Cephied Variables and distance determination to NGC4258

Divij Gupta

10628702

School of Physics and Astronomy

University of Manchester

Second year laboratory report

May 2022

This experiment was performed in collaboration with *Elisabeth Peak*, *10654157*

Abstract

This paper aims to find the distance to spiral galaxy NGC4258 using data from the Hubble Space Telescope taken in both the ultraviolet and infrared spectrum. The procedure used Python to analyse the ultraviolet raw data given a set of Cepheid candidates and used a known period-luminosity relationship to extract the absolute and apparent magnitudes for each Cepheid. These were then used to find a distance modulus which was corrected for interstellar dust using the infrared data. The final value for the uncorrected distance was found as 9.51 ± 0.18 Mpc while the correction reduced this to 6.22 ± 0.54 Mpc. The hubble constant was also calculated using a known recessional velocity and was found to be 70.8 ± 6.2 km/s per Mpc.

1 Introduction

Distance determination in Astrophysics is one of the most basic requirements for all experiments and theories and so having an accurate method to do this is essential. One of the first methods used for this distance determination were parallexes which worked well for relatively close objects but given their nature, another method had to be found to look at objects on a much larger distance scale. This is where Cepheid Variables are particularly useful.

Cepheids are blue giants just of the main sequence of stars in a region known as the instability strip in the Hertzsprung Russell diagram. Their key feature is that they pulsate periodically varying in brightness (up to 30,000 more luminous than The Sun) in a very regular pattern, they do not have the expansion and contraction stability of typical main sequence stars. Henrietta Swan Leavitt was the first to discover that the period and luminosity of these stars were directly proportional and follow a period-luminosity law, also known as Leavitt's law. [1] Applying this relationship, the variation is brightness of Cephieds can be fit to a photo-metric light curve to find the period which can then be used the star and hence estimate the distance to the galaxy it belongs to. It must be noted that this distance has to be corrected for interstellar dust which artificially increases the observed distance, details of this are discussed later. [2]

This paper aims to reproduce these steps to find the distance to galaxy NGC4258 using Cepheid Variables in the outer region of the galaxy. The outer region is specifically chosen to both minimise disturbances from the rest of the starts in the galaxy but also allow for as many cepheids as possible to be identified since relatively new stars, like Cehpeids, are found in the spiral arms of a galaxy.

2 Theory

The reason for the periodic pulsation of Cepheid Variables lies in their composition and specifically their amount of He^{2+} and He^{+}, He^{+} is relatively cooler and more transparent while He^{2+} is relatively hotter and more opaque absorbing more energy. From this it can be inferred that say for a given dim Cepheid for which there is initially more He^{2+}, the particles will absorb energy gaining kinetic energy causing the star to expand. This expansion leads to cooling which results in more He^{+} being present, as it is easier energetically to form, which makes the star appear brighter as He^{+} is more transparent. This however causes the star to lose more energy as it radiates over a larger surface area and contract heating it up which allows for more He^{2+} to be formed and continues the cycle. [3]

As mentioned, the distance determination of Cepheids requires two things to be known, its period and its intensity. Details of methods to find both are detailed in subsequent sections but the methodology used after their determination is given here for any one given Cepheid Variable. [2] The errors on all of these quantities are propagated as normal.

To begin with, the period can be used to determine the absolute magnitude of the Cepheid using a calibrated period-luminosity law, an extension to Leavitt's law, which for the purpose of this paper was taken to be:

$$M = -2.760(\pm 0.03)[log_{10}(P) - 1] - 4.218(\pm 0.02) \tag{1}$$

where P represents the period of the given Cepheid. The constant factors are empirically determined from known period-luminosity relations and are in general specific to a set of Cepheids in a given galaxy. The ones used here have been calibrated using the Large Magellanic Cloud, the distance to which has been found through a variety of other methods allowing for this calibration.

The intensity of radiation can be used to find the apparent magnitude of the Cepheid on average using a given calibration for the lens on the HST. It is taken that 7 photons correspond to 1 'count' at each pixel in an attempt to minimise the effect of random errors and that an apparent magnitude of $m_2 = 22.75$ corresponds to $F_1 = 1000$ where F_1 is the counts. Ultimately this gives the relations between observed counts and apparent magnitude to be:

$$m_2 = -2.5log_{10}(F_2) + 30.07 \tag{2}$$

where F_2 represents the intensity of the Cepheid in terms of counts. While the the number of counts may vary from day to day, due to the logarithm all readings are roughly the same and to get one value for the apparent magnitude of a given Cepheid, a weighted mean is taken.

Once a single value for both quantities a distance modulus, μ, can be calculated using:

$$\mu = m - M \tag{3}$$

where m and M are the apparent and absolute magnitudes respectively. From this a distance, d, can finally be calculated using:

$$d = 10^{\frac{\mu+5}{5}} \tag{4}$$

where μ is the aforementioned distance modulus.

This is however an overestimate and at this point it is important to consider the effect of interstellar dust which by absorbing some of the radiation makes the Cepheid dimmer that it should be. Correcting for this is known as dust extinction however it is not well documented for galaxy NGC4258 and so the formula applied here is a somewhat crude correction to the distance modulus of the form:

$$\mu_{true} = \mu_i - R|\mu_i - \mu_j| \tag{5}$$

where μ_i and μ_j are the distance modulii for two wavelengths respectively [4]. The R factor is known as the total-to-selective extinction ratio between the two wavelengths of light as is taken to be 2.45 [2] for the data given between ultraviolet and infrared light without an error. Using this 'true distance modulus', a 'true distance' to the Cepheid can then be found using equation 4.

Finally the Hubble constant can also be found using:

$$H_0 = \frac{v}{d} \tag{6}$$

where v is the recessional velocity of the galaxy found through Doppler shift and d is the aforementioned distance to the galaxy. For NGC4258, the recessional velocity was given to be 440 km/s.

3 Apparatus

There is a minimal amount of physical apparatus used in this experiment given its nature and most of it is analysis which has been done using Python. The specific non-standard packages used are given in-text where needed. To initially analyse and calibrate the Python script, SAPImageDs9 was used which is a standard data visualisation program used in astronomy. It is important to acknowledge where the data used here comes from.

Observations of galaxy NGC4258 come from the Hubble Space Telescope (HST) as it removes any effects from the atmosphere. The specific HST instrument used is the Wide Field Planetary Camera 2 (WFPC2) with 4 CCD chips, 3 large and 1 small. The data was collected over a series of 12 (non-concurrent) days in the 'u' band (ultraviolet) and over 3 (non-concurrent) days in the 'i' band (infrared), the later of which is used to correct for interstellar dust. It should be noted that the type of radiation is close to visible light in both cases. From the 4 chips, only the 3 large chips are examined for Cepheid variables one of which was found to contain none. [2]

4 Experimental procedure

As mentioned at the start of Section 3, this experiment was done primarily on python and this section will focus on the broad techniques that were used , the specifics of the code itself has been neglected. The dates for the 12 days of measurement were given as Julian dates relative to November 17th 1858 with the first and last values for the date being 50923.3 and 51281.6 Julian days respectively. [2]

Figure 1 shows the raw u-band data from the two chips for the first day, the i-band data is similar with different relative intensities and is not shown here. From this 15 Cepheid total were identified although the exact procedure for the determination of a Cepheid variable is not given here. Approximate coordinates for these Cepheids had been previously tabulated and for the purpose of this paper were taken as fact with no cross-check. [2] From this intensity data for each Cepheid was collected and the period was found so that the distance could be calculated according to the theory in Section 2. The first 2 subsections focus on the u-band data to get an initial value while the last includes i-band to correct for dust.

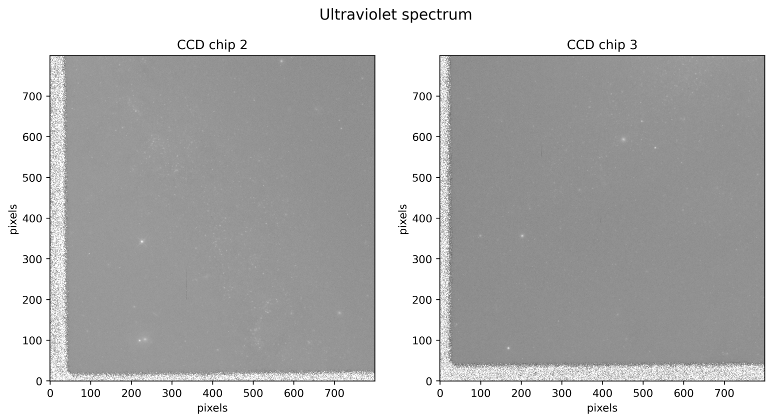

Figure 1: Raw data from the 2 relevant CCD chips in the ultraviolet spectrum on the first day of measurement. The image is 800x800 pixels.

4.1 Intensity data

For any given isolated Cepheid, the intensity of radiation for was found by drawing three regions around it. The inner ring was as small as possible to encompass just the Cepheid, the outer ring was as large as possible (up to a reasonable limit of roughly 10 pixels in radius) to encompass both the Cepheid and the background without including any other stars and the middle ring was slightly larger than the inner ring. These were first drawn manually for all 15 Cepheids for one day's worth of data and then used as a basis to overlay the regions for all other days. This was done to somewhat automate the process while still keeping the benefit of hand-drawn regions. The Python package used for this was Astropy Regions.

It was found that through the different days, the Cepheids shifted by a couple pixels in a random direction and so to correct for that the code re-centres the three rings on the brightest pixel in the innermost ring. While there are various other methods that could be used such as moving the regions according the the movement of a reference pixel, this was visually found to be a better solution. The drawn regions for all 12 days for the first Cepheid is given in Figure 2.

From this the sum of counts for each pixel within any given region was found as well as the area of the region. This was used to remove the background from the innermost region to get a value for the intensity of the Cepheid alone. The procedure for this involved finding the average background intensity between the 2 outer red rings, multiplying this by the area of the inner ring and subtracting this value from the sum of counts in the inner ring. As a crude way to remove any outliers, any pixel in the background region whose value was greater than the maximum value of the innermost region was masked. The python package used for all this was Aperture Photometry and data was collected for all Cephieds for all 12 days.

It has to be noted that the region areas used here are continuous despite the raw data being discrete, this is because each pixel is divided into subpixels automatically by the package so that a more accurate measurement is possible. The error on the region area was for this reason ignored but the error on the sum of

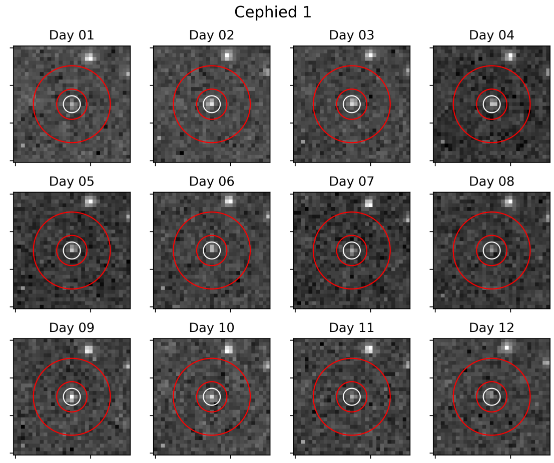

Figure 2: Data for the first Cepheid for all the days it was observed showing the three separate rings. The red rings are the outer two regions used to correct for background and the white inner ring is used to collect the intensity of radiation.

counts was taken as $\sqrt{sum}$ by estimating the counts to be in a Poisson distribution and was propagated accordingly to find the error on the 'true counts'.

4.2 Finding Period

Once the raw data was extracted from each Cepheid, the period of each Cepheid could be found by plotting the intensity count against the date, in other words plotting it's light curve. One way to find the period would be a visual estimate but in the interest of accuracy another method was adopted which is known as phase folding and distance minimisation.

The basic idea was to first fold the values around the point which corresponded to minimum intensity as this corresponded to the start of the light-curve's expected saw-tooth pattern. Then various periods would be trialed and for any given period the points whose days were greater than the period would be mapped back accordingly leaving a pattern over only one period. The point to point distance would then be calculated for every period guess and the one with a minimum value was considered to be the 'best fit' period for that Cepheid. It should be noted that the point to point minimisation used scaled normalised values otherwise the intensity counts would always dominate. [4] The aftermath of this for Cepheid 14 can be seen in Figure 3.

Although the plot is not saw-tooth like, partly due to the scarcity of the data points available to use, the periods seamed reasonable and so this was acceptable. Due to the wild variation, the error on each period is visually estimated as half the width of the main inverted Gaussian like curve seen in the lower plot of Figure 3. From this a period and error was found for each Cepheid.

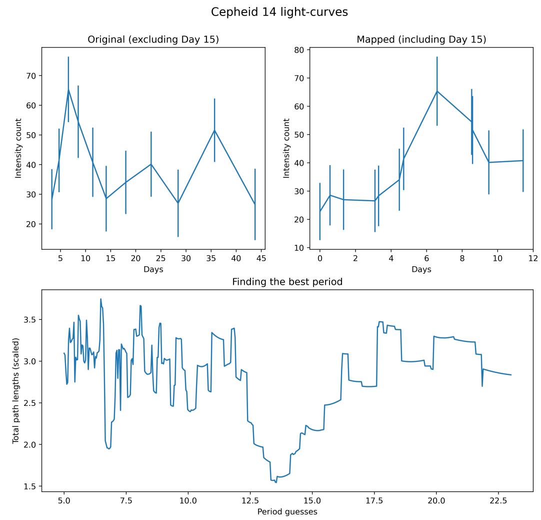

Figure 3: A light-curve for cephied 15 showing the original plot, the phase folded plot for the best period and a plot for how the point to point distance changes for each period guess. In the first plot Day 15 is omitted as it's data was taken at a much later date so the plot would be impossible to read with it. The days in this plot are also given relative to 50920 Julian days for easier visualisation.

4.3 Distance and Dust Extinction

The values for the period of each Cepheid together with all the raw intensity data was then used to calculate the uncorrected distance to the Cepheid as detailed in Section 2. This is done for each of the 15 Cepheids in question. To find an estimate for the uncorrected distance to NGC4258, a weighted mean of all these distances was used and the final value was found to be 9.51 ± 0.18 Mpc.

To correct for dust, the whole procedure detailed in Section 4.1 was repeated for the i-band data. This time however the same regions as before could be used to no manual drawing in Ds9 was necessary. Section 4.2 however was not repeated as only 3 days worth of data makes it practically impossible to find a period within any reasonable degree of accuracy. As a result the same period from u-band data had to be used for i-band data when finding the absolute magnitude with Equation 1. If separate period data could have been found, in theory Equation 1 would also had different factors depending since it depends on the choice wavelength of radiation. This is however only a minor concern as it is far from the greatest source of error as discussed later.

Overall, using the correction prodecure detailed at the end of Section 2, the final value for the corrected distance to galaxy NGC4258 was found to be 6.22 ± 0.54 Mpc. The corresponding Hubble constant was also found to be 70.8 ± 6.2 km/s per Mpc.

5 Discussion

Overall, it can be seen that the final distance found to NGC4258 is reasonable and when compared to values from other research papers and is within 2 σ of all of them. This also reinforces the idea that the dust correction was a necessary step as it resulted in a much more accurate value. It has to be noted however that in every case this value is an underestimate. The Hubble constant on the other hand is much more accurate and is well within 1 σ of any accepted value however this is partly due to the large error. [2]

This large error comes from the fact that almost every step has an associated error. Region determination while good could be further optimised by not relying on circular apertures although the masking detailed in Section 4.1 was a good workaround. Finding the period was by far the largest source of error and while employing phase folding and distance minimisation is effective, the curves are still far from the expected saw-tooth pattern and the point to point distance against period guess plot in Figure 3 shows how fickle the value is. One way to improve this is to fit known light curves directly and employ a chi-squared minimisation procedure however that is beyond the scope of this paper. To optimise the method displayed here, the errors on each point could be included in the point to point domination and a Gaussian could be fit to numerically determine the error on the period. [4] Besides this the other somewhat uncontrollable errors which have been excluded in the error analysis come from Equations 1 and 5 where the choice of numerical factors are up for debate and change depending on the source used. [2]

6 Conclusion

All of this likely means the error is an underestimate on the true error however both the corrected distance of 6.22 ± 0.54 Mpc and the Hubble constant of 70.8 ± 6.2 km/s per Mpc are still in reasonable agreement with accepted results. Increasing the error would only further increase this agreement albeit at the cost of precision and was not done here for this reason. The main improvements stem from much longer analytical procedures with some elements not available for use in this experiment and so the values obtained here are reasonable within the scope of this experiment.

References

[1] H. S. Leavitt and E. C. Pickering, "Periods of 25 Variable Stars in the Small Magellanic Cloud.," *Harvard College Observatory Circular*, vol. 173, pp. 1–3, Mar. 1912.

[2] J. A. Newman, L. Ferrarese, P. B. Stetson, E. Maoz, S. E. Zepf, M. Davis, W. L. Freedman, and B. F. Madore, "A revised cepheid distance to NGC 4258 and a test of the distance scale," *The Astrophysical Journal*, vol. 553, pp. 562–574, jun 2001.

[3] M. Williams, "What are cepheid variables?," Oct 2016.

[4] P. B. Stetson, "On the Automatic Determination of Light-Curve Parameters for Cepheid Variables," , vol. 108, p. 851, Oct. 1996.