

# To what extent does the period of classical Cepheid variable stars correlate with their absolute magnitude?

Word count: 2978

April 2025

## 1 Introduction

Before 1912, we could only accurately measure distances within a few hundred light years. No one knew the size of the Milky Way or whether stars existed outside the galaxy. However, in 1908, Henrietta Swan Leavitt published a paper noting a correlation between the period of certain variable stars, known as Cepheids, and their brightness. (Leavitt, [1908](#)) Four years later, she published another paper demonstrating a proper relationship between the star's brightness variation and its luminosity, enabling distance calculation by comparing their luminosity and brightness. (Leavitt & Pickering, [1912](#))

This discovery revolutionized astrophysics, enabling measurements beyond the Milky Way.

## 2 Research question

This investigation aims to explore the research question: “To what extent does the period of classical Cepheid variable stars correlate with their absolute magnitude based on the stars in the Large Magellanic Cloud?”. By analyzing observational data of Cepheid variables located in the Large Magellanic Cloud (LMC), the same region Henrietta Leavitt studied, and then calibrating the data using parallax, this IA will test the validity of the period-luminosity relationship using modern data and regression analysis.

### 3 Background Information

#### 3.1 Terminology

| Term                                       | Definition                                                                                                                                                                                                                                          |
|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Luminosity</b>                          | The energy a star emits, measured in watts ( $W$ ).                                                                                                                                                                                                 |
| <b>Brightness</b>                          | The total amount of energy from a star reaching a unit collecting area, measured in watts per meter squared ( $W \cdot m^{-2}$ ).                                                                                                                   |
| <b>Apparent magnitude (<math>m</math>)</b> | A unitless scalar that indicates how bright an object appears to an observer, defined using an inverse logarithmic scale. (Las Cumbres Observatory, <a href="#">n.d.</a> )                                                                          |
| <b>Absolute magnitude (<math>M</math>)</b> | A unitless measure of a star's luminosity. It equals the apparent magnitude the star would have if viewed from 10 parsecs. Related by: $M = -2.5 \log \frac{L}{L_0}$ , where $L$ is the luminosity of the star and $L_0$ is the luminosity of Vega. |
| <b>Period (<math>P</math>)</b>             | The time a variable star takes to go from one peak magnitude to the next. See Figure <a href="#">1</a> .                                                                                                                                            |
| <b>Extinction</b>                          | The loss of a star's apparent brightness due to interstellar dust between the star and the observer.                                                                                                                                                |

Table 1: Key astrophysics terms relevant to this investigation

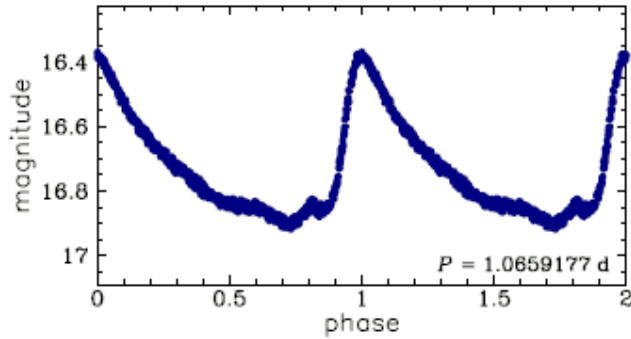


Figure 1: Light curve of OGLE-LMC-CEP-2491 (Soszyński, [n.d.](#))

### 3.2 Cepheid Variable Stars

Cepheid stars are a class of variable stars that pulsate primarily due to the kappa mechanism. When gravity compresses a layer of partially ionized helium in the star's outer envelope, it becomes more opaque, trapping heat and increasing pressure, which in turn causes further ionization that forces the star to expand.

As the star expands, it cools, and helium recombines, increasing transparency. This reduced opacity allows heat to escape, making the star brighter while decreasing internal pressure, which in turn causes contraction under gravity, restarting the cycle. These pulsations exhibit various standing wave patterns, known as modes. In this IA, we will utilize only the fundamental mode due to its simplicity and stability. (Dederick & Jackiewicz, 2017)

Overall, Cepheid pulsations act as a thermal engine, with opacity changes from helium ionization driving periodic expansions and contractions. This makes them predictable and suitable for distance measurements. (Unknown, n.d.) Figure 2 shows an image of one such star.

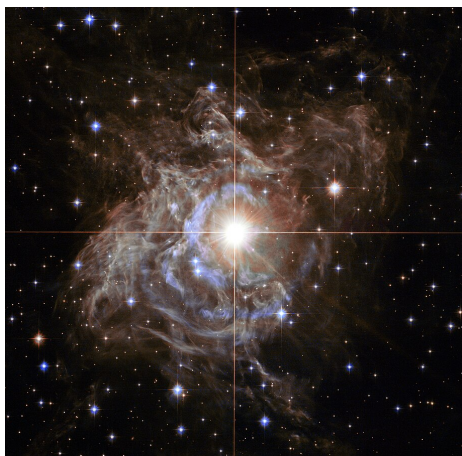


Figure 2: This image is sourced from Wikipedia contributors, n.d. In this video by European Space Agency, n.d.-c, one can see the change in brightness and how this causes an effect called a “Light echo” since the star is surrounded by nebula dust.

### 3.3 Equations and magnitude relationships

A key equation in the IA is the distance modulus formula, which states:

$$m - M = 5 \log d - 5 \quad (1)$$

Where

- $m$  is the apparent magnitude
- $M$  is the absolute magnitude

- $d$  is the distance between the observer and the star.

This formula enables us to calculate any of these three quantities if the other two are known.

The relationship between the luminosity of classical Cepheid variables and their pulsation periods was described in this way by Leavitt, a discovery critical to cosmic distance measurement. (STRUCTURES Blog, [2024](#))

$$M = a \log P + b \tag{2}$$

Where

- $P$  is the period.
- $M$  is the absolute magnitude.
- $a$  and  $b$  are constants that vary slightly across different photometric bands. For this experiment, we'll conduct our analysis based on near-infrared wavelengths and visual wavelengths (I-bandwidth and g-bandwidth, respectively).

In this IA, we'll trace Leavitt's steps using modern tools and data to identify this relation based on assumptions similar to hers.

## 4 Variables

| Type of Variable     | Description                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Independent Variable | Period of the Cepheid variable star ( $P$ ), measured in days. This is the time between two successive peaks in the light curve.                                                                                                                                                                                                                                                                                                                      |
| Dependent Variables  | <ul style="list-style-type: none"> <li>• Apparent magnitude (<math>m</math>) — used in the initial analysis with OGLE stars to establish the period–magnitude relationship.</li> <li>• Absolute magnitude (<math>M</math>) — calculated via parallax and extinction for Gaia stars and used in the final calibrated relationship.</li> </ul>                                                                                                          |
| Controlled Variables | <ul style="list-style-type: none"> <li>• <b>Location:</b> All OGLE stars are in the Large Magellanic Cloud (LMC), ensuring roughly equal distance to Earth and reducing distance-based variation in <math>m</math>.</li> <li>• <b>Star Type:</b> Only classical Cepheid variables pulsating in the fundamental mode are included to ensure the validity of Leavitt’s law. Other types, such as Type II or overtone Cepheids, are excluded.</li> </ul> |

Table 2: Summary of variables considered in this investigation

## 5 Method

### 5.1 Data Source

#### 5.1.1 The OGLE

Due to the nature of this IA and the requirement for heavy equipment to take measurements over several days, it’ll be conducted using secondary data. For the first half of the analysis, the Optical Gravitational Lensing Experiment (OGLE) will be used.

The project maintains one of the largest and most comprehensive catalogs of variable stars in the sky, including the Large Magellanic Cloud (LMC). OGLE

will provide a large sample size of Cepheids in the LMC, along with homogeneous observations from the same telescope, filters, and techniques, which will also be semi-processed, thereby enhancing the robustness of our work.

The data is processed by the OGLE team and made public through their online portal. (OGLE Team, [n.d.-b](#)).

### 5.1.2 Calibration Using Gaia Parallax Data

To generalize the relation between apparent magnitude and period into absolute magnitude and period, we must calibrate it in terms of absolute magnitude  $M$ . This is done using nearby Cepheid stars within the Milky Way, for which precise distances are known thanks to parallax measurements from the Gaia mission.

Gaia is a retired space observatory that measured the positions and parallaxes of stars. Parallax refers to the apparent shift in a star’s position against distant background stars due to Earth’s orbit around the Sun, as illustrated in Figure 3. This effect allows us to determine a star’s distance using the relation:

$$d = \frac{1}{p} \quad (3)$$

where  $d$  is the distance in parsecs and  $p$  is the parallax angle in arc-seconds.

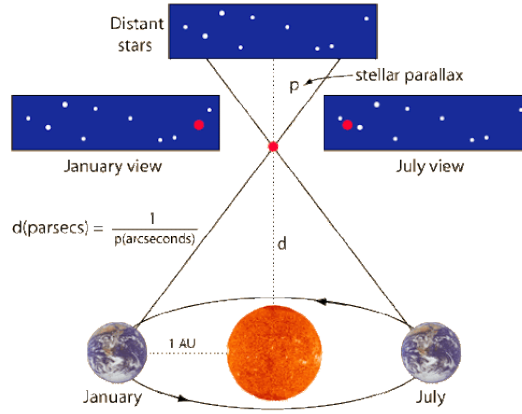


Figure 3: A diagram illustrating stellar parallax. (“Parallax”, [n.d.](#))

Once the distance is known, the absolute magnitude  $M$  of a Cepheid can be calculated from its apparent magnitude  $m$  using Equation 1, the distance modulus. By applying this method to several Galactic Cepheids, we can determine their absolute magnitudes and compare them to their pulsation periods. This allows us to calibrate Leavitt’s Law in terms of  $M$ , making the relationship valid for distance determination across the universe.

The data will be sourced from the ESA Gaia archive website. (“Gaia Archive”, [n.d.](#)).

## 5.2 Justification for LMC

The use of the LMC, as represented in Figure 4, to determine the relationship between apparent magnitude and period is based on an assumption made by Leavitt that those stars were far enough away but close enough together to ignore their relative change in apparent magnitude. Nowadays, thanks to her discoveries, it's known that the LMC is a dwarf galaxy orbiting the Milky Way, located at  $49.59 \pm 0.54$  kiloparsecs away from Earth. This allows us to quantitatively define the effect of assuming all stars in the LMC are at the same distance from Earth. (Pietrzyński et al., 2019)

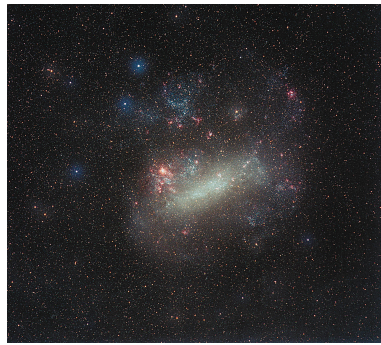
A difference in apparent magnitude  $m$  caused by a difference in distance  $d$  is given by this rearranged distance modulus:

$$\Delta m = 5 \log\left(\frac{d_2}{d_1}\right)$$

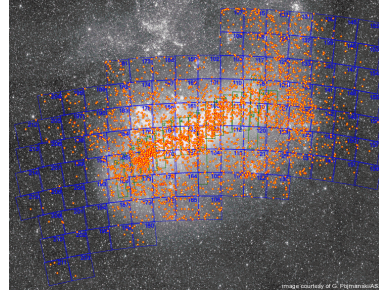
Where  $d_1$  and  $d_2$  are the two positions being compared for their apparent magnitude. The literature Van Der Marel and Cioni, 2001 indicates that the line-of-sight depth of the LMC is  $4.5 \pm 0.5$  kiloparsecs in the thickest parts. To obtain a conservative estimate, we'll be choosing the depth and distances that give the most significant uncertainty:

$$\Delta m = 5 \log \frac{(49.59 + 0.54 + \frac{(4.5+0.5)}{2})}{(49.59 - 0.54 - \frac{(4.5+0.5)}{2})} = 0.27$$

This result shows that, even in the most extreme case, the variation in apparent magnitude across the LMC due to its line-of-sight depth is less than 0.3 magnitudes—a relatively small difference when compared to the range of Cepheid magnitudes analyzed in this investigation. Nevertheless, this uncertainty in magnitude will be taken into account.



(a) Sourced from European Space Agency, [n.d.-b](#)



(b) Sourced from OGLE Team, [n.d.-a](#)

Figure 4: A comparison between an astrophotograph of the LMC and a map showing Cepheid stars within it.

## 5.3 Data Selection

### 5.3.1 For the apparent magnitude to period relationship

Our analysis will focus on classical Type I Cepheids in the LMC. Type I Cepheids are young, massive stars that follow a well-defined period-luminosity relationship as they pulsate in a highly predictable manner. This makes them ideal for verifying the  $P$ - $L$  (Period-luminosity) relation. Other Cepheids, such as Type II Cepheids, are older and have weaker  $P$ - $L$  relationships. Including them would add noise and reduce reliability. The specific Cepheids analyzed were downloaded from the OGLE Cepheids database search form, which allows users to filter Cepheid stars by type, pulsation mode, and location. (“OGLE Cepheids Search Form”, [n.d.](#)) After downloading the 2,483 stars that fulfill the filtered search’s requirements, the data is structured into five columns.

| Column      | Description                                                                                          |
|-------------|------------------------------------------------------------------------------------------------------|
| <b>ID</b>   | A code to identify each star.                                                                        |
| <b>Mode</b> | Indicates the pulsation mode; has been filtered to only include stars pulsating in fundamental mode. |
| <b>I</b>    | Represents the mean I-band apparent magnitude of the star.                                           |
| <b>P_1</b>  | The period of the star, stored in days.                                                              |
| <b>dP_1</b> | The uncertainty in the period measurement.                                                           |

Table 3: Columns included in the OGLE dataset after filtering. Unfortunately, the OGLE catalog doesn’t provide uncertainty measurements for the mean apparent magnitude, as they are considered negligible (Soszyński et al., 2015). Either way, the estimate of uncertainty due to different positions, as calculated in subsection 5.2, serves a similar purpose. Some data had impossible values, such as a magnitude of  $-99.99$ . These placeholder values indicated missing or erroneous measurements and were removed from the final dataset.

### 5.3.2 For the absolute magnitude to period relationship

To calibrate the absolute magnitude-period relation, we require a sample of nearby classical Cepheids for which precise distances, apparent magnitudes, and periods are known. For this purpose, we use data from the Gaia DR3 catalog. Unfortunately, Gaia doesn’t have I-band photometry, so we’ll use G-band instead. One of the main issues with this is that the g-band interacts with dust more than the I-Band, so we’ll require a correction for extinction.

To source the data from “Gaia Archive”, [n.d.](#) the advanced search (ADQL) function was used to retrieve data from the SQL database. The data were filtered to include only classical Cepheids in fundamental mode with parallax signal-to-noise ratios greater than 10 and no empty columns from the requested ones. The result included 344 rows with star data described using the following

columns:

| Column                                                                                                                                                                     | Description                                                                                        |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| <code>source.id</code>                                                                                                                                                     | A unique identifier for each star.                                                                 |
| <code>v.pf</code> , <code>v.pf_error</code>                                                                                                                                | Represent the primary pulsation period in days and its associated uncertainty.                     |
| <code>parallax</code> , <code>parallax_error</code>                                                                                                                        | The parallax angle and its uncertainty, both measured in arcseconds.                               |
| <code>phot.g.mean_mag</code> ,<br><code>phot.g.mean_flux_over_error</code> ,<br><code>ag_gspphot</code> , <code>ag_gspphot_lower</code> ,<br><code>ag_gspphot_upper</code> | G-band photometry data, including mean magnitude, uncertainty from flux, and extinction estimates. |

Table 4: Key columns retrieved from the Gaia DR3 dataset. Using these Cepheids, we can calculate absolute magnitude via parallax and apply the distance modulus. This results in a set of data points of  $\log P$  vs.  $M$ , which will be used to fit a new regression model. The slope from the OGLE relation will be reused, while the intercept will be updated to reflect absolute magnitude  $M$  instead of apparent magnitude  $m$ .

#### 5.4 Data Processing, graphing, and regression

After selecting the data, the OGLE file contains 2,446 Cepheid stars that meet the processing requirements. To identify a pattern, we first graph the relationship between the two key variables: apparent magnitude  $m$  and period  $P$ . A linear regression fit is performed using Python, minimizing the squared differences between the data points and the predicted values. Figure 5 shows the fit. The model's accuracy is evaluated using  $R^2$ , which indicates how well the line describes the variation in the data.

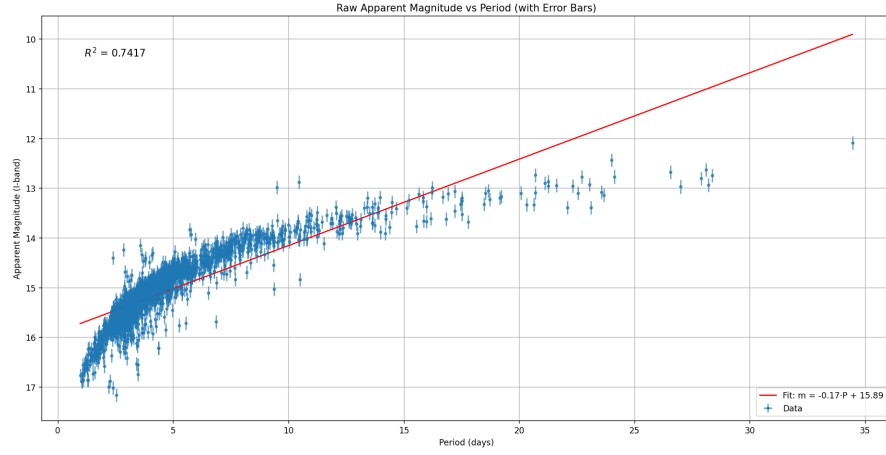


Figure 5: Linear regression fit of apparent magnitude vs. period. The fit yields  $R^2 = 0.74$ , and  $m = (-0.174 \pm 0.002) \cdot P + (15.895 \pm 0.012)$ . While some correlation is evident, the distribution suggests that a better fit may be obtained with a logarithmic transformation of the period. Re-plotting the data as  $m$  vs.  $\log(P)$  significantly improves the correlation, as visible in 6:

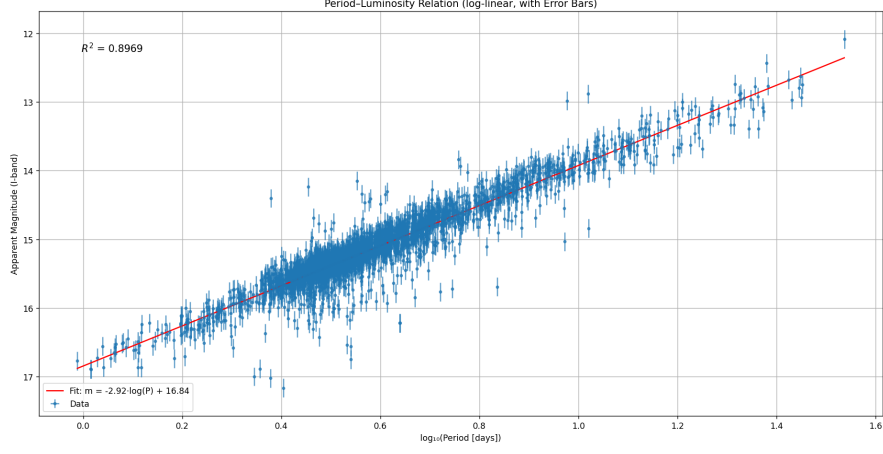


Figure 6: Logarithmic fit of apparent magnitude vs.  $\log(P)$ . The model yields  $m = (-2.92 \pm 0.02) \cdot \log(P) + (16.843 \pm 0.013)$ , with  $R^2 = 0.90$ . While the error bars for  $\log P$  are not visible due to being too small, this higher  $R^2$  value confirms that a log-linear model better represents the relationship between period and magnitude, aligning with Henrietta Leavitt’s original discovery (see Section 3.3 and Equation 2). The relationship shown above is based on the apparent magnitude  $m$ , which depends on distance. However, this IA aims to find a relationship using absolute magnitude  $M$  as it would allow for the measurement of distances in the universe. For this, we can use the Gaia data.

Using the distance modulus 1, one can substitute  $M$  into the equation:

$$m = M + 5 \log_{10}(d) - 5 + A_\lambda$$

where:

- $d$  is the distance in parsecs, calculated using Gaia parallax via Equation 3,
- $A_\lambda$  is the interstellar extinction in the observed band — in the following case, the G-band, so  $A_G$ ,
- $m$  is the observed apparent magnitude in the G-band.

Rearranging the formula, we obtain:

$$M = m - 5 \log_{10}(d) + 5 - A_G$$

This enables us to calculate extinction-corrected absolute magnitudes,  $M$ , for each star in the Gaia dataset. We then combine these with their corresponding periods to calibrate Leavitt’s Law.

Rather than refitting both slope  $a$  and intercept  $b$ , we fix the slope to the OGLE-derived value of  $a = -2.92$ , which is more reliable due to the OGLE

sample's size and uniformity. We then calculate the intercept  $b$  using the Gaia data.

This is achieved by rearranging the Leavitt relation:

$$M = a \log(P) + b \quad \Rightarrow \quad b = M - a \log(P)$$

And then, computing the average:

$$b = \text{mean}(M - a \log(P))$$

This gives a calibrated, distance-aware version of Leavitt's Law:

$$M = -2.92 \log(P) + b_{\text{Gaia}}$$

This two-step method combines the statistical strength of OGLE (for the slope) and the distance-based accuracy of Gaia (for the intercept), resulting in a refined absolute period-luminosity relation.

To compute the distances from Gaia parallaxes, we use Equation 3:

$$d = \frac{1}{p}$$

The final step utilizes this distance and the extinction-corrected G-band apparent magnitude to calculate  $M$ , after which we determine the best-fitting intercept  $b$  using a fixed-slope regression method. This is all accomplished by the Python program, which also calculates the uncertainties. Figure 7 shows the results of the regression fit.

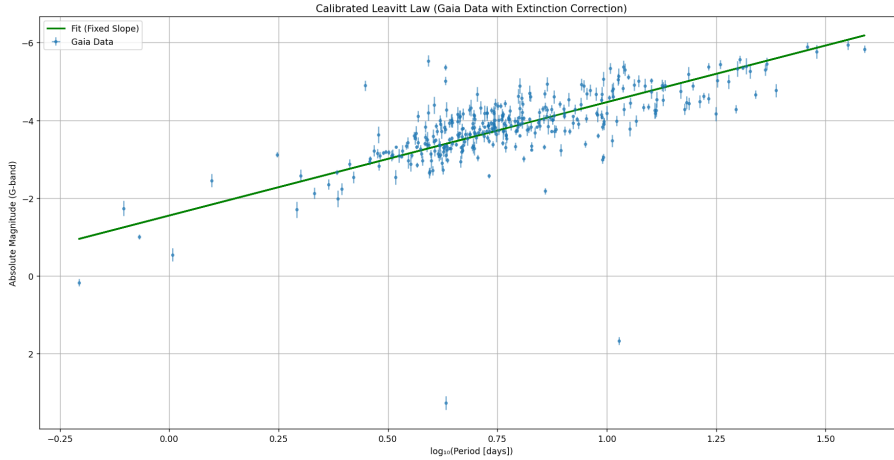


Figure 7: The fit for  $M = -2.92 \cdot \log P + b$ , finding the minimal  $R^2$  at  $M = (-2.92 \pm 0.020) \cdot \log P + (-1.55 \pm 0.04)$  at  $R^2 = 0.44$

For example, this is the calculation of the uncertainties for each star involved in the program. Each star has the uncertainty values for:

$\Delta \log P$  and  $\Delta M$  The uncertainty of  $\Delta \log P$  is relatively straightforward. As table 4 shows, the values of uncertainty for the pulsation period are included in the dataset. Hence, the uncertainty can be calculated:

$$\Delta \log_{10} P = \frac{1}{\ln 10} \cdot \frac{\Delta P}{P}$$

Since Gaia gives very high-quality period data, the random error in these measurements is very low, and its error bars are not visible in the graph.

However, the calculation of uncertainty becomes significantly more complicated for the absolute magnitude. The uncertainty in absolute magnitude,  $\Delta M$ , arises from the propagation of errors in the apparent magnitude  $m$ , the distance  $d$ , and the extinction correction  $A_G$ . Since the equation used is 1:

$$M = m - 5 \log_{10}(d) + 5 - A_G$$

we apply error propagation rules for functions involving logarithms. Assuming these uncertainties are uncorrelated, the total uncertainty in  $M$  is given by:

$$\Delta M = \sqrt{(\Delta m)^2 + \left( \frac{5}{\ln 10} \cdot \frac{\Delta d}{d} \right)^2 + (\Delta A_G)^2}$$

However, since distance  $d$  is computed from parallax via  $d = \frac{1}{p}$ , its fractional uncertainty becomes:

$$\frac{\Delta d}{d} = \frac{\Delta p}{p}$$

Therefore, the formula simplifies to:

$$\Delta M = \sqrt{(\Delta m)^2 + \left( \frac{5}{\ln 10} \cdot \frac{\Delta p}{p} \right)^2 + (\Delta A_G)^2}$$

Each Gaia star in the dataset includes values for  $\Delta p$ ,  $\Delta m$  (estimated using the inverse of flux-over-error), and extinction uncertainty via the provided upper and lower bounds for  $A_G$ . These values were used to compute  $\Delta M$  for every individual star.

The dominant contribution to the error typically came from the uncertainty in parallax, especially for more distant stars. In contrast, extinction and apparent magnitude errors were comparatively small. The error bars for  $M$  in Figure 7 reflect this propagated uncertainty.

## 6 Equipment

### 6.1 The OGLE catalog's telescope

The measurement equipment is a telescope dedicated to the OGLE catalog, featuring a 1.3-meter Ritchey–Chrétien. Figure 8 illustrates its inner workings.

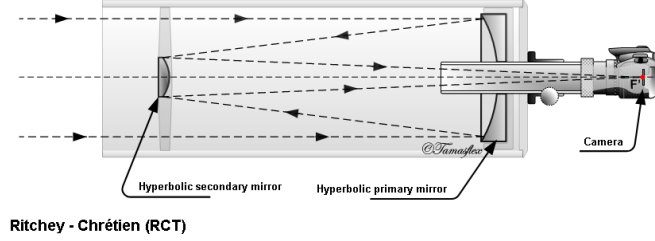


Figure 8: Sourced from Wikimedia Commons contributors, [n.d.](#) These telescopes offer sharp, wide-field imaging, making them ideal for monitoring thousands of variable stars. The telescope also features a Charge-Coupled Device (CCD) camera. These sensors convert photons into electrical charges using a silicon chip. These charges are then read out and digitalized to form an image. (MSSL, University College London, [n.d.](#))

## 6.2 Gaia

The Gaia space observatory, launched in 2013, carried the largest CCD camera sensor ever launched into space at the time. (European Space Agency, [n.d.-a](#)) By rotating and utilizing a complex mirror system, the probe scanned the sky to create a map of all stars with great precision, enabling it to measure even the differences in parallax angles.

## 7 Results

This section presents the results of the analysis conducted to determine the relationship between the pulsation period of classical Cepheid variable stars and their apparent and absolute magnitudes. The analysis was carried out in two main parts:

1. A log-linear regression using OGLE data to derive a relationship between period and apparent magnitude.
2. A calibration of this relationship into absolute magnitude using Gaia parallax data.

The regression of apparent magnitude against  $\log(P)$  using OGLE data yielded a clear fit (Figure 6):

$$m = (-2.92 \pm 0.02) \cdot \log(P) + (16.843 \pm 0.013)$$

With  $R^2 = 0.90$ , this strong correlation supports Leavitt's law (Equation 2). The dataset comprises 2,446 stars from the LMC, and although some outliers are present, most points fall within the expected bounds, indicating that the modeled impact of LMC depth is realistic.

However, because apparent magnitude  $m$  depends on both distance and extinction, a more universal relationship requires calibration to absolute magnitude  $M$ . For this, data from 344 nearby Cepheids in Gaia DR3 was used, including parallax, extinction, and apparent magnitude in the G-band.

To ensure consistency with the OGLE slope,  $a = -2.92 \pm 0.02$  was fixed. Although the G-band introduces more extinction-related uncertainty than the I-band, the intercept was determined through linear regression, yielding:

$$M = (-2.92 \pm 0.02) \cdot \log P + (-1.55 \pm 0.04)$$

The lower  $R^2 = 0.44$  reflects a more significant level of random uncertainty in the Gaia data, attributed to parallax, extinction, and photometric variability. Still, the overall trend remains consistent with Leavitt’s law, supporting the existence of a log-linear relationship.

Since Leavitt’s law is empirical rather than a theoretical formula derived from first principles, the slope and intercept depend on the sample of stars used and how extinction, metallicity, and blending are handled. Every measurement of it reports different values. According to the literature by Clementini et al., 2016, solved using data from Pietrzyński et al., 2013, suggests similar values to those measured in the essay for the **G-band**, reaching:

$$M = (-2.818 \pm 0.032) \cdot \log P - (1.129 \pm 0.020)$$

Whereas our measurements stated:

$$M = (-2.9173 \pm 0.020) \cdot \log P + (-1.5497 \pm 0.0374)$$

Although the results are not entirely consistent due to very small uncertainties, they are generally close. However, other studies suggest different results for different bands. Soszyński et al., 2015 is performed on the *I-band* and hence coincides with  $a = -2.95$ , but strongly differs in calibration, as  $b = 16.8$  in the study, since the study utilizes a calibration in the **I-band** and we did so in **G-band**. These differences highlight that while the general form of the relation is consistent, the precise coefficients depend on the photometric band, extinction correction, and stellar sample.

| Dataset                    | Equation                                            | $R^2$        |
|----------------------------|-----------------------------------------------------|--------------|
| OGLE (LMC, I-band)         | $m = (-2.92 \pm 0.02) \log(P) + (16.843 \pm 0.013)$ | 0.8969       |
| Gaia (calibrated, G-band)  | $M = (-2.92 \pm 0.02) \log(P) + (-1.550 \pm 0.037)$ | 0.4424       |
| <b>Literature (G-band)</b> | $M = (-2.82 \pm 0.03) \log(P) + (-1.13 \pm 0.02)$   | Not reported |
| <b>Literature (I-band)</b> | $m = (-2.95 \pm 0.03) \log(P) + (16.80 \pm 0.01)$   | Not reported |

Table 5: Comparison of regression results from this investigation and literature values in different photometric bands.

## 8 Conclusion

This investigation explores the research question: “To what extent does the period of classical Cepheid variable stars correlate with their absolute magnitude?” This relationship, first discovered by Henrietta Leavitt, is fundamental in astrophysics, enabling astronomers to calculate stellar distances using the distance modulus formula (Equation 1).

Using OGLE and Gaia data, a log-linear model consistent with Leavitt’s Law was established:

$$M = (-2.9173 \pm 0.020) \cdot \log P + (-1.5497 \pm 0.0374)$$

OGLE data showed a strong correlation between apparent magnitude and period ( $R^2 = 0.8969$ ), which, given the uniform distance of LMC stars, implies a strong underlying relationship with absolute magnitude. The Gaia calibration supported this but with a much weaker correlation ( $R^2 = 0.4424$ ), likely due to differences in the photometric band, increased uncertainty in parallax, and extinction.

Overall, the investigation successfully reconstructed the period–absolute magnitude relationship. While the Gaia calibration introduces random uncertainty, it still supports the general form of Leavitt’s law, indicating that this relationship remains valid even with less-than-ideal data. Despite the limitations of the Gaia dataset, the results affirm the continued use of classical Cepheids as reliable standard candles in cosmology.

## 9 Evaluation

In this section, we will conduct a final review of the reliability of the results, identify possible sources of uncertainty that may have influenced them, and discuss ways the investigation could have been enhanced.

### 9.1 Assumption of LMC uniformity

The investigation assumed all stars in the LMC were at the same distance from Earth. This assumption, as investigated in section 5.2, introduced an error. This was quantified using LMC data, adding random uncertainty to the regression. The calculations not only quantify the error as representing a change of  $\pm 0.13$  magnitudes but also justify this assumption as this error represents a small proportional error to the magnitudes ranging from 17 to 12. Moreover, this assumption was used to find the relationship, making it even more meaningless.

### 9.2 Extinction and photometric band differences

Gaia data is in the G-band, while OGLE used the I-band. The G-band, being more sensitive to dust, may introduce systematic errors due to extinction, affecting Gaia’s apparent magnitude  $m$ . While Gaia data on extinction partially

mitigated these errors, the impact this had on  $R^2$  is hard to estimate. Furthermore, the broader G-band wavelength results in greater variability, introducing random errors.

### 9.3 Gaia parallax uncertainty

Gaia’s data accounts for uncertainty in parallax measurements. Excluding those with large uncertainties, distance errors significantly affect the logarithm of the distance modulus for distant stars. These errors are predominantly random, as Gaia was a precise experiment continually checked for systematic errors by its team. For a measurement with a 10% parallax uncertainty, this adds  $\pm 0.22$  magnitudes of uncertainty to the star’s luminosity or brightness.

### 9.4 Fixing the slope

Keeping the slope aligned with accurate OGLE data may hinder the results. OGLE and Gaia data differ in bands, potentially causing discrepancies. Still, OGLE data underpins this IA’s conclusion; independent Gaia results would stray from Leavitt’s original observation, potentially introducing errors related to G-band properties discussed in Section 9.2. If the fit is done for the Gaia data only and without restrictions, the slope turns out to be  $-2.5 \pm 0.2$ , and the  $R^2 = 0.45$ , demonstrating this choice does not have a very big impact on the value of  $R^2$ .

### 9.5 Instrumental limitations

Gaia and OGLE used different instruments, photometric bands, and sensitivities. As the investigation depends on secondary data, there was no control over data collection; in some cases, we had to implement workarounds, such as calculating the error of  $M$  based on the position of the LMC as discussed in section 5.2, since OGLE did not provide the errors themselves. Out of all instrumental uncertainties, the parallax error had the biggest impact on the final result, significantly lowering  $R^2$ , since the LMC and the choice of fixing the slope did not have such a big impact as to drop  $R^2$  that much. Photometry also had an impact, but in the form of systematic errors, which become harder to quantify without doing more experiments.

### 9.6 Suggestions for improvement

Potential methods to enhance this investigation may involve employing multi-band photometry to eliminate band discrepancies and utilizing Cepheids with minimal extinction and steeper parallax angles. If this minimizes the error in the Gaia measurements, a distinct regression fit without slope restrictions could be performed.

## 9.7 Final Comments

Overall, a strong relationship exists between period  $P$  and apparent magnitude  $m$ , supported by the OGLE results. These measurements affirm Leavitt’s law. As discussed in 9.4, it’s not the calibration for  $M$  that impacts precision and reliability, causing the low  $R^2$ . Instead, it seems to be caused by other instrumental uncertainties. Even so, it does not undermine the findings on the relationship but instead affects result accuracy. On a final personal note, conducting my first astrophysics research has been a rewarding experience. While I need to explore other physics fields before specializing, this IA helped me overcome my long-standing fear of a field I admired. I look forward to expanding my knowledge in it soon.

## 10 References

### References

- Clementini, G., Ripepi, V., Leccia, S., Mowlavi, N., Lecoœur-Taïbi, I., Marconi, M., Szabados, L., Eyer, L., Guy, L. P., Rimoldini, L., De Fombelle, G. J., Holl, B., Busso, G., Charnas, J., Cuypers, J., De Angeli, F., De Ridder, J., Debosscher, J., Evans, D. W., & et al. (2016). Gaia data release 1: Testing parallaxes with local cepheids and rr lyrae stars. *Astronomy & Astrophysics*, 595, A133. <https://doi.org/10.1051/0004-6361/201629583>
- Dederick, E., & Jackiewicz, J. (2017). A possible mechanism for driving oscillations in hot giant planets. *The Astrophysical Journal*, 837(2), 148. <https://doi.org/10.3847/1538-4357/aa61a8>
- European Space Agency. (n.d.-a). *Focal plane - gaia - cosmos* [Accessed April 2025]. <https://www.cosmos.esa.int/web/gaia/focal-plane>
- European Space Agency. (n.d.-b). *The large magellanic cloud* [Accessed April 2025]. <https://esahubble.org/images/heic0411d/>
- European Space Agency. (n.d.-c). *Light echoes around rs puppis* [Accessed April 2025]. <https://esahubble.org/videos/heic1323a/>
- Gaia archive* [Accessed April 2025]. (n.d.). European Space Agency. <https://gea.esac.esa.int/archive/>
- Las Cumbres Observatory. (n.d.). *What is apparent magnitude?* [Accessed April 2025]. <https://lco.global/spacebook/distance/what-apparent-magnitude/>
- Leavitt, H. S. (1908). 1777 variables in the magellanic clouds. *Annals of the Astronomical Observatory of Harvard College*, 60, 87–108. <https://articles.adsabs.harvard.edu/pdf/1908AnHar..60..87L>
- Leavitt, H. S., & Pickering, E. C. (1912). Periods of 25 variable stars in the small magellanic cloud [Accessed April 2025]. *Harvard College Observatory Circular*, 173, 1–3. <https://ui.adsabs.harvard.edu/abs/1912HarCi.173...1L>

- MSSL, University College London. (n.d.). *Ccd operation and readout* [Accessed April 2025]. [https://www.mssl.ucl.ac.uk/www\\_detector/ccdgroup/opttheory/ccdoperation.html](https://www.mssl.ucl.ac.uk/www_detector/ccdgroup/opttheory/ccdoperation.html)
- Ogle cepheids search form [Accessed April 2025]. (n.d.). [https://ogledb.astrouw.edu.pl/~ogle/OCVS/ceph\\_query.php](https://ogledb.astrouw.edu.pl/~ogle/OCVS/ceph_query.php)
- OGLE Team. (n.d.-a). *Classical cepheids in the large magellanic cloud* [Accessed April 2025]. [https://ogle.astrouw.edu.pl/cont/4\\_main/var/ogleiii/ceplmc/](https://ogle.astrouw.edu.pl/cont/4_main/var/ogleiii/ceplmc/)
- OGLE Team. (n.d.-b). *Ogle project homepage* [Accessed April 2025]. <https://ogle.astrouw.edu.pl/>
- Parallax* [Accessed April 2025]. (n.d.). <http://hyperphysics.phy-astr.gsu.edu/hbase/Astro/para.html>
- Pietrzyński, G., Graczyk, D., Gallenne, A., Gieren, W., Thompson, I. B., Pilecki, B., Karczmarek, P., Górski, M., Suchomska, K., et al. (2019). A distance to the large magellanic cloud that is precise to one per cent. *Nature*, 567(7747), 200–203. <https://doi.org/10.1038/s41586-019-0999-4>
- Pietrzyński, G., Graczyk, D., Gieren, W., Thompson, I. B., Pilecki, B., Udalski, A., Soszyński, I., et al. (2013). An eclipsing-binary distance to the large magellanic cloud accurate to two per cent. *Nature*, 495(7439), 76–79. <https://doi.org/10.1038/nature11878>
- Soszyński, I. (n.d.). *Ogle atlas of variable star light curves - classical cepheids* [Accessed April 2025]. <https://ogle.astrouw.edu.pl/atlas/classical-Cepheids.html>
- Soszyński, I., Udalski, A., Szymański, M. K., Skowron, D., Pietrzyński, G., Poleski, R., et al. (2015). The ogle collection of variable stars. classical cepheids in the magellanic system. *Acta Astronomica*, 65, 1–1. <https://arxiv.org/pdf/1601.01318>
- STRUCTURES Blog. (2024, March). *Leavitt’s law – or: The woman who changed astronomy* [Accessed April 2025]. [https://structures.uni-heidelberg.de/blog/posts/2024\\_03\\_womens\\_day/index.php](https://structures.uni-heidelberg.de/blog/posts/2024_03_womens_day/index.php)
- Unknown, U. (n.d.). *Radial stellar pulsations* [Accessed April 2025]. <https://www.astro.princeton.edu/~gk/A403/pulse.pdf>
- Van Der Marel, R. P., & Cioni, M. L. (2001). Magellanic cloud structure from near-infrared surveys. i. the viewing angles of the large magellanic cloud. *The Astronomical Journal*, 122(4), 1807–1826. <https://doi.org/10.1086/323099>
- Wikimedia Commons contributors. (n.d.). File:ritchey-chrétien.png – wikimedia commons [Accessed April 2025]. <https://commons.wikimedia.org/wiki/File:Ritchey-Chr%C3%A9tien.png>
- Wikipedia contributors. (n.d.). File:Heic1323a-1243686232.jpg [Accessed April 2025]. <https://en.wikipedia.org/wiki/File:Heic1323a.-1243686232.jpg>