Lab 6. The Distance to a Star Cluster

1 Introduction

This first half of this century saw astronomers make great strides in understanding the nature of the universe. The reason for this is obvious: in 1900, the largest telescope in the world was $\sim$ 1-m in diameter; by 1950, a 5-m telescope (the Hale telescope on Mt. Palomar) had been built. As a result, astronomers could see 25 times fainter than they could previously (how many magnitudes is that?) and this led to a series of discoveries about the physical nature of the universe.

As great as this increase was, what came next was even more dramatic. Between 1950 and 1990, astronomers increased their ability to explore the universe by almost a factor of 50 — not by building bigger telescopes, but by improving the quality of their detectors. In the 1940's, the most common detector in astronomy was the photographic plate, which, under the best conditions, is 2% efficient. (That is, 98 out of every 100 photons hitting a photographic emulsion are ignored.) Today, the detector of choice is a CCD, or charge-coupled device. High quality CCDs capture almost 90% of incoming photons. This increase in *quantum efficiency* from 2% to 90% is responsible for the explosion in optical astronomy over the past three decades.

In this lab, you will be exposed (no pun intended) to data taken with a CCD. You will use this data to perform photometry on stars and measure the distance to a cluster of stars in the Galaxy. In doing this, you will perform many of the tasks (and possibly encounter some of the problems) optical astronomers deal with every day. But first, you will need some background about photometry, photometric measurements, and CCDs.

2 CCDs

A CCD is an electronic two dimensional detector, much like a TV camera, but laid out in a grid. You can think of the device as containing tens of thousands of picture elements, or *pixels,* in rows and columns like a chessboard. Within each pixel are two layers of silicon: one containing a large number of excess electrons, and another deficient in electrons. When a pixel is exposed to light, the incoming photons deposit their energy onto the electrons and cause them to jump out of the "donor" layer into the "receptor" layer. Here, the electrons are stored until the exposure is finished. All that is then required is to "read out" these electrons: since there is a one-to-one correspondence between incoming photons and ejected electrons (*i.e.,* every detected photon ejects one, and only one electron) counting the electrons is equivalent to counting the number photons detected.

The phrase "timing is everything" may be trite, but it is the key to understanding how CCDs count their electrons. When an exposure is completed, a small positive charge is placed at the corner of the chip, and at the bottom of each column. This causes the electrons stored in the

Figure 1: The operation of a CCD may be likened to a "bucket brigade," reading out electrons pixel by pixel.

receptor layer to flow towards the charge in a follow-the-leader type fashion, one pixel after the other (see Figure 1). At the corner of the CCD is an amplifier, which reads the current carried by the electrons in each pixel. By measuring the charge passing the amplifier as a function of time, the user can determine the number of photons collected at any place on the CCD.

Of course, CCDs are not perfect. If an extremely bright star is imaged onto pixel of the CCD, so many electrons may be ejected out of the donor layer that the receptor layer becomes full and actually resists additional electrons. When this happens, the electrons can "bleed" from one pixel to the next, causing the charge to be distributed over an entire column. Another limitation comes from the process of translating the *analog* charge to a *digital* readout: many systems have some maximum digital number, which cannot be exceeded. (For example, if the analog-to-digital converter has a maximum capacity of 15 bits, the largest number that can be recorded by the device is 32767 (2^{15} - 1). Most electronics get around this to some extent by binning (essentially recording only every n^{th} electron), but obviously, some information is lost in the process.

There are a host of other oddities and problems associated with CCDs which I will not detail here. (Be assured, however, that if you stay in astronomy long enough, you'll encounter all of them.) For further information on CCDs, I refer you to the article by Kristian and Bloucke, (1982), *Scientific American*, **247**, 66.

3 Photometry

This lab concerns itself with determining the distance to a star cluster via the technique of *main sequence fitting.* In the previous lab, you discovered a relationship between the intrinsic brightness (absolute magnitude) of a star and its temperature, or color. Most hot stars are intrinsically bright; most cool stars are faint. As a result, when plotted in a color-magnitude diagram, the overwhelming majority of stars fall on a line. We call this line the *main sequence.* The color-magnitude diagram you plotted last time is called the *HR diagram* after the pioneer astronomers Ejnar Hertzprung and Henry Norris Russell.

The existence of a main sequence gives astronomers a powerful tool for estimating the distances to stars. Stellar parallax is exceedingly difficult to measure and, as you have discovered, can only be applied accurately within $\sim$ 20 pc of the Sun. However, it is easy to determine the color of a star — just measure the apparent magnitude of the star in two different filters (say, B and V) and subtract them. If the star is like most stars and is sitting on the main sequence, the color of the star uniquely specifies its absolute magnitude. In order to get the distance to the star, one just compares the star's apparent and absolute magnitudes by applying

$$(m - M) = 5 \log d - 5. \tag{1}$$

Thus, the observed star is "fit" to the main sequence.

Of course, not all stars are on the main sequence, so if you use this technique on some randomly chosen star, there's a chance that you will get the wrong answer. (In fact, as you will learn later on, many of the stars visible to the naked eye do *not* lie on the main sequence — these stars are rare in space, but are so intrinsically bright, they can be seen at extremely large distances.) However, because all the stars of a star cluster are at the same distance, when you measure the entire group, and plot apparent magnitude versus color, the non-main sequence stars become obvious.

4 The Theory of Photometry

Photometry of astronomical sources can be a complicated progress, involving many steps and corrections. Fundamentally, however, it is pretty simple. For the sake of completeness, I will sketch out the theory of what's happening.

Consider the light coming from a star like the Sun. This star will put out approximately 4×10^{33} ergs each second. (Again, to avoid using big numbers, astronomers call this amount of light 1 solar luminosity, or $1L_\odot$.) If there were no losses anywhere, the flux (power per unit area) we would measure on the earth would be

$$F_o = L/4\pi d^2, \tag{2}$$

where d is the distance to the star. However, in practice, we never measure this amount of light. The reasons are several.

The first place that light can be lost in space. Littered about in interstellar space are tiny dust grains that absorb light, much like the smog in Los Angeles. In some directions, like towards the center of our Galaxy, the dust is so thick that is extremely difficult to see through; in other directions, it is barely noticeable. Still, it is a rare observation that is not affected in some way by interstellar extinction. For now, let's call the fraction of light making its way through the interstellar dust, f_{ISM}. For the purposes of this experiment, we will assume that interstellar extinction is small, *i.e.,* $f_{ISM} \approx 1$.

After traveling through space, the light from the star reaches the Earth's atmosphere. Again the light is attenuated by dust scattering. The amount of this extinction depends on two things: how much dust (and what type of dust) is in each cubic centimeter of air, and how much air the starlight is going through. As you can see in Figure 2, if a star is being observed on the zenith, the amount of air the light is going through is relatively small (1 *airmass*). However, if the star is on the horizon, the light path goes through several airmasses. Because of this, the light we receive on the earth is less than the amount of light reaching the top of the earth's atmosphere by

$$F = F_o e^{-\kappa \mathcal{M}}, \tag{3}$$

where κ is an extinction coefficient depending on the amount of dust in the atmosphere, and $\mathcal{M}$ is the airmass.

After the light makes it through the atmosphere, it hits the telescope. No mirror or lens is perfect: even with a well polished mirror, $\sim 20\%$ of the light might be lost. If a telescope has two or three reflecting surfaces, only 50% of the light might make it through. Let's call the fraction of light making it past the mirrors and lenses f_m. Next, the light encounters a filter, possibly the blue (B) or visual (V) filters you encountered in the previous lab. No filter is 100% transparent, even within its acceptance bandpass. Hence some fraction of light will be lost reflecting off the filter. Let's call the fraction of light making it through the filter filter f_f. After the filter, there is, finally, the detector. As implied earlier, detectors have a wide range of efficiencies. Photographic plates can record 1 or 2% of the light, photomultiplier tubes (devices based on the photoelectric effect) can detect up to 25% of the light, and CCDs may be 90% efficient. Let's call the fraction of light recorded by a detector f_d.

Although we may think we know something about the numbers given above, in practice, we probably do not know them precisely. Furthermore, all the fractions, f_m, f_f, and f_d, as well as the atmospheric extinction coefficient, κ, and the interstellar extinction, f_{ISM}, are wavelength dependent, hence the the percentage of blue photons detected will not be the same as that for red photons. In addition, blue light carries more energy than red light, but when detected, both only produce one electron in the detector.

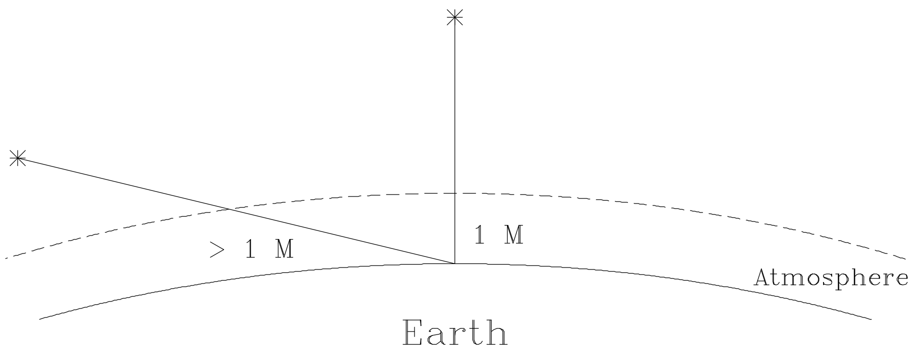

Figure 2: Change of Airmass with zenith distance.

How do astronomers handle these problems and inefficiencies? It's really quite simple. Suppose we define F_o as the monochromatic flux of a star hitting the top of the Earth's atmosphere. (The units of this number in the mks systems will be joules-m^{-2}-sec^{-1}-Å^{-1}.) The number of counts a detector on the ground will detect is

$$\text{Counts} = F_o\, t\,(\frac{\pi D^2}{4})(\frac{\lambda}{hc})\Delta\lambda\, e^{-\kappa\mathcal{M}} f_{ISM}\, f_m\, f_f\, f_d, \tag{4}$$

where t is the exposure time, D is the diameter of the telescope, c the speed of light, h, Planck's constant, λ, the mean wavelength of light transmitted by the filter, and $\Delta\lambda$ is the filter bandpass, *i.e.,* the width of the filter. (Confirm by dimensional analysis that this does, indeed gives counts.) Dividing both sides by the exposure time and writing in terms of magnitudes, (4) becomes

$$-2.5\log\mathcal{C} = m - C + 1.086\kappa\mathcal{M} - 2.5\log(\frac{\pi D^2}{4}) - 2.5\log(\frac{\lambda}{hc}) - 2.5\log\Delta\lambda(f_{ISM} f_m f_f f_d), \tag{5}$$

where $\mathcal{C}$ is the count rate. Now, presumably, f_{ISM}, f_m, f_f, f_d, D, $\Delta\lambda$ and κ will not vary over a single night of observation, so we can combine all the unknowns to yield,

$$m = -2.5\log\mathcal{C} + A\mathcal{M} + K, \tag{6}$$

where $A = 1.086\kappa$. This is now a fairly simple equation. To solve it, suppose that, in addition to observing your unknown (program) objects, you also observe a *standard star* with a known apparent magnitude. Suppose also that you observe this star several times during the night, through several different air masses. In this case, since you know m, and K is a constant, you can solve for A by noting how the number of counts you detect decreases as the airmass increases. Moreover, at the same time you solve for A, you can solve for K. When this is done, you know everything, and you can apply A and K to all the observations of your program objects, to transform the observed count rate into apparent magnitude.

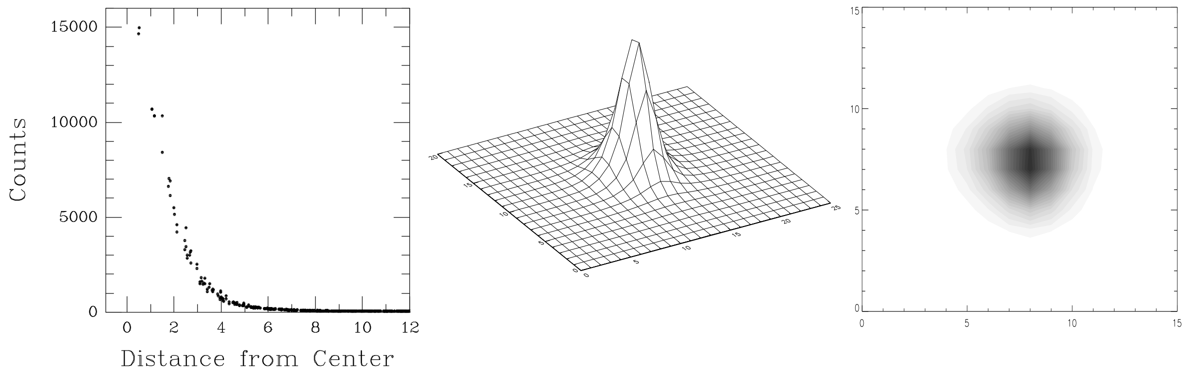

Figure 3: Three different representations of a point spread function.

5 Photometry with CCDs

The stars are so far away that for all intents and purposes they are point sources. Yet, as you have no doubt observed on any astronomical photograph, the stars are not infinitely small points of light. Rather, their light is spread out over some area. The cause of this (at least on Earth) is the atmosphere. As the light from a star makes its way through our atmosphere, it gets refracted in all different directions by cells of air with different temperatures. The result is a slightly blurred image, and the effect is called "seeing." A typical image profile is shown in Figure 3. Typically, at a good astronomical sight, 90% of a star's light may fall within 1 arcsecond. At the best sites, this light may be condensed into ~ 0.4 arc sec. The image profile of a point source is called a *point spread function,* or PSF.

When one performs stellar photometry with a CCD, one measures the light from a star within some radius. To get all the light from the star, this radius has to be large — in fact, no matter how large an aperture you choose, there will be some light falling outside of it. However, by the time one goes five or six times the *full-width at half-maximum* (FWHM for short) of the PSF, the amount of light the observer is missing is inconsequential.

If detectors only recorded starlight, that would be the end of the story. You would choose a large aperture around each star, sum the number of counts within the aperture, divide by the exposure time of the image, take the log, and multiply by -2.5. The result would be a raw, instrumental magnitude. All you would then have to do is find the constants that relates this raw magnitude to apparent magnitude. Unfortunately, life is not quite so easy. In addition to measuring starlight, detectors also measure background light from the sky. (This background may be anything from the glow of many faint background stars and galaxies to scattered moonlight.) Thus, before any calculations can be performed, this contribution must be removed, via subtraction. To do this, normally an estimate of the sky counts per pixel is made from an annulus surrounding each star, and this number is subtracted from each pixel in the stellar aperture. But this leads to another problem. The sky can never be estimated perfectly, and, if one chooses a large aperture for the photometric measurements, a small error in the sky estimate can lead to a relatively large error in the measured magnitude. (Think about it — a large aperture contains many pixels, and if there's a small error in performing the sky subtraction, the contributions of these pixels will add up.) The only way to avoid this problem is to choose small apertures for the star measurements. You might think that such a solution is terrible, since it might exclude a sizeable percentage of the light from each star. This is true, but for most telescopes and detectors, the point spread function is identical for every star. So, once again, as in the above section, if you choose a small aperture, you will be excluding the same fraction of energy from each star, and this number will be buried in one of the many factors of equation (4). Thus, no harm done.

6 The lab

In this lab, you will be presented with CCD images of the outer parts of the globular cluster 47 Tucanae, taken through the B and V filters with the Danish 1.5-m telescope at the European Southern Observatory on La Silla, in Chile. You will use these data to measure the V magnitudes and $(B-V)$ colors for as many stars as possible in the field. You will then combine these measurements with the HR diagram information of the previous lab, and determine the distance to this star cluster.

Two words of warning. First, a large amount of CCD photometry is bookkeeping. While the programs here keep track of many of the variables for you, you may still run into trouble if you attack the lab haphazardly. So stay organized. Second, the programs used in this lab are fairly complex, and may sometimes crash. If that happens, remember to type **retall** before continuing.

6.1 Procedure

1. To perform this lab, you will need to new IDL programs, one to display an image of a star field, called **display.pro**, and another to perform photometry, **phot.pro**. Copy them over to your directory as you have in the past, *i.e.,* by typing

   cp /home/faculty/rbc/display.pro   display.pro

   and

   cp /home/faculty/rbc/phot.pro   phot.pro

(If you are already in IDL, precede these commands with a dollar sign.) Now introduce IDL to the programs by typing

.run display

and

.run phot

Also, to make life easier, use the mouse to move your IDL window down towards the bottom of the screen. (You'll find out why later.)

2. Up until now, the data files you have come in contact with have been in ascii format, and could be viewed by eye. Obviously, these files are extremely useful, but they also take up a relatively large amount of space on disk. Consequently, image files, such as you will be working with here, are usually not stored this way. Instead, the data files this week are in a format called FITS (Flexible Image Transport System), which is a standard in astronomy and used by every observatory in the world. A FITS file cannot be typed to the screen (*i.e.,* you can't see the numbers with **$page**) and cannot be read with the commands of previous labs. IDL, does have its own FITS-reader, however.

3. The two data files you will need to read are **/home/faculty/rbc/47tuc_b.fits** (a B frame), and **/home/faculty/rbc/47tuc_v.fits** (a V frame). Choose one and read it with the command

   data = **readfits**(*filename*)

   where *data* is the array you are reading into, and *filename* is the FITS file, which MUST BE ENCLOSED IN QUOTE MARKS.

4. You can now use the **display** command to view your image. The command takes 4 arguments, but only the first is mandatory; the others are optional. The format for the command is

   display, *data, z1, z2, code*

   The variable *data* is obviously the array you defined with the **readfits** command. The last variable, *code* is also straightforward: it is either, **'p'** or **'n'** (ENCLOSED IN QUOTES!) depending on whether you want to display the picture as a positive, with white stars on a black sky, or as a negative, with black stars on a white sky. (Astronomers almost always work with negatives, but I've left the choice to you.) The variables *z1* and *z2,* however, need some explaining. CCDs usually have a very large *dynamic range.* That is, they can record a large range of numbers, typically from 0 to 32767. Since this is much more than an image display can handle, the computer has a problem: how to map 32767 numbers into a video monitor with 255 shades of grey. Does it do it linearly and have one shade represent all the numbers from 0 to 127, a second shade represent the numbers from 128 to 255, *etc.*, all the way up to 32767? This can be pretty inefficient, especially if almost all your numbers are small, and only only one or two pixels are near maximum. Another possibility is to assigned one shade to 0, the second shade to to 1, the next to 2, *etc.*, all the way up to 255, and have all numbers bigger than 255 treated as if they were 255. (In other words, 0 is black, 255 is white, and anything whiter than white is still white.) Obviously, this might work if most of the pixels are small, but if the values are large, you'll be left with a blank screen. Consequently the only way to ensure a good display is for the user to manually set the minimum (*z1*) and maximum (*z2)* limits for display. If you omit *z1* and *z2,* **display** will attempt to assign them reasonable values, but to see the faintest, or study the brightest stars, you will have to experiment on your own. Play with display and compare what you see on the screen to the finding chart given in this writeup. Make sure you get the orientation of the field correct.

5. You can now run the photometry program by typing

   phot, *data, star_radius, inner_sky_radius, out_sky_radius, filename*

   Here, *data* is again the data array, *filename* is an optional output file to record your measurements which MUST BE ENCLOSED IN QUOTES, *star_radius,* is the radius within which the counts from the star are summed, and *inner_sky_radius* and *outer_sky_radius* define the annulus within which the level of the background sky is determined. Normally, *star_radius* should be chosen to be slightly larger than the full-width half-maximum, and the sky annulus should be defined to be well outside this region. Run **phot** first without the *filename* option. Put your cursor on a fairly faint star and click the mouse button. The program will plot the number of counts in the CCD pixel *vs.* the position, both as a radial profile and a surface plot. It will also ask you for a comment. (For now, you can just hit a carriage return, but later you might want to use this feature.) Now choose a slightly brighter star. Note that the PSF is identical, except for the scale. Now choose and extremely bright star. If the star is bright enough, you will see that the center of the star is saturated over the maximum number of counts that can be recorded (approximately 25,000). These stars cannot be measured.

6. When you run the program and click the mouse button, the program will 1) find the center of the star identified with the cursor, 2) count up the total number of counts within your *star_radius,* 3) estimate the sky by finding the median value of the pixels between *sky_inner_radius* and *sky_outer_radius,* 4) subtract the sky from each pixel within the *star_radius,* 5) compute a raw instrumental magnitude via

   $$m = -2.5 \log \mathrm{Counts} + K \tag{7}$$

   and 6) output the results to the screen and (optionally) to your output file. Experiment with **phot** until you feel confident about how it is working and what aperture radii are reasonable. Then restart the program, giving it an output *filename,* and measure as many stars as possible. After you are done, you can view this file with the **$page** command. (Note: **phot** will allow you to stop and start and often as you wish. If you give it the same *filename,* it will just append the new data to the old file; if you give it a different *filename,* it will open up a new file.

7. Repeat steps 3 through 6 with the data from the other frame. Store the results in a different file. At this stage, you may also want to edit your data files, in order to delete bad measurements.

8. You now have two sets of stellar photometry, one for the B frame and one for the V frame. However, unless you were extremely careful or have a fantastic memory, the order of the stars in your B file will not be the same as that for the V file. To help you match up the two files, type

   .run /home/faculty/rbc/starmatch

   This program will prompt you for the names of your two output files and ask you to give a tolerance, which it will use to match up the coordinates. (You need this because the two images of 47 Tuc do not exactly line up. To prove this to yourself use **display** and **phot** to measure the position of the same star in both frames. They will be different.) The **starmatch** program will use the tolerance you put in to try and match a star at a given x and y in one file to a star at $x \pm tol$ and $y \pm tol$ in the other. As output, **starmatch** produces three new variables: a 5-column array called **bdata** that contains the numbers of the B-band file, a similar 5 column array named **vdata** that holds the values of the V file, and a 4-column array called **appmag**, that contains a) the index of the star in B-file in column 0, b) the star's raw B magnitude in column 1, c) the index of the corresponding star in the V-file in column 2, and d) the star's raw V magnitude in column 3. In addition, **starmatch** also prints **appmag** to the screen and tells you how many stars were not matched in the two files.

9. Until now, you have been working with raw, instrumental magnitudes, (*i.e.,* you have ignored the airmass and constant term in equation 6). To place your instrumental magnitudes on a standard system, you will need to compare your measurements to those of some standards. Fortunately, there are two stars with known magnitudes on your frames: these are identified on the finder at the back of this lab, and tabulated at the end of this section. To use these stars, consider that equation (6) is really two equations,

$$B = -2.5 \log \mathcal{C}_{\mathcal{B}} + A_B \mathcal{M} + K_B$$

$$V = -2.5 \log \mathcal{C}_{\mathcal{V}} + A_V \mathcal{M} + K_V. \tag{8}$$

   and the airmass through which you observed your standards is exactly the same as that of your program stars. (After all, they are on the same frame!) So, for each filter, compare the standard stars' instrumental magnitudes to their true magnitudes and compute $A\mathcal{M} + K$. You can then use this number to calculate the apparent B and V magnitudes for all the stars you have measured.

10. Finally, make a color-magnitude diagram using your measured apparent magnitudes. Compare this to the HR diagram of the previous lab. What distance modulus $(m - M)$ is required to make the two diagrams overlap? What is the distance to the cluster?

    As mentioned above, the presence of interstellar dust can complicate matters for an astronomer. Interstellar dust absorbs light at all wavelengths, but is much better at absorbing blue light than red light. As a result, objects seen through dust will appear redder and dimmer. The ratio of *total absorption* in V to reddening $(E(B-V))$ is about 3. How will computed distance change is there is, say, 0.5 mag of V extinction towards 47 Tuc?

Star A: $V = 15.31 \quad (B-V) = 0.82$
Star B: $V = 16.86 \quad (B-V) = 1.24$