Calculating the Period-Magnitude Relationship of Type 1 Cepheids to Find the Distance to the LMC

Mahebub Aalam Khatri
(Northwestern University)

November 19, 2022

Abstract

Type 1 Cepheids are yellow supergiant stars whose luminosities vary over a period of 1 to 100 days and exhibit a period-magnitude relationship. This relationship can be calibrated from stars within the Milky Way whose distances are within the range of parallax measurements and then extrapolated to find the distances to Cepheids in other galaxies. Here, I calculated the period-magnitude relationship in the I-band using 80 Cepheids within the galactic disk of the Milky way. Using parallax measurements from Gaia-EDR3 and lightcurves from OGLE, I found the relationship to be $M_I = (-2.32 \pm 0.73) \cdot \log(P/10) - 5.14 \pm 0.27$ where P is the period in days and M_I is the I-band absolute magnitude. Using this relationship, I found the mean distance to Cepheids in the Large Magellanic Cloud to be 52.1 ± 10.3 kpc.

1 Introduction

Getting accurate distance measurements are central to astronomy as they are necessary for calculating many other attributes of astronomical objects including their size, luminosity, and temperature. The need for accurate distances extends out all the way to measurements of the most distant galaxies which can be used to measure Hubble's constant and in turn, the dark energy content of the universe Ryden (2017).

Unfortunately, no single set of measurements is enough to scale the whole breadth of distances in astronomy. Instead, the distance ladder is used to calibrate measurements to increasingly distant objects using measurements from lower "rungs" of the distance ladder Ryden & Peterson (2017). Gaia, which can measure parallaxes with an uncertainty of just 10 microarcseconds, can only provide measurements for stars up to 2 kiloparsecs away with an apparent magnitude of 17 with 10% uncertainty and stars as far away as the galactic center with an uncertainty of 20% Lindegren et al. (2018); Mignard (2019). This makes parallax measurements good for distances within the Milky Way. However, measuring parallaxes to other galaxies would require parallax measurements with an uncertainty of less than 0.01 microarcseconds, far beyond the capabilities of Gaia and any other instrument planned to date Ryden (2017). Instead, the period-magnitude relationship of Cepheid variable stars can be used to measure the distances to neighboring galaxies.

Cepheid variables are yellow supergiant stars whose luminosities vary over a period of 1 to 100 days and are luminous enough to be seen in nearby galaxies Guidry (2019). Type I (also known as "classical") Cepheids are population I stars which are young and metal-rich, meaning they have an abundance of elements heavier than helium. Meanwhile, Type II Cepheids are population II stars which are older and metal-poor Ryden & Peterson (2017). The two types have different period-magnitude relationships due to the differences in their chemical make up. Type 1 Cepheids tend to be more luminous and can be seen father away which is why they are often chosen to measure distances, as is the case in this paper. Their period-magnitude relationship can be established using the Type I Cepheids that lie within the range of parallax measurements. That relationship can then be extrapolated to find the distances to Type I Cepheids in nearby galaxies Guidry (2019).

Henrietta Swan Leavitt was the first to discover and use Cepheids and their period-magnitude relationship to prove that the Large Magellanic Cloud (LMC) was too far away to be within our galaxy Guidry (2019). This paper aims to repeat Leavitt's original measurements using modern parallax and light curve observations of Type 1 Cepheids.

2 Data

The data for this paper was acquired from the Optical Gravitational Lensing Experiment (OGLE) (which can be found here: https://www.astrouw.edu.pl/ogle/ogle4/OCVS/), as well as Gaia's Early Data Release 3 (Gaia-EDR3) and the Infrared Science Archive's (IRSA) Galactic Dust Reddening and Extinction database which were both queried using the python package astroquery.

OGLE provides a list of all the Type 1 Cepheids in the Milky Way. From this list, I selected all the Cepheids that had both OGLE and Gaia-EDR3 IDs as I needed to access data about the Cepheids from both databases. Then, I picked out Cepheids for which OGLE had I-band data since OGLE has more I-band light curves than V-band. Finally, I selected only those Cepheids which oscillated in the fundamental mode which would make calculating their periods much easier as I would not have to worry about any of the overtones. Then, I used the Gaia-EDR3 IDs of the remaining Cepheids to query Gaia-EDR3 for those stars' right ascension (RA), declination (DEC), parallax, and parallax error.

I converted the parallax measurements to distances using the formula $d = 1000/\pi_{\text{mas}}$ where the 1000 in the numerator accounts for the conversion from milliarcseconds to arcseconds since Gaia gives parallax in milliarcseconds. Then, I filtered out all of the Cepheids that had a distance error to distance ratio of greater than 0.1 to curb errors, especially as some Gaia data has errors for distances larger than the distances themselves. I also removed Cepheids that had negative distances. This filtering left me with 101 Cepheids in the galactic disk.

To get extinction data, I used the RA and DEC provided by Gaia for each Cepheid to look up the corresponding extinction table from IRSA and extracted the extinction in the CTIO I-band which approximates OGLE's I-band filter. I knew that all of the remaining Cepheids had distances too short to be within the bulge of the Milky Way so the extinction for Cepheids in the direction of the bulge would be over-corrected. Therefore, I removed Cepheids that had an extinction greater than 5 magnitudes to approximate the cutoff for extinction towards the bulge. This reduced my list of Cepheids to calculate the period-magnitude relationship to 80.

OGLE also provides data for all of the Cepheids within the LMC and a separate identification file that provides their RA and DEC. After downloading the data and matching OGLE's LMC Cepheids to their corresponding RA and DEC, I was able to get the extinction for each Cepheid from the IRSA. For this data however, since the LMC is not facing Milky Way's galactic bulge, I did not need to do any filtering on the extinction.

3 Methods and Analysis

The analysis was done in a python jupyter notebook using mainly the *astropy*, *astroquery*, and *pandas* packages. I also used the *uncertainties* package to help calculate the uncertainties.

After the initial process of filtering out the Cepheids, I was left with a pandas data frame that had 80 Cepheids with their corresponding Gaia-EDR3 and OGLE IDs, RA and DEC, and distance and distance error measurements. The first step after this was to calculate their periods. For each Cepheid in the data frame, I loaded in the corresponding OGLE data that contained a table of observations with the time of observation, I-band magnitude, and I-band magnitude error. An example of this data for OGLE-GD-CEP-0028 is plotted in Figure 1, left. I passed that data to the astropy function `LombScargle` and used `autopower` with a minimum frequency of 1/24 days (corresponding to 1 hour) and a maximum frequency of 50 days. This returned a list of frequencies and their corresponding powers on the Lomb-Scargle periodogram (L-S). One over the frequency (i.e. the period) of OGLE-GD-CEP-0028 is plotted against the corresponding power in Figure 1, right. From here, to get the period, I took the inverse of the frequency that had the corresponding index of the maximum power. To confirm this method, I plotted the phase folded data (Figure 2) and saw that it indeed gives an accurate measure of the period, returning a sinusoidal curve with the slow increase and rapid decrease in magnitude that is characteristic of Cepheids.

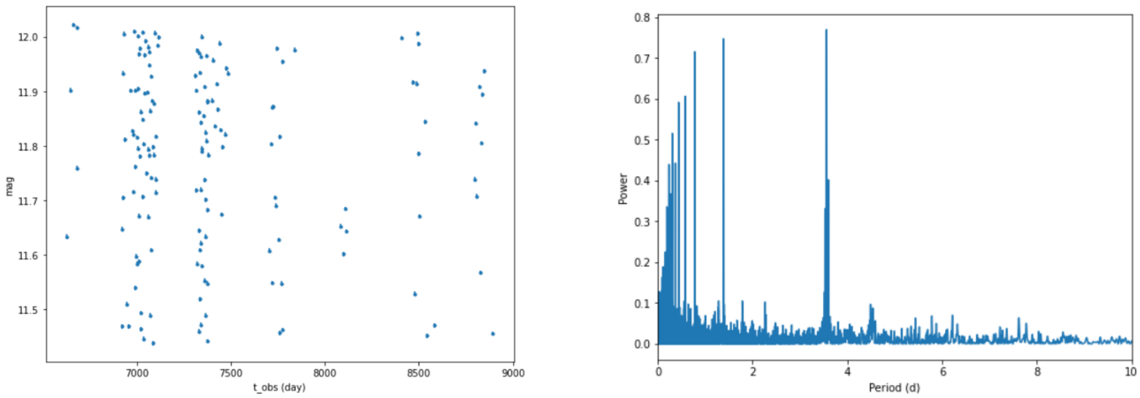

Figure 1: Left) Raw light curve of OGLE-GD-CEP-0028, plotting time of observation vs I-band magnitude with error bars. Right) The output of the Lomb-Scargle periodogram for OGLE-GD-CEP-0028 plotting period vs power. The period with the maximum power is the true period.

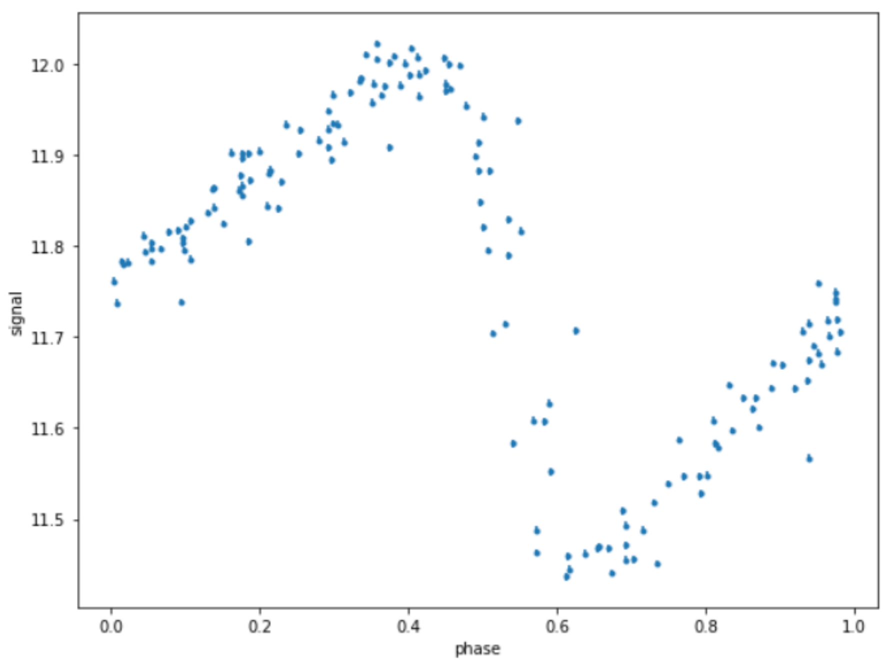

Figure 2: Phase folded plot (I-band magnitude vs phase) of OGLE-GD-CEP-0028 with the period determined by L-S.

For more specifics on the inner workings of L-S, refer to VanderPlas (2018). Now that I had the periods of each Cepheid that had Gaia data, I could work towards calculating their absolute magnitudes. The first step was to get the extinction. For every Cepheid, I used the RA and DEC provided by Gaia to get the extinction table from IRSA's Galactic Dust Reddening and Extinction database and took the extinction in the CTIO I filter which approximates the extinction in OGLE's I-band filter. Finally, with a data frame containing the apparent I-band magnitudes, distances, and extinction for each Cepheid, I could calculate each of their absolute I-band magnitudes using the equation

$$M = m_{\mathrm{obs}} - 5\log_{10}\left(\frac{d}{10\mathrm{pc}}\right) - A$$

where m_{obs} is the mean I-band apparent magnitude, d is the distance in parsecs, and A is the extinction term that I got from IRSA. To find the period-magnitude relationship, I fit a linear line to a plot of Log(period/10 days) vs I-band Absolute Magnitude (Figure 3).

Now that I had a period-magnitude relationship, I could move on to calculating the mean distance to Cepheids in the LMC. First, I created a list of all the Cepheids in the LMC that oscillated in the fundamental mode and connected each Cepheid with its RA and DEC from a separate identifier file. Then, just like for Cepheids in the Milky Way, I calculated the period of each Cepheid in the LMC and got the extinction using its RA and DEC from the IRSA dust extinction database. Then, using the period-magnitude relationship and the calculated period of each Cepheid in the LMC, I determined an estimate for what the absolute I-band magnitude of each Cepheid in the LMC should be. Finally, I could calculate the distance to each Cepheid using the following equation:

$$d = 10^{\frac{m_{\mathrm{obs}} - M - A + 5}{5}}$$

where m_{obs} is the observed mean I-band magnitude, M is the calculated absolute magnitude from the period-magnitude relationship, and A is the extinction that we get from IRSA.

At each step in this process, I calculated the errors using the *uncertainties* package. For every equation that I wanted to calculate errors for, I would provide every variable that had an error associated with it as a `ufloat` (essentially a tuple of the value and the uncertainty or error expressed as a float) to the function `std_devs` which returned the error for that equation.

4 Results

I calculated the period-magnitude relationship (Figure 3) to be

$$M_I = (-2.32 \pm 0.73) \cdot \log(P/10) - 5.14 \pm 0.27$$

where P is the period in days and M_I is the I-band absolute magnitude. Using this relationship, I found the mean distance for Cepheids in the Large Magellanic Cloud to be 52.1 ± 10.3 kiloparsecs. This confirms that the LMC is too far away to be within the Milky Way and is a Galaxy in its own right.

To briefly discuss errors, they are likely heavily under-estimated. I did not have access to errors in the extinction nor did I calculate error in the period. Especially regarding extinction, the IRSA provides extinction along the entire line of sight in a given direction which may have a tendency to over-estimate the extinction since these

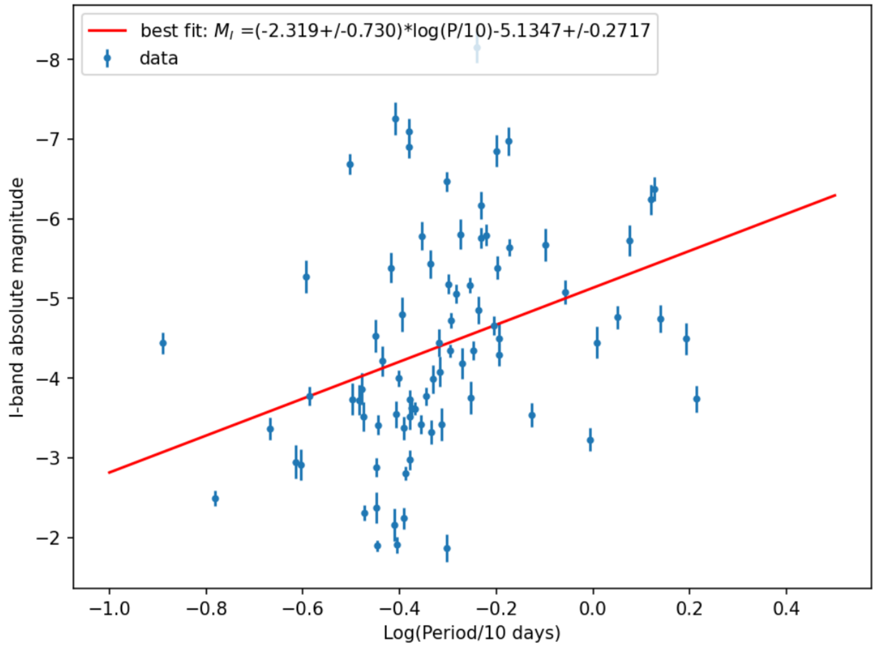

Figure 3: The period-magnitude relationship is represented by the red line

stars are relatively close. And, in general the uncertainty in the extinction may in part be the culprit for how spread out the Cepheids' absolute I-band magnitudes are compared to the best fit. Finally, it is important to consider that after all of the filtering, I was left with only 80 Cepheids to calculate the period-magnitude relationship.

5 Conclusions

Type 1 Cepheids are variable stars that exhibit a period-magnitude relationship. In this paper, I calculated the relationship using parallax data from Gaia and light curves from OGLE. I used L-S to calculate the periods of the stars and accounted for extinction from the IRSA galactic dust extinction database. I found the period-magnitude relationship to be $M_I = (-2.32 \pm 0.73) \cdot \log(P/10) - 5.14 \pm 0.27$. I used a similar process to calculate the periods of Cepheids in the LMC and their extinction. Then using the period-magnitude relationship, I found the average distance to Type 1 Cepheids in the LMC to be 52.1 ± 10.3 kiloparsecs.

In order to get better results, future improvements to this project could incorporate error measurements for the periods and spend more time accounting for and filtering extinction measurements. They could also make the filtering methods more sophisticated in order to retain more than just 80 Cepheids to calculate the period-magnitude relationship. This projected could also be extended to find higher rungs in the distance ladder using the period-magnitude relationship determined here.

References

Guidry, M. 2019, Stars and Stellar Processes (Cambridge University Press)

Lindegren, L., et al. 2018, Astronomy & Astrophysics, 616

Mignard, F. 2019, Comptes Rendus Physique, 20, 140

Ryden, B. 2017, Introduction to Cosmology (Cambridge University Press)

Ryden, B., & Peterson, B. M. 2017, Foundations of Astrophysics (Pearson Education)

VanderPlas, J. T. 2018, The Astrophysical Journal Supplement Series, 236, 16