## Data Analysis: Instructions

- **Do not touch envelopes until the start of the examination.**
- The data analysis examination lasts for 3 hours and is worth a total of 125 marks.
- There are **Answer Sheets** for carrying out detailed work and **Working Sheets** for rough work, which are already marked with your student code and question number.
- ***Use only the answer sheets for a particular question for your answer. Please write only on the printed side of the sheet. Do not use the reverse side.*** If you have written something on any sheet which you do not want to be evaluated, cross it out.
- Use as many mathematical expressions as you think may help the evaluator to better understand your solutions. The evaluator may not understand your language. If it is necessary to explain something in words, please use short phrases (if possible in English).
- You are not allowed to leave your work desk without permission. If you need any assistance (malfunctioning calculator, need to visit a restroom, etc.), please draw the attention of the supervisor.
- The beginning and end of the examination will be indicated by the supervisor. The remaining time will be displayed on a clock.
- At the end of the examination you must stop writing immediately. Put everything back in the envelope and leave it on the table.
- Once all envelopes are collected, your student guide will escort you out of the examination room.
- A list of constants and useful relations are included in the envelope.

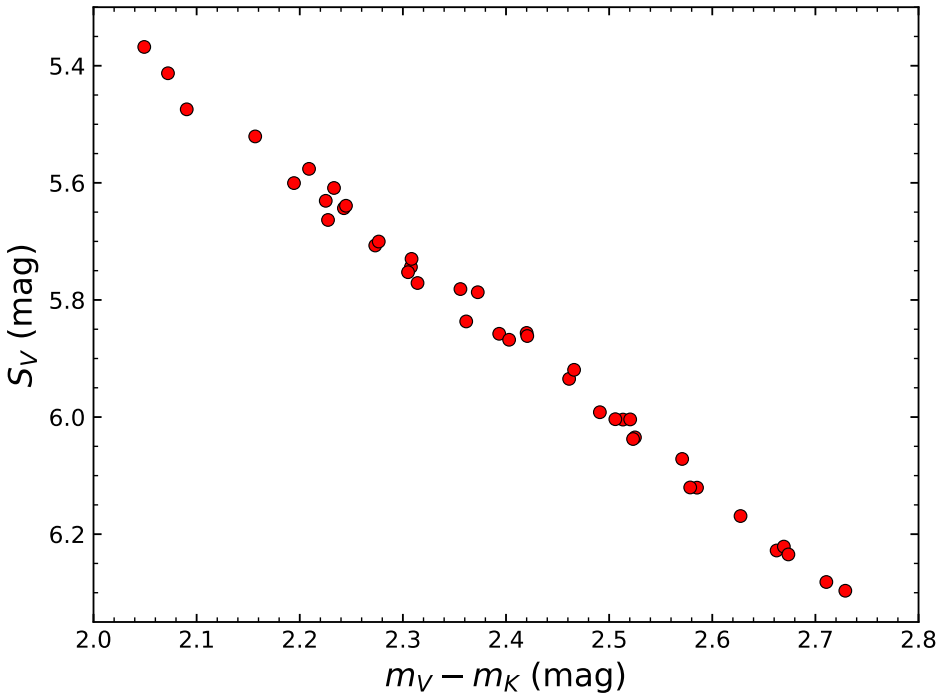
## Data Analysis 1: 'Distance to the Large Magellanic Cloud'

In 2019 an international collaboration led by Polish astronomers measured, with very high precision and accuracy, the distance to the Large Magellanic Cloud (LMC), a satellite galaxy of the Milky Way. In this way they set the zero point of the extragalactic distance scale, which allowed for a very precise measurement of the Hubble constant. Their method involved measuring the distances to 20 eclipsing binary stars in the LMC, using the concept of the surface brightness $S_V$ of a star defined as:

$$S_V = m_V + 5 \log_{10} \theta,$$

where $m_V$ is the magnitude of a star in the optical $V$ band and $\theta$ is the angular diameter of the star on the sky in milliarcseconds (mas).

The quantity $S_V$ can be understood as the magnitude of a star with an angular diameter of 1 mas. An empirical relation has been established between $S_V$ and the colour index $(m_V - m_K)$, where $m_V$ and $m_K$ are magnitudes in the $V$-band and infrared $K$-band. This is shown in the figure below for giant stars of spectral types G and K.




Using this relation, the distance to an eclipsing binary system can be determined by deriving the physical radii of the components (using photometry and spectroscopy), and comparing these with the angular diameters predicted by the $S_V$ – $(m_V - m_K)$ relation.

The table below gives the parameters of three detached eclipsing binary stars. $R_1$ and $R_2$ are the radii of each component, $V_{1+2}$ and $K_{1+2}$ are the total brightness in magnitudes of the binary in the $V$- and $K$-bands, and $L_2/L_1$ is the luminosity ratio of the components in each band.

| **source ID** | $R_1$ $[R_\odot]$ | $R_2$ $[R_\odot]$ | $V_{1+2}$ [mag] | $K_{1+2}$ [mag] | $L_2/L_1$ $(V)$ | $L_2/L_1$ $(K)$ |
|---|---|---|---|---|---|---|
| OGLE LMC-ECL-03160 | 17.03 | 37.42 | 16.73 | 14.10 | 2.80 | 4.23 |
| OGLE LMC-ECL-10567 | 24.60 | 36.64 | 16.15 | 13.83 | 1.41 | 1.99 |
| OGLE LMC-ECL-18365 | 37.30 | 15.94 | 16.27 | 14.01 | 0.206 | 0.188 |

Apply the method outlined above to the three eclipsing binary systems and calculate the distance to the LMC in kiloparsecs. Estimate the total error of the result. Assume that the fitting of the $S_V - (m_V - m_K)$ relation contributes to a bias of up to 0.8% in all measurements simultaneously.

(Total: 50 points)

Hint: in your calculations keep at least three significant figures and two decimal places. Assume that interstellar extinction is negligible and that the angular size of the LMC is small.

## Solution

*Distance calculation*

The information from the table can be used to derive individual magnitudes of both components according to equations:

$$m_1 = m_{1+2} + 2.5 \log_{10}(1 + L_2/L_1)$$

$$m_2 = m_{1+2} + 2.5 \log_{10}(1 + L_1/L_2).$$

We will use the third system (OGLE LMC-ECL-18365) as an example to demonstrate the calculations in detail. The values for this system are as follows:

*Magnitudes:* $m_{V,1} = 16.47$ mag, $m_{K,1} = 14.20$ mag, $m_{V,2} = 18.19$ mag, $m_{K,2} = 16.01$ mag.
*Colours:* $(m_V - m_K)_1 = 2.27$ mag, $(m_V - m_K)_2 = 2.18$ mag.

The second step is to determine the surface brightness $S_V$ for both components using the figure. A least square fit to the data in the figure results in a linear function:

$$S_V = 1.346\,((m_V - m_K) - 2.407) + 5.869 \text{ [mag], or}$$

$$S_V = 1.346(m_V - m_K) + 2.629 \text{ [mag].}$$

Uncertainties of the coefficients are $1.346 \pm 0.017$ mag and $2.63 \pm 0.04$ mag.
For this system, this gives $S_{V,1} = 5.69$ mag and $S_{V,2} = 5.57$ mag.

However, participants will have two other ways of determining $S_V$.
1) The first way is a graphical way by using a ruler and a pencil in order to draw the 'best-fit' line on the figure. Then $S_V$ follows from an intersection of the $x = (m_V - m_K)$ vertical line and the 'best-fit' line.
2) The second way is to determine the coefficients of the best-fit line $y = ax + b$ by using coordinates of two points on the figure. The points should be far from each other.
For example, the coordinates of the second and the penultimate points are: ($x_1$, $y_1$) = (2.07, 5.41) and ($x_2$, $y_2$) = (2.71, 6.28). This results in:

$$a = \frac{y_2 - y_1}{x_2 - x_1} = 1.36$$

$$b = \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1} = 2.60$$

Using these coefficients we have: $S_{V,1} = y = 2.27 \cdot a + b = 5.69$ mag and $S_{V,2} = y = 2.18 \cdot a + b = 5.56$ mag, so within 0.01 mag of the 'precise' results.

The third step is to calculate angular diameters of components by using the equation presented in the problem. By modifying the equation defining $S_V$ we obtain:

$$\theta = 10^{0.2(S_V - m_V)}$$

Subsequently we get: $\theta_1 = 10^{0.2(5.69-16.47)} = 0.00698$ mas and $\theta_2 = 10^{0.2(5.56-18.19)} = 0.00298$ mas.

The fourth step is to calculate the distance to each target. As components form a physical binary, their distances should be very similar. This is an independent check of the method and calculations. As the angles under which we see stellar discs are very small ($\sin\theta \approx \theta$) we can safely use a linear relation between the angular and physical diameters of an object. We therefore

calculate the distance $D$ as $D = kR/\theta$, where $\theta$ is expressed in mas, $R$ in solar radii and $k$ is a conversion factor. The conversion factor results from a fraction $(2R_\odot/1\text{ kpc})/(1\text{ mas}/1\text{ rad}) = (2R_\odot/1\text{ AU}) = 9.30 \cdot 10^{-3}$.

The distances for the third system are $D_1 = 9.30 \cdot 10^{-3} \cdot 37.30/0.00698 = 49.70$ kpc and $D_2 = 9.30 \cdot 10^{-3} \cdot 15.94/0.00298 = 49.75$ kpc.

We then repeat the calculation following the same scheme for the first and second systems, obtaining for the first system $D_1 = 49.33$ kpc and $D_2 = 49.07$ kpc, and for the second system $D_1 = 49.56$ kpc and $D_2 = 49.30$ kpc. The unweighted mean of all distances is 49.45 kpc.

*Uncertainties*

*'Statistical' part.*
The standard deviation of the sample is $s = 0.24$ kpc. The standard error of the mean is $s/\sqrt{6} = 0.09$ kpc.
OR:
The mean distances to the three eclipsing binaries are: 49.73 kpc, 49.20 kpc and 49.443 kpc. The standard deviation is $s$=0.22 kpc, and the standard error of the mean is $s/\sqrt{3}$=0.13 kpc.

*'Systematic' part.*
All distances are inversely proportional to angular diameters derived from the $S_V$ – $(m_V - m_K)$ relation. Thus their accuracy is limited by the precision of the relation. That gives the 'irreducible' part of the error: $0.008 \cdot 49.45 = 0.40$ kpc.

Finally: the distance is $49.45 \pm 0.09 \pm 0.40$ kpc; the uncertainty is completely dominated by the precision of the $S_V$ – $(m_V - m_K)$ relation.

*Author's suggestion of scoring*

A full proper solution is scored with 50 points.

*Individual scores:*
Correct formulas for individual magnitudes: 4 p
Derivation of V and K magnitudes for 6 components: 6 p
Determination of the trend line of Sv from the figure (by line fitting or by ruler): 7 p
Calculation of the SV quantity for 6 components: 6 p
Correct formula for angular diameter: 2 p
Determination of angular diameters of 6 components: 6 p
Correct formula for the distance calculation 4 p
Calculation of distances to 6 components: 6 p
Calculation of the final distance: 3 p
Errors: 'statistical': 3 p
Errors: 'systematic': 3 p