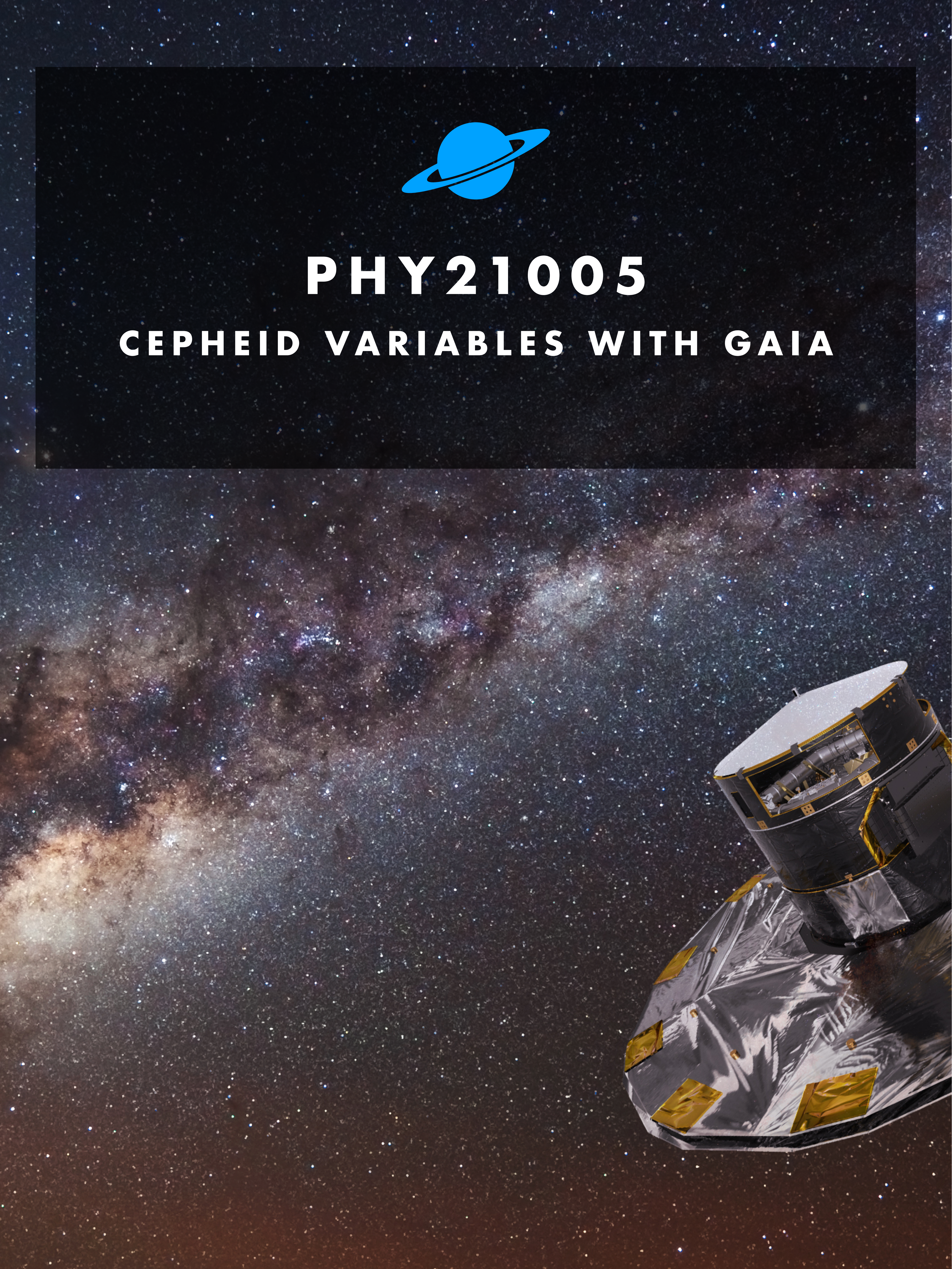
PHY21005

CEPHEID VARIABLES WITH GAIA



INTRODUCTION

GAIA

Gaia is a global space astrometry mission, with the aim of making the largest and most precise 3D map of the Galaxy - surveying more than a billion stars.

Gaia aims to monitor each of its target stars around 70 times over a five-year period. It charts the positions, distances, movement and brightness change. Distances are measured via parallax, whilst repeated visits can detect new variable stars and measure precise photometry for previously known objects.

TASK:

The Gaia mission is outlined in detail at the mission website: https://sci.esa.int/web/gaia/home. Read the mission summary and learn more about Gaia.

CEPHEID VARIABLES

Cepheid variables are pulsating stars, with pulsation periods in the range 1-100 days. There are actually more than one type of cepheid variable: **classical cepheids** are yellow giants and supergiants with spectral types of F6-K2, **type II cepheids** are older, metal poor stars around half the mass of the Sun.

Cepheids are important because their pulsations exhibit a period-luminosity relation, first discovered by Henrietta Swan Leavitt in 1908. Cepheids with longer periods are more luminous. Since Cepheids are large and bright, they are an important distance indicator. Once the period-luminosity relationship is measured, the luminosity of a Cepheid in a galaxy can be inferred from its period; the distance is then measured from the difference between apparent magnitude (flux) and absolute magnitude (luminosity).

In practice this is not as simple as it sounds:

- there are different sub-classes of Cepheids that exhibit different period-luminosity relationships;

- metallicity also affects the period-luminosity relationship
- extinction must be corrected for when finding the apparent magnitude.

TASK:

Start with the Wikipedia entry for Cepheid variables, which is quite detailed and unusually well referenced. Read this page and the references within and make notes on:

- the different types and sub-types of Cepheid variable;
- why Cepheid variables pulsate;
- whether all Cepheid types fall on the same period-luminosity relationship.

LAB OUTLINE

The use of Cepheid variables as distance indicators is one of the keys to local measurements of the Hubble constant. This is not historical astronomy, but cutting-edge research. The Hubble constant as measured from local measures, such as Cepheids and Ia supernovae is around 74 km/s/Mpc, whereas the value suggested from analysis of the Cosmic Microwave Background is nearer 67 km/s/Mpc.

We still don't know whether this reflects systematic errors, or the existence of new physics. See this blogpost and paper for a discussion.

Because Gaia makes repeated measurements of the brightness of stars, and uses parallax to measure distance, the Gaia survey is ideally placed to help resolve this tension by improving our use of Cepheids to measure distance. The Gaia mission is not complete, but there are early data releases, which we will analyse to find the period-luminosity-metallicity relationship of Cepheid variables.

In this project we will:

1. Learn how to query astronomical databases using ADQL
2. Query Gaia Data Release 3 (DR3) to find Cepheid variables in the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way.
3. Fit a period-luminosity relationship to the Cepheids in each galaxy

4. Use the measured metallicities of the three galaxies to measure a period-metallicity-luminosity relation for Cepheid variables.

ASTRONOMICAL DATABASES AND ADQL

These days, there are huge astronomical datasets from missions like Gaia or SDSS available for astronomers to analyse. These datasets reside in large online databases; knowing how to query them is a vital skill for a modern astronomer.

In order to select data from a database you use a query, written in a query language. By far the most common query language is SQL. A special dialect of SQL for astronomy is called AQDL and this is used for astronomical databases. Almost every query you can write in ADQL also works in SQL.

For this section, you will need to refer to the reference manual for AQDL. That is rather complex, so you may also prefer this ADQL cookbook which is written with Gaia data in mind.

In Gaia DR3 a large number of Cepheid variables have been identified according to their lightcurves and luminosity. We are going to learn how to query them.

You will need to refer to the following sources of information:

The documentation for Gaia Data Release 3, in particular:

- the papers explaining how Cepheid variables were identified (Rimoldini et al 2018), and verified (Clementini et al 2018) - these papers are for Gaia DR2 but the method is unchanged.
- the "data model" describing the contents of the 'gaia_source' table;
- the data model describing the contents of the 'vari_cepheid' table;
- the data model describing the contents of the `astrophysical_parameters` table.

You can run ADQL queries on the Gaia tables using the astroquery module in Python. The task below guides you through this using some online tutorials.

TASK: Your first AQDL queries.

We will use Python notebooks based in Google's Colab to carry out this lab. Go to https://colab.research.google.com/ and create a new notebook. We will need to install the astroquery module. In the first cell, type and run:

```
!pip install astroquery;
```

The astroquery module should now be installed. You can check this by typing in the next cell:

```
import astroquery
```

which should not return any errors.

Use this tutorial to see how to run ADQL queries of Gaia in astroquery, and - by reading examples in the ADQL cookbook - perform the following queries and display the results:

1. Select the top 10 rows of the `gaiadr3.vari_cepheid` table. Look at the data model for this table and select only the columns for the `source_id`, fundamental mode period (and error), the metallicity (and error), the pulsation mode and the classification of the type of Cepheid.

   Note: the types in the table mean the following. DCEP=Classical (δ) Cepheid, ACEP=Anomalous Cepheid, T2CEP=Type II Cepheid.

2. Select the top 10 rows of the `gaiadr3.gaia_source` table. Only select the columns for ra, dec, parallax (and error), G-band magnitude and Rp-band magnitude.

3. Select the top 10 rows from the `gaiadr3.gaia_source` table, only choosing objects in the Small Magellanic Cloud (SMC). For our purposes this is all the stars within a 28x4 degree rectangle centred on RA=16°, DEC=-73°.

JOINING TABLES

You'll have noticed that the Cepheid table contains information on the period, metallicity, etc, but no information on the position or parallax. This information is stored in the \`gaia_source\` table. A **relational database** is so-called because the tables are tied together by related columns between them. In the case of Gaia data, the "source_id" column contains a unique number that identifies a particular object. This column exists in both tables, allowing us join tables together.

> **TASK:**
>
> Read this tutorial on **joins** in SQL and use a join on the "source_id" column to perform the following queries:
>
> 1. Select all Cepheids in the SMC. **From the \`vari_cepheid\`** table, select the columns for source_id, fundamental period (and error), metallicity (and error), the type classification and the pulsation mode. **From the \`gaia_source\` table,** select columns for ra, dec, parallax (and error), G-band magnitude, Bp-band magnitude and Rp-band magnitude. From the \`astrophysical_parameters\` table, select the column for metallicity (mh_gspphot).
>
>   Save the table to your Google Drive in VOTABLE file format.
>
> 2. Repeat the selection for Cepheids in the Large Magellanic Cloud (LMC). Use a 30x9.5 degree rectangle centred on RA=82.5°, DEC=-68.25°.
>
>   Save the table to your Google Drive in VOTABLE file format.
>
> 3. The Cepheids in the LMC and SMC are too far away to have reliable parallaxes, but the Cepheids in the Milky Way are not. Select all Cepheids where the \`parallax_over_error\` column in the \`gaia_source\` table is >5.
>
>   These are stars with $\pi/\sigma_\pi > 5$ - i.e stars with accurate parallaxes. They will all be in the Milky Way. Save the table to your Google Drive in VOTABLE file format.

ANALYSING THE DATA

Now we have selected Cepheids from the SMC, LMC and Milky Way using the Gaia archive, it is time to try and measure the period luminosity relation. Before we start, we should inspect the data. We will start with the SMC cepheid data.

Task:

For the SMC Cepheids, plot histograms of the two metallicity estimates on the same axis. The metallicity in the `vari_cepheid` table is derived from the shape of the light curve. The estimate in the `astrophysical_parameters` table is derived from low resolution spectroscopy by Gaia. Compare the histograms to the measured metallicity of [Fe/H] = -0.98 ± 0.06 from Choudhury et al (2018, MNRAS, 475, 4279).

Do the metallicity estimates in the Gaia tables look reliable?

Task:

For the SMC Cepheids, plot the logarithm of the fundamental period on the x-axis, and the G-band magnitude on the y-axis. Plot the Classical Cepheids, Anomalous Cepheids and Type II Cepheids in different colours.

Do the data show a linear relationship?

Does it show much scatter and if so, why?

WESENHEIT MAGNITUDES

Without a knowledge of the extinction and reddening to each Cepheid, how can we correct for the extinction? The answer is that, for a given value of $R = A_V/E(B-V)$, there is a combination of magnitude and colour you can make that is *independent* of extinction; see equation 1 of Madore et al (1982). This is known as the Wesenheit magnitude and it doesn't change with extinction - provided the value of R is correct.

For the Gaia photometric system, and assuming R=3.1, the Wesenheit magnitude is given by Ripepi 2019 as:

$$W = G - 1.9(G_{BP} - G_{RP})$$

Task:

For the SMC Cepheids, plot the logarithm of the fundamental period on the x-axis, and the Wesenheit magnitude on the y-axis.

Has the quality of the linear relationship improved?

FITTING THE DATA

Task:

Fit a straight line to the data of W_{abs} vs $\log_{10} P$ for the SMC and LMC Cepheids. W_{abs} is the absolute Wesenheit magnitude. The distance moduli for the SMC and LMC are 18.98 and 18.47 respectively (Pietrzynski et al 2013).

Only fit the data for classical Cepheids pulsating in the fundamental mode.

Are the two relationships consistent with each other?

MILKY WAY CEPHEIDS

For the Milky Way, we cannot assume all the Cepheids are at the same distance, and so we will have to use the parallax to compute distances, find the distance modulus and therefore and calculate W_{abs}.

Task:

Plot W_{abs} vs $\log_{10} P$ for the Milky Way Cepheids.

Use a different colour for each class.

Is there a good relationship between period and W_{abs} for the various classes of Milky Way Cepheids?

Are there any outliers to the relationship - i.e objects that seem to lie far away from the other Cepheids? Based upon what you have read about how Cepheids are identified in Gaia, can you think of a reason why other types of star might be wrongly classified as Cepheids by Gaia?

Assuming the fainter objects are mis-classified Cepheids, fit a period-luminosity relationship using only the **reliable** classification for Classical Cepheids pulsating in the fundamental mode.

Does this relationship agree with the ones for the SMC and LMC?

PERIOD-LUMINOSITY-METALLICITY

We have seen that the Gaia estimates of metallicity are unreliable. Find good literature sources for the metallicity of the SMC and LMC, and an average metallicity for the Milky Way.

Task:

Plot graphs of the intercept and gradient of your P-L relationships as a function of metallicity. Is there a clear relation between these measured values and the metallicity?

If not, can you think of reasons why the measured slopes and intercepts of the period-luminosity relations could be innacurate?

What about your metallicities? Is it reasonably to assume that all stars in the SMC and LMC have the same metallicity? What about stars in the Milky Way?