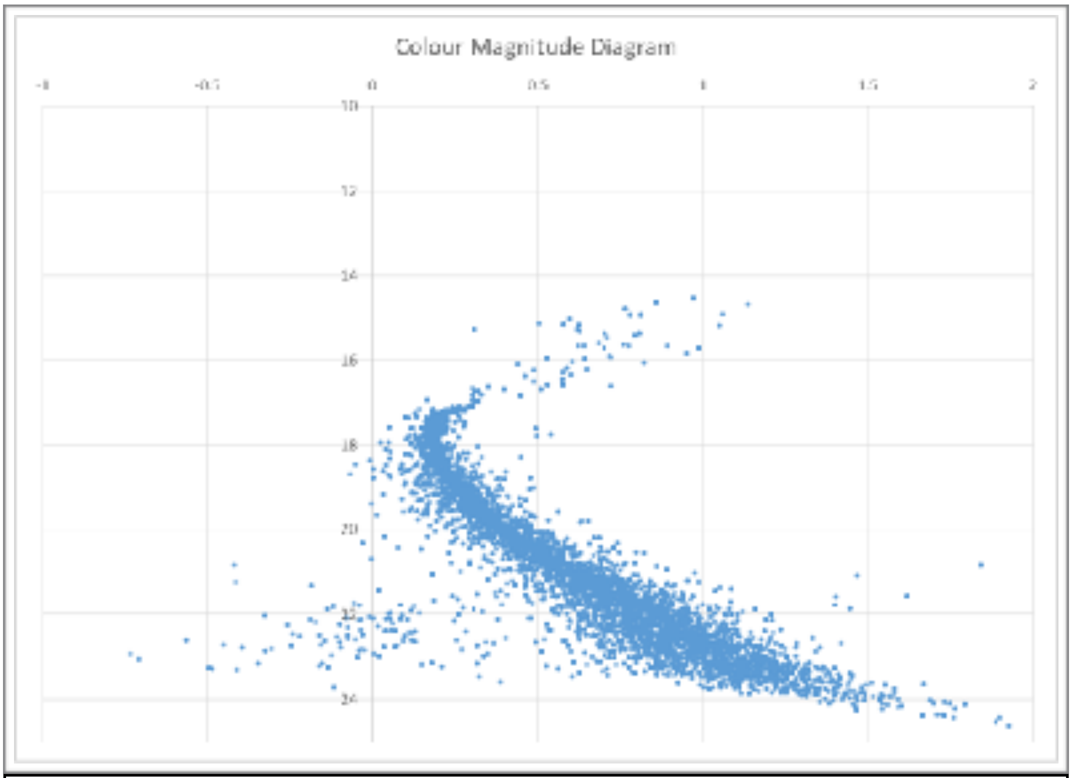


Figure 2: Horizontal axis is (MagAP2/f606w-f814w). Vertical axis is Mag/f606W.

It shows up much more clearly in Figure 3 which is based on data for 150,000 stars. Here we avoid the excessive overlap of plotting points by creating a sort of 2-dimensional histogram in which the colour indicates the density of points. (In fact, we base the colour on the *logarithm* of

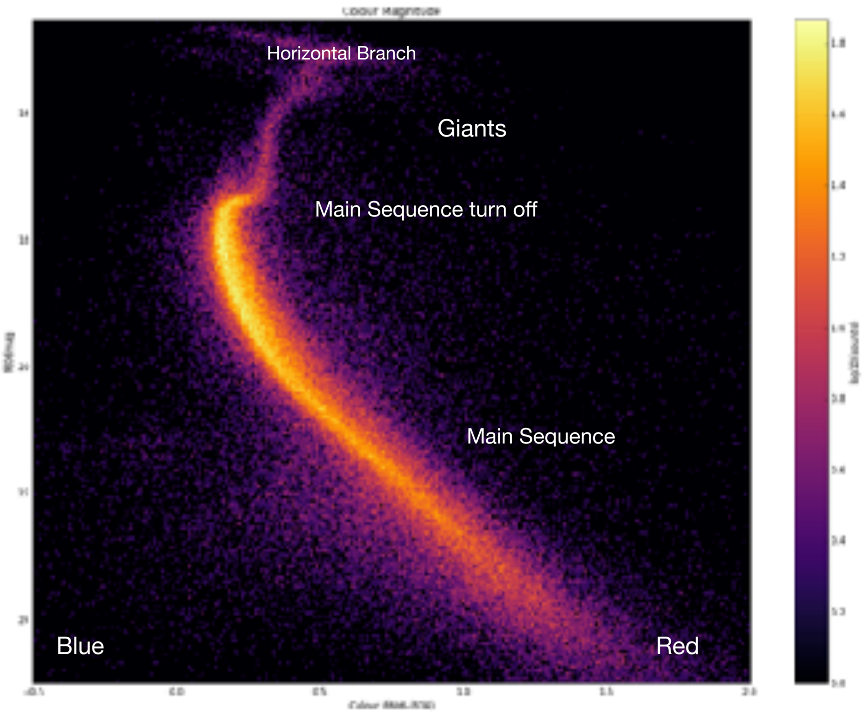


Figure 3: A colour magnitude diagram of NGC104 produced with Python's MATPLOTLIB "hexbin" option (brightness represents log of point density).

the density, which shows detail in the less populated areas while also avoiding excessive sideways spread of the main sequence.)

I wrote a short Python programme which made use of the MATPLOTLIB library to plot Figure 3. This is an extremely useful skill for any aspiring scientist or engineer who is likely to find themselves needing to handle large amounts of data. (This is, increasingly, most of us.) While Excel is very convenient for handling relatively small amounts of data, once you find yourself dealing with thousands of rows it is almost always easier to write short Python scripts. I will supply the Python program that I employed for plotting.

You may have noticed that the data in the Hubble Legacy archive seems to divide into two classes: observations made with the older “wide field and planetary camera” tagged “WFPC2” in the “Detector” column. These tend to have a few thousand data points. Later observations were made with the “Advanced Camera for Surveys” and “Wide Field Camera” tagged “ACS/WFC in the “Detector” column. This is a more recent and more sensitive and higher resolution instrument, so these observations may contain hundreds of thousands of data points. These are more useful observations because the position of the main sequence is more accurately determined and there are enough data points to show some of the short-lasting phases of stellar evolution.

In particular Figure 3 is now showing a structure known as the “horizontal branch” right at the top of the graph - a distinct spreading left and right amongst the giants (only just becoming visible even with this amount of data). At a certain stage in stellar evolution (when the core starts to burn helium) a star will vary in colour, going first blue then back to red, while keeping almost the same luminosity. This is a relatively brief phase so we only see this detail when we can plot many tens or hundreds of thousands of points - and it has only been possible to collect data such as this with instruments like Hubble combined with automatic photometry. (Even the most obsessively dedicated experimenters would be unable to measure more that a few thousand stars one by one.) Note that if we tried to plot 150,000 points on a conventional scatter plot (such as I show in Figure 2) a large part of the graph would simply appear black.

The position of the horizontal branch, when it is visible, can also be useful standard candle for estimating globular cluster distances because the behaviour arises directly from the fundamental physics of stars so it should mark *the same absolute magnitude* in the HR diagram of all globular clusters. Furthermore, because this track is nearly horizontal it is almost unaffected by the reddening effect of dust. (Though the vertical position is still subject to the dimming effect of “extinction”.) In principle this can be a more accurate method of determining distance than main sequence fitting alone because we have only one extinction correction to worry about. However, even with the HST we can only collect sufficient photometry data to reveal the horizontal branch for the brighter and more populous clusters.

Both Figure 2 and Figure 3 clearly show the Main Sequence “turn-off” where the higher mass stars (those that evolve more quickly) have completed their main sequence phase. As clusters get older this turn-off point should move down the Main Sequence towards the redder low-mass stars. It should be possible to arrange your globular cluster HR plots in some type of age sequence. Is there much variation in age? Consider how dust reddening might affect your results?

Look up the astronomical coordinates for your clusters (in terms of right ascension and declination). Now look up the trigonometric formula for converting RA and Dec to Galactic latitude and longitude. (Galactic latitude is the angle of the line of sight with the Milky Way disk plane. Galactic Longitude is 0 degrees when looking towards the galactic centre.) We might expect that our line of sight passes through more dust for a low latitude line-of-sight, especially when looking towards the galactic centre. Do we see any correlations between the errors in our distance estimates and these galactic coordinates?

Appendix A: How Astronomers Measure Light

Astronomers traditionally measure light in terms of *magnitudes*[9]. This dates right back Hipparchus, and ancient greek astronomer who believed that he could distinguish six difference brightness in stars. 1st magnitude stars were the brightest and 6th the dimmest.

The brightness of stars we see from Earth are now known as *apparent magnitudes* because we now know that stars are at different distances, so a bright star may have a rather low absolute luminosity, but be fairly close, while a dim star could have a very high intrinsic luminosity but be rather far away. When astronomers were able to measure distances to stars they also introduced the term *absolute magnitude*, which is a measure of the brightness a star would have at a standard distance of 10 parsecs[10].

However, our eyes do not respond linearly to light (that is, twice as much light does not produce twice as much sensation). Measurements show that the scale is logarithmic, with each magnitude step corresponding to a multiplication in the amount of light. These days the magnitude scale is *defined* such that a change of five magnitudes corresponds exactly to a factor of 100, so one magnitude is a factor of exactly 2.512.

Up until the development of photography, astronomers estimated magnitudes by looking through telescopes and comparing the apparent brightness of stars with stars of know magnitude. (Many amateur astronomers who study variable stars still do this.) Professional astronomers nearly always use photography and this introduced a problem, because the human eye is responsive to a range of wavelengths, but more sensitive in the red and yellow than to blue light, whereas the first photographic emulsions were much more sensitive to blue light than red light. How could information from photographs be consistently related to the older visual magnitude scale?

To deal with this issue, astronomers *defined* the star Vega to have apparent magnitude 0 when photographed through *any* colour filter. (Hence, all stellar magnitude at any colour are related to that of Vega in the same wavelength range.) There are two problems with this approach:
- Using modern instruments, Vega turns out to be slightly variable in brightness.
- It does not relate easily to standard physical units.

Modern astronomers therefore often refer their measurements to the *AB Magnitude*[11] system which has a direct definition in terms of flux density.

Flux density is the absolute amount of electromagnetic energy falling on a square meter perpendicular to the direction to the object of interest, per Hz of frequency[12]. This unit is now also used by optical astronomers who increasingly work on projects that span the entire electromagnetic spectrum, but beware, sometimes they quote energy per unit of wavelength, which is related but slightly different. (Google the STMAG system.) You may come across any of these in published data, but you have to realise that making accurate and reproducible measurements of very small amounts of light is a very tricky business, and require a good deal of care.

---

[9] See https://en.wikipedia.org/wiki/Magnitude_(astronomy) for a more detailed history.

[10] Professional astronomers measure distance in *parsecs*, where one parsec is the distance at which an object would appear to shift on the sky (its *parallax*) by one arc-second as the Earth moves from one side of its orbit to the other. One parsec is 3.26 light years or about 3×10^{16} meters. See also Appendix D.

[11] https://en.wikipedia.org/wiki/AB_magnitude

[12] https://en.wikipedia.org/wiki/Jansky Conventionally measured in Janskies (Jy), which is not, however, a standard SI unit, but is directly related: $1 \text{ Jy} = 10^{-26} \text{ W/m}^2\text{/Hz}$.

Appendix B: How Astronomers Measure Colour

Scientists need methods of measurement than can be reproduced (both by themselves and others). When you are reading popular astronomy books or watching Sky At Night on TV it is all too easy to forget that making astronomical observations, particularly of faint objects, in a reproducible way is actually a very difficult technical feat requiring a great deal of care and hard work. A night on the telescope may be followed by weeks or months of detailed analysis in order to be certain that all the data you intend to report is consistent and believable.

The very best way of measuring the colours from an astronomical object is by taking a *spectrum*, that is splitting the light into all its component wavelengths using a *spectrograph* and separately measuring the amount of light at each wavelength. (Look up these term if you do not understand them.)

Astronomers can get a lot of information about the physical processes in stars and galaxies by studying spectra, but it often takes a long time to measure a spectrum (you need to wait for enough photons to arrive in every part of the wavelength band, so spectra of faint objects such as distant galaxies are difficult to capture).

If astronomers want to survey the properties of as many objects as possible as quickly as possible then they take photographs through coloured filters. The Sloan Digital Sky Survey, for example, takes five photographs of the same area of the sky through five different coloured filters. Each filter is sensitive to a broad band of wavelengths and so the telescope will collect a lot more photons in a given amount of time. Furthermore, because we can are looking at a large area of the sky, containing many stars (and sometimes many galaxies) we can scan the images (usually automatically these days) to determine the characteristics of a great many objects from a small number of photographic observations.

Filters and Calibration of Magnitude Estimation

HR diagrams have been traditionally plotted using the UBV filter system colours (and in particular just the B and V filters). This is also known as the *Johnson-Morgan* system (and sometimes as Johnson/Cousins after the addition of standard R and I filters by Kron and Cousins).

The B and V filters are centred on 415 and 520 nanometers wavelength respectively, but in practice they pass a broad band of wavelengths which has a maximum transmittance at the specified wavelength, but falls away on either side. If you want to compare observations made at one telescope with those at another you would ideally wish to have filters at both telescopes that pass light in exactly the same way, and there is, in fact, a precise specification of how the filters are to be constructed, of what materials, and how they should be tested.

This is important because the amount of light emitted by stars varies with wavelength. Our Sun has a peak in the middle of the visual band (greenish) but falls way to the blue and to the red. If you observe the Sun in the blue with a filter that had a different width to the standard (or even a different variation of transmittance with wavelength) you would observe a different total amount of light passing through.

Unfortunately, even though we try as hard as we can to standardise filters, we cannot standardise telescopes and detectors, which add their own idiosyncrasies in terms of passing certain colours better than others. Furthermore, although the UBV filters have been important for the study of stars there are a number of reasons other filter sets are more suitable for different types of professional astronomical observations.

The Sloan Digital Sky Survey filter, for example, are optimised for detecting faint galaxies (partly because they are better at rejecting light emitted by the Earth’s atmosphere - they do this by carefully arranging for little gaps between the wavelengths where one filter leaves off and another starts).

- The “u” filter of the SDSS is most sensitive in a band of wavelengths slightly shorter than the bluest colour our eyes can see in the **ultraviolet** region.
- The “g” filter has a band pass in the **green** region.
- The “r” filter passed wavelengths in the **red** region.
- The “i” filter passes in the **near infrared** (i.e. just beyond the visible region at the red end of the visible spectrum).
- The “z” filter passes **far infrared** - that is even longer wavelengths than the “i” filter.

Hence, if our photograph contains young blue stars they will look bright on the “u” and “g” photographs and relatively less bright in “r”, “i” and “z”. In contrast dusty star-forming regions often appear very bright in the “i” and “z” plates and are dim on the “u” and “g” plates. (Dust absorbs the light emitted by newly formed stars, which heat up and re-emit the energy at long wavelengths.) Stars like our own Sun would appear brightest in “g” plates because it emits most energy in the central part of the visible light range.

Astronomers do not in general attempt to measure the absolute amount of light coming through a filter from a star or galaxy. Instead they compare the amount they detect with an observation of a known standard star. Good standard stars give out a steady amount of light and have a smoothly varying spectrum. We need standard stars in every part of the sky for convenient and frequent calibration, but we need at least one star that has had its light spectrum measured with extreme care and related back to absolute physical units. Astronomers used to relate everything to the star Vega (which was *defined* to have apparent magnitude 0 through every filter) but we now know that in fact its light intensity does vary by a very small amount, so there is a more modern “AB” system which ties observations to absolute flux densities.

The practicing astronomer at the telescope, however, still needs his standard reference stars, though the hard work of regular calibration is often these days left to the full time professional technical staff at large telescopes, while research astronomers have the convenience of picking up already calibrated data.

All this works well because most stars are pretty close to being “black body” radiators, which means that the relative amount of light they emit at each wavelength (forgetting spectral lines for the moment) is determined *only* by their surface temperature. Turning this round, we can, in principle, use the ratio of the amount of light from *any* two known wavelengths to determine the surface temperature.

Hence, even if we take photographs using the SDSS filter set, for example, we can use this to infer the temperature of a star. We can then use our temperature estimate to infer what we would have seen through, for example, the UVB system, and plot standard HR diagrams even if we are not using UBV filters. Standard inter-conversion formula are available - see below.

Color Indices

Astronomers therefore measure colour by looking at the *difference* between the light emitted by astronomical objects when photographed through standard colour filters compared to the average amount of light.

The normal colour index for plotting HR diagrams is B-V, that is the difference in magnitude in the B and V bands. Note that these measures are normally quoted in *magnitudes*[13], which are a logarithmic measure defined such that a change of a factor of 100 in the absolute amount of light reaching the Earth is equivalent to 5 magnitude steps - or each step is a change by a factor of 2.512 (because $2.512^5 = 100$). This is very convenient because a difference in magnitudes is not affected by distance. (A star that is twice as bright as another is still twice as bright if you move it to 10 times the distance, so the magnitude difference, being a multiplying factor, is still the same.) It is also convenient on the telescope because observational magnitudes are always derived by

---

[13] https://en.wikipedia.org/wiki/Magnitude_(astronomy) and see also Appendix A.

comparing the amount of light from one object with the amount of light at the same colour from a standard reference star.

However, with the SDSS data, we might choose to define a “ug” colour index as

$$C_{ug} = u - g$$

where u and g are measures of the amount of light in the u and g bands.

We could, of course, also define different colour measures by comparing green and red, or red and infrared (or even green and infrared). Each of these different measures might be useful for a different purpose. If we want to quickly identify young stars, we might use the C_{ug} and pick the biggest values, if we want to find dust clouds we might use the red/infrared colour as our guide.

We could plot a form of HR diagram using C_{ug} but it would not look exactly like the graphs plotted with the B-V colour index (though it would share many characteristics and would certainly still be useful for a number of purposes). There are various recipes[14] for translating from the SDSS ugriz system into other magnitude systems, such as UBV. Similarly, we can translate from Hubble filter colours to standard UBV colours[15].

This is a complex process, because the best transformation depends on the type of object that you are observing (so you would need to use a different transformation for a star and for a certain type of galaxy and a different recipe again for quasars). The numerical manipulation are also complex. You might also expect that it would always introduce an extra degree of uncertainty into the magnitude estimation. All this is well beyond the scope of our projects.

While ideally it would be better to use observations that require less manipulation, in practice we have to make use of the data that is actually there. Although Hubble has in its extensive filter set[16] some that are quite close to the UBV filter standards[17], for various sound reasons they are not used as extensively as other filter groups. In particular, extensive studies of globular clusters have been made with the F606W and the F814W filters. F606W is close to (though rather wider than) the Johnson-standard “V” colour, but F814W is much further into the red. This does make good observational sense, because globular clusters are mostly old stars and a lot of the stars are much redder than those around us in the galaxy. Conversely, they do not have so many blue stars (those that would show up well in the Johnson “B” filter). We do not waste HST time making observations that do no show very much.

---

[14] http://www.sdss3.org/dr8/algorithms/sdssUBVRITransform.php

[15] http://iopscience.iop.org/article/10.1086/444553/pdf which I include here because it is right to note the appropriate technical source. This is a long, detailed and complex scientific report: even postgraduate students would need to spend a long time digesting the contents in order to be able to use them.

[16] http://www.stsci.edu/hst/wfc3/ins_performance/ground/components/filters

[17] The nearest filters on Hubble to B and V are the Wide Field Camera 3 (WFC3) filers F438W (also identified as WFPC2 B) and F555W for visible V (depending on the camera in use). These do not have precisely the same wavelength-passing characteristics as the Johnson-standard B and V filters but are close. F606W passes a rather wider range of light than F555W. F606W, however, seems to be more widely used in the globular cluster observations - perhaps because its wider bandpass means that it is possible to pick up fainter stars.

Appendix C: Identification of Astronomical Objects

One of the first things and astronomer does with a new instrument is to produce a new catalogue of the things that he or she can now see. This goes right back to the ancient Greeks, Chinese and Indians but for our purposes the game really gets going with Charles Messier, a famous comet hunter of the 17th Century. He spent a great deal of time looking for bright smudges on the sky, and repeatedly found that he and others were tentatively identifying the same "nebulae" as potential comets. (Nebulae - which just means "clouds" - also also bright smudges on the sky, but unlike comets they do not move with respect to the fixed stars.)

Messier therefore produced the first catalogue of the positions of bright nebulae, of which he had identified 103 by 1771, recording their characteristics accurately so he and others would not be again confused. (Later astronomers added another seven "Messier" objects that met the original selection criteria, but which he had missed.)

We now know that there are three different types of object in the Messier catalogue:

- True nebulae - glowing clouds of hot gas. These are within our own galaxy and relatively close as astronomical distances go.
- Galaxies - huge and distant collections of perhaps a trillion stars, some very like our own Milky Way.
- Globular Clusters - tight collections of a few hundred thousand stars gravitationally bound together, but orbiting our own Milky Way as a coherent group.

Many Messier objects are old friends to astronomers, who often refer to them (with universal recognition) by their catalogue numbers. So "M1" is the Crab Nebulae, which is a supernova remnant, M3 is a globular cluster, and M31 is the "Great Nebulae in Andromeda" the nearest large spiral galaxy to our own. Catalogue aim to be complete down to a certain visual magnitude and Messier objects are in effect the 110 brightest nebulae in the sky and therefore well studied by both amateur and professional astronomers.

As instruments got better they discovered more and more nebulae - particularly galaxies. So the next most common list of interesting objects is the "*New General Catalogue of Nebulae and Clusters of Stars*" compiled in 1888 by John Dreyer - building on the earlier work of William, Caroline and John Herschels who produced a "*General Catalogue of Nebulae and Clusters of Stars*". The 1888 version has 7840 objects, all identified by NCG numbers and aimed to be complete down to a specified visual magnitude, so it also includes the objects in the Messier catalogue. Hence, M31 is also known as NGC 224. (In practice a lot of objects were missed, and observations in the Southern Hemisphere were particularly patchy.) Most of the galaxy photographs you see in popular astronomy publications are almost certainly drawn from the NGC catalogue, because they are relatively bright and accessible in modern telescopes.

Professional astronomers now have many other catalogues including radio source, X-ray source. The GAIA satellite has catalogued the characterises of billions of stars, and automated surveys such as SDSS are easily capable of collecting information on millions of galaxies. The same objects appear in many of these catalogues, so professionals have to use on-line cross-referencing databases to keep track of the different names used for the more commonly studied objects, and most of the on-line data search tools let you use any valid name. Hence, the globular cluster NGC104 is more commonly known in the literature as "47 Tucanae".

The general rule-of-thumb when writing up work is to use the oldest commonly employed identifier, so everyone talks about "M31" not "NGC 224".

Appendix D: Methods of Distance Measurement

The basis for nearly all astronomical distance measurement is parallax - the apparent variation in a star's position on the sky as the Earth moves around its orbit. This is a very difficult experimental technique, and the uncertainty in measuring the parallax angle translates directly into an uncertainty in the distance estimate. A star that shows a parallax of 1 arc-second. (1 arc-sec is about about 1/100th of the diameter of a human hair held at arms length.)

Professional astronomers measure distance in **parsecs**[18]: which is the distance at which a star shows 1 arcsec of parallax. A star that shows 0.1 arcsec of parallax would therefore have a distance of 10 parsecs. (One parsec is about 3.26 light years or 3×10^{16} meters.)

Earth-based telescopes have to look at the sky through the atmosphere, which tends to blur stars into a fuzzy disk about 0.5 arc-seconds in diameter, but by making many, many measurements of the average position we might estimate the parallax of a star to an accuracy of about 0.01 arc-seconds. So a star 100 parsecs away (326 light years) would be just about measurable - but the distance would have very big uncertainties.

Observations made from space are very much better because they are not distorted by the atmosphere, and we can now reliably measure distances out to a few 1000 parsecs. The GAIA satellite in particular is measuring the distance of hundreds of millions of stars very accurately - including some in nearby globular clusters. (GAIA's accuracy is equivalent to the angle subtended by a human hair 1000km away!) The Hubble Space Telescope has also been used to measure parallaxes for a small number of globular cluster stars. This will work with reasonable accuracy for clusters that are between us and the centre of the Milky Way

Unfortunately, many globular clusters are still outside the parallax range. We have to find different ways of measuring their distance.

The general idea of the **astronomical distance ladder** is to use parallax measurements to understand nearby bright objects sufficiently well that we can recognise the same type of object in much more distant locations. If we know enough about them to estimate their intrinsic luminosity we can measure the apparent brightness as seen from Earth and estimate their distance. This is generically known as the "standard candle" method.

The real problem here is that intrinsically bright objects (those that can be seen from a long way away) tend to be very rare and often outside parallax range. Hence, we have to build our ladder step by step, using nearby, not so rare, but also not so bright objects to calibrate our understanding of somewhat brighter, rarer and more distant objects. Each step is difficult and each adds its own uncertainties, so the final distance estimates are sometimes rather uncertain.

We are going to try to use a method sometimes known as "main sequence fitting". It is particularly useful with globular clusters because it is hard to find alternative standard candles. The general idea is that stars that are still on the main sequence in globular clusters should be pretty similar to main sequence stars in our locality. We first plot of *apparent* magnitude against colour index, and then we *guess* a distance, work out what the absolute magnitudes of each star would be for that assumed distance and compare with the HR diagram for local stars. If the main sequence trend it too high, we increase the distance, if it is too low we decrease the assumed distance, and so on

---

[18] The parsec as a unit made a lot of sense when we had greater uncertainty in the size of the Earth's orbit. (The radius of the Earth orbit, being so fundamental, is know as the "Astronomical Unit" or AU) Every distance in astronomy ultimately relied on parallax measurements and therefore every distance quoted really scales from the value of the AU. Whenever size of the AU was remeasured more accurately every absolute distance estimate dependent on it would also change (whereas the value in parsecs would be unchanged). It perhaps makes somewhat less sense now we know the Earth's orbit to a very high degree of precision as a result of radar measurements, but the parsec is now thoroughly embedded in astronomical practice and the literature. On the plus side, 3×10^{16} meters is a conveniently memorable distance for back-of-the-envelope physics calculations, and also almost exactly the distance that light travels in 10^8 seconds.

until it lies correctly. This technique is better known as **main sequence fitting** - and often appears under this name in the astronomical literature.

This is actually much easier than it sounds, because *by definition* the apparent magnitude and the absolute magnitude are related by the formula:

$$M_{abs} = m_{app} - 5.log_{10}\left(\frac{d_{pc}}{10}\right)$$

where M_{abs} is the absolute magnitude, m_{app} is the apparent magnitude and d_{pc} is the distance in parsecs. Rearranging we have:

$$m_{app} - M_{abs} = 5.log_{10}\left(\frac{d_{pc}}{10}\right) = 5.\left(log_{10}(d_{pc}) - 1\right)$$

You may see this difference $(m_{app} - M_{abs})$ referred to as the *distance modulus*, μ. It frequently appears in the literature on globular clusters because professional work needs to take account of absorption of light by interstellar dust, which is normally documented as an additive correction to be applied to the distance modulus. (See the discussion of "extinction" below.) Rearranging the formula again (and ignoring potential corrections) we can get the distance:

$$d_{pc} = 10^{\left(1+\frac{m_{app} - M_{abs}}{5}\right)}$$

We have, in fact, made a couple of implicit assumptions here that are not entirely accurate.

Firstly, stars in at least some globular clusters tend to be very old and have fewer "metals" in their composition. (Again we have to apologise for astronomical terminology: to an astronomer a "metal" is any element heavier than helium - so carbon, nitrogen and oxygen are considered to be "metals". Learn to live with it - but do not tell your chemistry teacher!) Metals are made inside stars, so seeing metals mean that we are looking at a second or third generation star. First generation stars, formed from primeval material are of "low metallicity". Even very small amounts of metals affects the physics of stars to a small but significant extent because elements like carbon have many electrons and when these become loose in the core they make it harder for light to get out, so the core is hotter and the fusion reactions burn a bit more fiercely. Hence, for the same mass a higher metallicity star is a little more luminous. This means that the main sequence for a low-metallicity cluster is actually just a little lower on the graph than you would expect, in comparison with nearby stars in the Milky Way. We can ignore this problem because we will calibrate our distance estimates against a globular cluster of known distance which we assume has similar metallicity. Professionals use small corrections related to the relative ages of the clusters, but we will not be working to this level of accuracy.

Secondly, we are assuming that the colour index that we see in our telescope is the same as that of the light leaving the cluster. This is true for globular clusters well above the galactic plane, so we are looking upwards out of the "fog" and are little affected. The lower the *galactic lattitude* (the angle of the line of sight with the galactic plane) the more dust we will encounter before getting above the fog. The dust blocks the light making things look dimmer, but it also blocks more blue light than red light so object tend to look much redder. (This is also why sunsets look red: we are looking at the low angle sun through a lot of atmospheric dust.) This is the effect is known to astronomers as "extinction". Unfortunately, of course, the clusters that are furthest away are all likely to be at low galactic attitudes - especially if they are on the other side of the galactic centre. These are just the clusters whose distance we do not find it easy to measure by the parallax method and where we most need an alternative approach like main-sequence fitting.

Professional astronomers have various methods of estimating the amount of extinction in different viewing directions and would use these to apply corrections to the globular cluster observations. Generally these involve subtracting a small amount from apparent magnitudes with the correction being different for different filter colours. This tends to shift the main sequence line up and also to the left.

In practice, there are a number of other methods of measuring distance to the closer clusters (though some of them are not very accurate) so the professionals compare the various distance estimates with main sequence fitting and sometimes use the distance error as a way of *estimating* the extinction in various directions. You might find it interesting to compare your main-sequence-fitting distance estimates with the professional estimates, and see if we can spot any correlation of the distance error with altitude above the galactic plane.

Appendix E: Programme 10775 Globular Clusters

The 'HLA VisitNum' corresponds to the VisitNum column in the Hubble Legacy Archive database, and to a sub-string in the HLA dataset name. 'HB' indicates that the "Horizontal Branch" should be visible on the HR plot. All the ACS/WFC data have a clear main sequence with a turn-off point.

Name	HLA VisitNum	Horizontal Branch?	Name	HLA VisitNum	Horizontal Branch?
PALOMAR1	**1**		**NGC6651**	38	HB
NGC5053	2		**NGC6681**	39	
LYNGA7	4		**NGC6717**	40	
NGC6779	5	HB	**NGC6723**	41	
NGC6366	7		**NGC6981**	42	
NGC1261	9	HB	**NGC6144**	43	HB
NGC1851	10	HB	**NGC6218**	44	
NGC2298	11	HB	**NGC3201**	46	
NGC5186	12	HB	**NGC2808**	47	HB
NGC5927	14	HB	**NGC6656**	48	HB
NGC5986	15	HB	**NGC4147**	49	
NGC6093	16	HB	**NGC5024**	50	HB
NGC6101	17	HB	**NGC7089**	52	HB
NGC6304	18	HB	**NGC5272**	53	
NGC6308	19	HB	**NGC7078**	54	HB
NGC6584	21	HB	**NGC7079**	55	
NGC6714	23	HB	**NGC5905**	56	HB
TERZAN7	24		**NGC6204**	57	HB
ARP2	25		**NGC6341**	58	
TERZAN8	26		**NGC6352**	59	
NGC6934	27	HB	**NGC104**	60	HB
PALOMAR12	28		**NGC6254**	62	
NGC362	30		**NGC6809**	63	
NGC4590	32		**NGC6121**	64	
NGC6171	33		**NGC6357**	65	
NGC6362	34		**NGC5139**	a7	HB
NGC6535	35		**NGC6838**	a8	
NGC6541	36	HB	**NGC4833**	ac	
NGC6637	37	HB	**NGC288**	ad	