

Measuring distances to the Palomar 5 and Palomar 13 globular clusters using RR Lyrae variable stars

Luke Bischoff

Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

Abstract. RR Lyrae stars resident in two globular clusters in proximity to the Milky Way Galaxy, Palomar 5 and Palomar 13 are analysed to calculate the distances to the globular clusters. PSF photometry was performed on a series of 12 epochs using data taken from the *Spitzer* IRAC in the mid-infrared bands, 3.6 μm (channel 1) and 4.5 μm (channel 2). Apparent magnitudes of detected stars were calculated and averaged using a Gaussian local estimation algorithm (GLOESS) which were corrected for extinction effects. PL relations were fit to the data in both clusters and distance moduli were obtained from the intercepts of the PL fit and subsequently their distances. Distances of $d = 20.37 \pm 0.36$ (random) ± 1.13 (systematic) kpc for Pal 5 and $d = 24.25 \pm 0.44$ (random) ± 1.35 (systematic) kpc for Pal 13 were found to be consistent with previous work.

1. Introduction

1.1. SMHASH program and *Spitzer*

The measurement of distances to objects throughout the universe is notoriously difficult due to the lack of information available and remains an active area of research within astrophysics. Observations made using the *Spitzer* Space Telescope as part of the *Spitzer* Merger History and Shape of the Galactic Halo program (SMHASH) alongside the Carnegie RR Lyrae Problem (CRRP) can be used to help measure distances in the universe to a higher accuracy than before. The SMHASH program primarily aims to construct a three-dimensional map of the Milky Way Galaxy [1] by using mid-infrared observations of RR Lyrae (RRL) variable stars in globular clusters and dwarf spheroidal satellite galaxies (dSphs) of the Milky Way (MW) and so these programs are contingent on accurate calculation of distances to these globular clusters and dSphs.

In this paper, distances to the Palomar 5 (Pal 5) and Palomar 13 (Pal 13) globular clusters are calculated using observations from the Infrared Array Camera (IRAC) on board the *Spitzer* Space Telescope following the process outlined in sections 2 and 3. Observations were made during the Warm Mission of IRAC in 2013, with data collected using the 3.6 μm and 4.5 μm bands, after the telescope's ability to maintain cryogenic cooling to sustain its larger two waveband cameras as part of its array of four wavebands.

1.2. RR Lyrae variables

RR Lyrae stars are Population II variable stars, meaning they are relatively old and metal-poor stars that vary with regular periods of around a few hours to a day [2]. RRLs are frequently found in globular clusters in large quantities which reside in the galactic halo of the MW galaxy. Starting life on the main sequence with masses typically less than that of the Sun, RRLs have evolved past their red-giant stage to lie on the instability strip of the Hertzsprung-Russell diagram [3]. There are three main types of RRL: (i) RRab types typically pulsate with a larger period and amplitude and in the fundamental frequency, (ii) RRC types pulsate in the first overtone, typically with smaller amplitudes but in a more sinusoidal fashion and (iii) RRd types which pulsate simultaneously in the fundamental frequency and first overtone.

Due to their varying brightness with regular periods, RRL stars are used as standard candles for measuring distances within the galactic neighbourhood. Standard candles are useful as they provide markers to calibrate distances to cosmological objects, which gradually builds into the cosmic distance ladder. Previous work has typically focussed on Cepheid variable stars, discovered by Henrietta Swan Leavitt in 1908, which pulsate with a longer period than RRLs, on the order of weeks compared to hours. Cepheids are brighter than RRLs and so have typically been used as standard candles however, due to numerous advantages, such as their abundance in globular clusters in the MW galaxy and their lifespan, RRLs are particularly suited to use for accurate measurements of distance and as standard candles [4]. Period-luminosity (PL) relations show the relationship between a star's period to its magnitude and is the key to calculating distances. The cosmic distance ladder details a process of using various cosmological objects, such as Cepheids and RRLs, to calibrate different scales of distances ranging from the galactic neighbour of around 10 kpc all the way scales of 100 Mpc in an aim to yield accurate measurement of the Hubble constant, H_0 . There appears to be a fundamental disagreement on the value for H_0 between different methods of calculation, such as using the cosmic microwave background radiation (CMB), or indeed Cepheid variables [5]. It is noted that calculations involving Cepheids carry a larger uncertainty compared to CMB calculations, potentially due to the reliance of the optical wavebands for measuring distances to Cepheids. In contrast to this, distances to RRLs are calculated using data collected in the mid-infrared bands, and so PL relations for RRLs provide a number of advantages over Cepheids due to a number of reasons, namely (i) dust extinction effects are weaker, (ii) effects of metallicity are damped, (iii) more sinusoidal and symmetrical light curves and smaller amplitudes improve the precision of average apparent magnitudes, and (iv) a narrower PL dispersion [1]. Therefore, RRLs are particularly suited to high accuracy distance measurements and improve the accuracy to around 1%, the necessary first step on the cosmic distance ladder to calculating a value for H_0 to an accuracy of 2.4% [2]. The ability to solve the disagreement on the value of H_0 and obtain an accurate value is an area of active research with implications affecting the current understanding of the scale, expansion, and evolution of the universe.

2. Data selection and photometry

2.1. Data selection

The IRAC on board *Spitzer* observed a region around the centres of the Pal 5 and Pal 13 clusters, however both Pal 5 [6] and Pal 13 [7] exhibit long tidal streams extending beyond their centres due to interaction with and accretion to the MW galaxy, and so RRLs within these parts of the clusters were not observed. Each cluster was observed in each waveband of 3.6 μ m and 4.5 μ m corresponding to IRAC channels 1 and 2 respectively and over a duration of 12 epochs, with a frame in each channel per epoch, which was calibrated such that at least one full cycle for each RRL in the frame was observed. Prior to use, the raw data was pre-processed to basic calibrated data (BCD) by *Spitzer* Science Centre using an established pipeline to reduce the images into a useable state. This process includes flat fielding of the images and the removal of biases [8]. Mosaic frames of each epoch in each channel were created, which is a stitching together of the individual BCDs, a task completed prior to usage by the supervisor of this project, Dr Victoria Scowcroft resulting in Flexible Image Transport System (FITS) files used. A pixel phase effect can be exhibited on individual BCDs due to differences in flux depending on the location within the pixel a point source falls because of quantum efficient variations [9], however correction of this effect is neglected because the combination of BCDs smooths out this pixel phase effect.

The positions for right ascension (RA) and declination (dec), periods in days and RRL type of previously known RRLs was obtained from the *Catalogue of Variable Stars in Galactic Globular Clusters* (CVS) maintained by Christine Clement [10]. Further data was gathered from the *Gaia* Data Release 2 (*Gaia* DR2) [11] by considering a region of 10' radius around the centre of the cluster and searching for stars marked *variable*. This was then checked against the CVS to identify any stars missing from the CVS, as well as information of their periods obtained by scanning the *gaiadr2.vari_rrlyrae* list containing information on the type and periods of 140784 RRLs. If a match was found between the region search and DR2 RRL lists in *Gaia* by matching *source_ids*, then the periods in *Gaia* were used instead of those in the CVS as these periods were obtained in 1962 [12] due to the fact RRL periods can shift over the course of human lifespans.

Master mosaic images were also created for each IRAC channel to create FITS files of a median image using data from across all 12 epochs. These master mosaics are particularly useful due to their higher signal-to-noise ratio as a result of this combination of data. The data for the SMHASH and CRRP projects were optimised for IRAC channel 1 in the 3.6 μm band and so the channel 2 data, in the 4.5 μm has a lower signal-to-noise ratio which presented issues, particularly regarding the individual epoch data which is explained more in section 2.4 [1].

The individual FITS files contained the flux data in MJysr^{-1} , and header information regarding the exposure time, observation date and time and a flux conversion time. This information, minus the observation date and time, was also present for the master mosaic. This information was used to convert the data into data number counts, DN, using equation 1,

$$\text{counts (DN)} = \text{flux} \cdot \frac{\text{exposure time}}{\text{flux conversion factor}} \quad (1)$$

where flux is measured in MJysr^{-1} , exposure time in seconds and the provided conversion factor in $\text{MJysr}^{-1} \cdot \text{DN}^{-1} \cdot \text{s}$.

2.2. Photometry

Photometry was then performed on the data to obtain the apparent magnitudes of point sources detected in the frames. Photometry involves measuring the amount of light that falls within a desired aperture radius, in this case in data number counts. There are two main methods for performing photometry that were considered in this project, aperture photometry and point-spread function (PSF) photometry. Aperture photometry involves considering the flux of light measured through an aperture centred on the star, alongside a second aperture of larger radius drawn to create an annulus around the star. The average background flux measured in the annulus region is then subtracted from the inner aperture to obtain the flux of the star. A problem of optimisation is present here, particularly regarding the radii of the inner and outer apertures, as such a radius that is just large enough to capture light from the star but not too large such as to include nearby stars but a radius that is not too small such as to omit light from the star. The second method, PSF photometry, involves using a model to measure the flux of the stars under an iterative process whereby the measured stars are essentially removed from the image to create a residual image which may show yet more stars that were not found in the previous iteration, particularly noticeable in crowded fields. PSF photometry was chosen over aperture photometry due to the crowded nature of globular clusters and so PSF photometry was far more effective than aperture photometry.

2.2.1 PSF model. A point-spread function can be thought of as the image profile produced by a particular instrument, such as IRAC, when a point source is incident on it [13]. Anderson and King introduce the idea of an *effective* PSF (ePSF) model in their paper [13] which is a net PSF describing the amount of light from a point source falling on an individual pixel. The ePSF model was built from the master mosaic frame such that it can be used to identify point sources within the master frame, as well as the epoch frames, with profiles that are expected to be stars. A new model was constructed for each IRAC channel for both globular clusters as the model is dependent on the sample of stars present in each frame. Individual epoch frames were used to develop ePSF models to be used in the photometry of that corresponding epoch, however some epoch frames had a particularly low signal-to-noise ratio resulting in an inaccurate ePSF model which caused later defects in obtaining accurate apparent magnitude values. Therefore, the master frames only were used to develop ePSF models due to their higher signal-to-noise ratio which improved the model quality.

Stars used to construct the ePSF model were detected using a source detection algorithm, DAOSTarFinder, which implements the DAOFIND method of the greater DAOPHOT procedure [14] and is built into the photutils [15] package in python. Bright stars, above a threshold of 50σ from the background, were chosen for the ePSF model due to their higher signal-to-noise ratio over more faint sources to ensure an accurate ePSF model was derived for each IRAC channel. Stars residing near the edge of the frames were removed from the detections and each candidate star for the ePSF model were manually examined to remove stars with defects that were contaminating the ePSF model. Stars removed from the model typically exhibited crowding from nearby stars, blending effects which could be two or more unresolved stars or a galaxy and stars containing artefacts such as bad pixels or cosmic rays which were still being detected after parameters to reduce these effects were implemented as part

of the source detection algorithm (see section 2.2.2). Once removed from the list of candidate stars, the ePSF model was recalculated and saved ready for use in photometry.

2.2.2 PSF photometry. A master source list was created from the master mosaic frame for each IRAC channel in each globular cluster by using the star detection algorithm and performing a PSF routine based on the PSF model built previously, however, in this case neglecting the photometry as it was not required for this frame.

Several parameters can be used to aid with reducing the number of false positive detections made by the star finding algorithm, the first of which is the full width at half maximum (FWHM) value of stars in the frames. The radial profiles of several uncrowded and well sampled stars were measured by to gain an appreciation for the range of values for the FWHM of stars in the frames, with a value of 5.0 pixels deemed acceptable. The next parameter to consider is the threshold value, which determines the level above the background noise should a detected source be to be determined a star. As described at the start of section 2.2, there is another efficiency problem here as a compromise is required such that selecting a threshold value too high results in false negative detections but a value too low such that random noise is detected as stars and the process becoming computationally time expensive. A value of 6σ above the background noise was considered acceptable compromise for source detection in the master frame and epoch frames. Another parameter considered is the `round` parameter which considers the profiles of the sources and aims to discard sources with directional biases which would indicate the source is, for example, a galaxy. The final parameter is the `sharp` parameter which also considers the profile of the point source and aims to avoid detection of bad pixels or cosmic ray hits. The values of these parameters were determined by visual examination of the sources detected in the frames.

Photometry was then performed using an iteratively subtracted PSF routine which follows the process described in [14] whereby the star detection algorithm is run in accordance with the parameters described above, in order to ultimately measure the magnitudes of the sources in data numbers. The ePSF model constructed in each channel is then fit to the epoch data using the Levenberg-Marquardt algorithm and least squares statistic built into `astropy` [16]. A fixed centroids process to reduce the uncertainty in the magnitude measurements [1] was undertaken by using initial positions from the master frame, which has more well sampled data and a higher signal-to-noise ratio, to improve precision of the detected positions of the sources within the PSF photometry routine. A residual image was obtained after subtraction of the detected and fitting sources from the image ready for the next iteration of the process and values for the magnitudes in data numbers were obtained for stars detected in that epoch.

The process was slightly modified for IRAC channel 2 of Pal 13 to consider only regions of 50×50 pixels around known RRLs using the CVS and *Gaia* data due to an extremely bright source in the frames causing the original process to fail. The PSF photometry routine was then considered in the same fashion as described in the previous paragraph and this modification was taken forward as the primary method.

The coordinates in x and y were converted into RA and dec values using the given World Coordinate System (WCS) in the header of the FITS files and transformed using the International Celestial Reference System (ICRS) because x and y coordinates for each frame are unique to that frame only and so coordinates common to all frames and epochs were required to enable identification of the same sources in each frame. The PSF photometry process was then completed on all 12 epochs for each IRAC channel in sequence, each time matching the detected stars in the epoch to the master frame using the WCS, gradually building a large table containing the positions of each star and the magnitudes (in DN) and uncertainties for each epoch. The uncertainties in the measured magnitudes (in DN) were calculating as part of the PSF photometry routine and arise from the precision of the fitting process of the ePSF model [14].

2.3. Calculating apparent magnitudes

Apparent magnitudes m , were calculated using equation 2,

$$m = ZP - 2.5 \log_{10} \left(\text{counts} \cdot \frac{\text{aperture correction} \cdot \text{flux conversion factor}}{\text{exposure time}} \right) \quad (2)$$

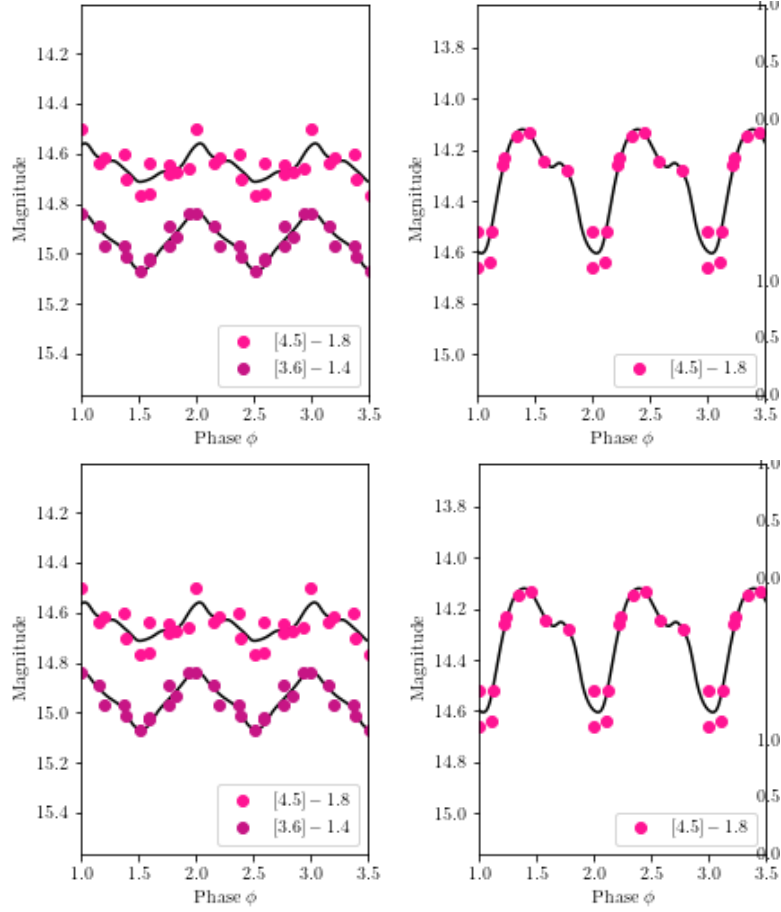


Figure 1. A selection of light curves with indicated IRAC channels from Pal 5 (top) and Pal 13 (bottom). [Note: placeholder light curves: add error bars, change colour and dimensions of light curves].

where ZP is the zero-point magnitude for each IRAC channel, which is 18.80 for channel 1 and 18.32 for channel 2 (in the IRAC Handbook [9]) and an aperture correction applied as a result of difference between the apertures used to calibrate IRAC and the aperture size of 6 used to perform photometry, also in the IRAC Handbook [9], $counts$ is the measured counts from photometry and the flux conversion factor and exposure time used to convert counts in data numbers back into flux. With the apparent magnitudes now calculated and the uncertainties in DN propagated through into magnitudes, the list of stars were matched to the CVS to identify the rows of stars and magnitude data correspond to RRLs.

2.4. Light curves

Light curves are plotted to identify how the stars vary and a selection of light curves obtained from Pal 5 and Pal 13 are shown in figure 1. A Gaussian local estimation algorithm (GLOESS) [17] has been used to fit the curves shown and to calculate an average apparent magnitude for each identified RRL for use in the PL relations. The phase of each RRL has been calculated by dividing the time of observation, in modified Julian Date (MJD) by the period and subtracting the floor of this value and the apparent magnitudes of the RRLs has been used to obtain the average apparent magnitudes [18].

The average magnitudes were corrected for extinction effects caused by dust in the MW galaxy between the observer and the globular cluster which results in a reddening of the apparent magnitudes which would influence the measured distance to the globular cluster. The corrections for dust extinction were obtained from the IRSA Dust Extinction Service Queries built into the `astroquery` module in python [19] which uses a MW dust map developed in 2011 [20]. Extinction magnitudes, A , were subtracted from each RRLs average apparent magnitude to account for this effect. It is noted that because the effects for extinction in were less than 0.1% of the general sample's average apparent magnitudes and

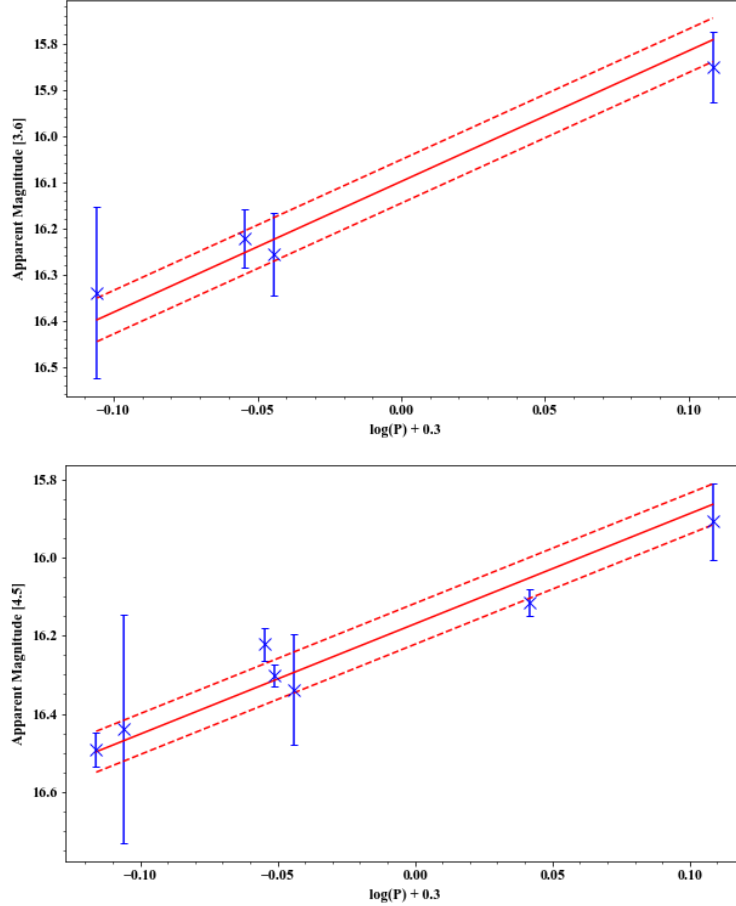


Figure 2. PL relations of RRLs in channel 1 (top) and channel 2 (bottom) for Pal 5. Dashed lines indicate 1 standard deviation of points around the PL fit as a measure of the dispersion of the cluster sample.

the lack of uncertainties provided for the extinction correction values, the value itself has been taken as the uncertainty in extinction, A .

The uncertainty in the average apparent magnitude was calculated by using equation 3,

$$\sigma_{(m)} = \sqrt{\frac{\sum \sigma_i^2}{N^2} + \sigma_{fit}^2 + A^2} \quad (3)$$

where σ_i^2 is the uncertainty in each individual epoch's apparent magnitude, N is the number of epochs present in the light curve, σ_{fit}^2 is the standard deviation of points around the fit and A the value of extinction [21]. This completed the process for obtaining average (mean) apparent magnitudes, and their uncertainties, for use in the PL relations.

3. Period-Luminosity relations and distance measurements

3.1. Building PL relations for Pal 5 and Pal 13

A PL relation shows the relationship between the magnitude of a variable star with its period and enables a comparison between apparent magnitude and absolute magnitude to be made by plotting the apparent magnitude data and fitting a PL relation to the sample data and then using a calibrated PL relation from the literature to obtain two PL relations in the same mathematical form using the following process. PL relations for mid-infrared bands for RRL populations are relatively recent [1] and have continually evolved to become more accurate and to take advantage of the benefits that the mid-infrared bands allow, such as less dependence on extinction, as discovered in the previous section, and muted metallicity effects when compared to the optical bands [3]. The PL relations shown in figures 2 and 3 for the data in Pal 5 and Pal 13 have been fit using the PL relation data provided in *Table 3* of Neeley

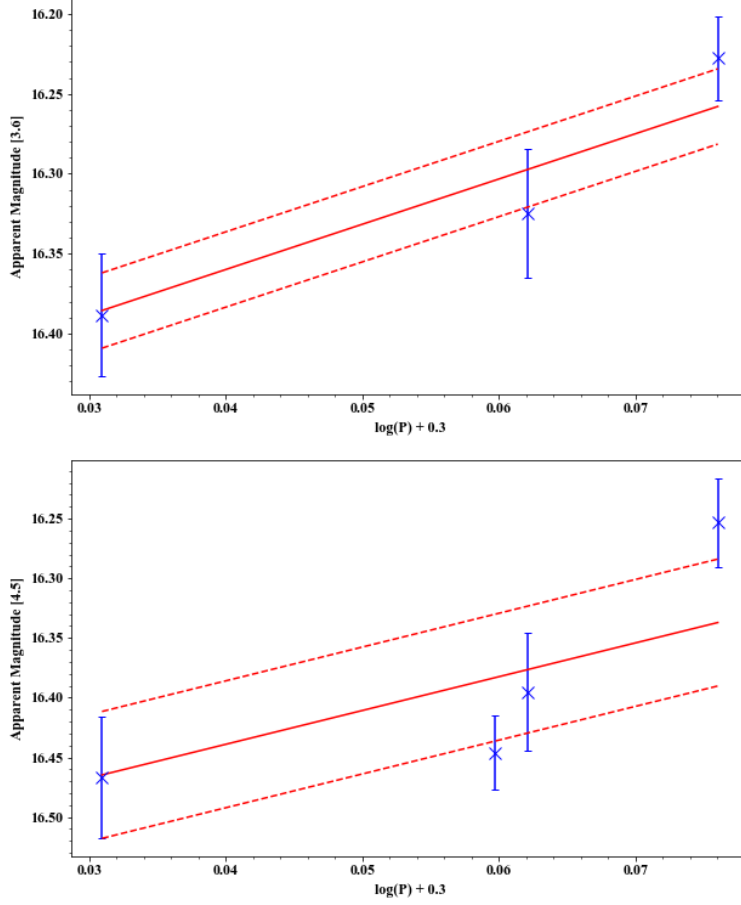


Figure 3. PL relations of RRLs in channel 1 (top) and channel 2 (bottom) for Pal 13. The dashed lines indicate 1 standard deviation of points around the PL fit as a measure of the dispersion of the cluster sample.

et al (2019) [22], the most recent and complete set of PL relations covering IRAC channels 1 and 2 as of 2020.

The calibrated PL relations in Neeley *et al* (2019), for both apparent magnitude and absolute magnitude take the form in equation 4,

$$M = a + b(\log_{10} P + 0.3) \quad (4)$$

where M is the absolute or apparent magnitude of an RRL in the sample, a is the y-intercept of the PL fit, b is the gradient of the PL fit and P is the period of the RRLs.

However, prior to plotting the data in which the PL relations will be fit to, RRLs identified as type RRC stars, which pulsate in the first overtone as opposed to the fundamental frequency, were fundamentalised to enable comparisons between the RRC and RRab stars [23], using equation 5,

$$\log_{10} P = \log_{10} P_{FO} + 0.127 \quad (5)$$

where P is the fundamentalised period and P_{FO} is the first overtone period; the logarithms have been taken in due to the form of the PL relations. The exact form of the calibrated PL relations in absolute magnitudes for both IRAC channels are shown as equations 6 and 7,

$$M_{[3.6]} = -0.40(\pm 0.03) - 2.78(\pm 0.38)(\log_{10} P + 0.3) \quad (6)$$

$$M_{[4.5]} = -0.41(\pm 0.03) - 2.83(\pm 0.39)(\log_{10} P + 0.3) \quad (7)$$

where the coefficients for a and b from equation 4 are given in Table 3 of [22]. It is now possible to fit a PL relation to the sample data, which is in apparent magnitude, by fixing the gradient b as constant, while allowing the y-intercept to be a free variable. Therefore, the PL relations shown in figure 2 and figure 3 are of the forms as equations 8 and 9 for each IRAC channel,

$$m_{[3.6]} = a_0 - 2.78(\pm 0.38)(\log_{10} P + 0.3) \quad (8)$$

Table 1. Summary of the results and uncertainties for the intercept m , distance modulus μ and distance d obtained for Pal 5 (top) and Pal 13 (bottom).

Pal 5			
IRAC Channel	m [mag]	μ [mag]	d [kpc]
[3.6]	16.10	16.51	20.04
$\pm \sigma_{\text{rand}} (\pm \sigma_{\text{sys}})$	± 0.03	$\pm 0.04 (\pm 0.12)$	$\pm 0.37 (\pm 1.11)$
[4.5]	16.16	16.57	20.69
$\pm \sigma_{\text{rand}} (\pm \sigma_{\text{sys}})$	± 0.03	$\pm 0.04 (\pm 0.12)$	$\pm 0.35 (\pm 1.15)$
Pal 13			
IRAC Channel	m [mag]	μ [mag]	d [kpc]
[3.6]	16.47	16.88	23.80
$\pm \sigma_{\text{rand}} (\pm \sigma_{\text{sys}})$	± 0.02	$\pm 0.04 (\pm 0.12)$	$\pm 0.38 (\pm 1.32)$
[4.5]	16.55	16.96	24.69
$\pm \sigma_{\text{rand}} (\pm \sigma_{\text{sys}})$	± 0.03	$\pm 0.04 (\pm 0.12)$	$\pm 0.49 (\pm 1.37)$

$$m_{[4.5]} = a_0 - 2.83(\pm 0.39)(\log_{10} P + 0.3) \quad (9)$$

where m is the average apparent magnitude and a_0 is the y-intercept yet to be found.

The intercepts were found by using a least-squares fit to obtain intercept values for Pal 5 as $a_{0[3.6]} = 16.10 \pm 0.03$ mag and $a_{0[4.5]} = 16.16 \pm 0.03$ mag in channels 1 and 2 respectively and for Pal 13 as $a_{0[3.6]} = 16.47 \pm 0.02$ mag and $a_{0[4.5]} = 16.55 \pm 0.03$ mag in channels 1 and 2 respectively. The dispersion of the PL fit (the dashed lines in figure 2 and figure 3) was taken as $\pm 1\sigma$ of the samples data in each case.

3.2. Calculating distances to Pal 5 and Pal 13

The distance to a globular cluster can be calculated using a distance modulus, which is the difference between the apparent magnitude and the absolute magnitude,

$$\mu = m - M \quad (10)$$

where μ is the distance modulus, m is apparent magnitude and M is absolute magnitude. Consider the PL relations in equations 6 to 9, for IRAC channel 1 the distance modulus can be calculated by subtracting equation 6 from equation 8, and likewise for IRAC channel 2 by subtracting equation 7 from equation 9, giving a general distance modulus, and substituting the relevant IRAC channel data into equation 11,

$$\mu = m - M = a_0 - a \quad (11)$$

where a_0 is the calculated intercept from fitting the PL relation for each IRAC channel and a is the intercept given in the Neeley *et al* (2019) [22] paper for the calibrated PL relation for each IRAC channel. Therefore, distance moduli values were calculated for Pal 5 to be $\mu = 16.51 \pm 0.04 (\pm 0.12)$ mag in channel 1 and $\mu = 16.57 \pm 0.04 (\pm 0.12)$ mag in channel 2, where the first uncertainty is the random uncertainty, defined as the uncertainties in the given calibrated intercepts a , and the calculated intercepts a_0 , and given by equation 12,

$$\sigma_{\text{rand}} = \sqrt{\sigma_a^2 + \sigma_{a_0}^2} \quad (12)$$

where σ_a is the uncertainty in the calibrated intercept from Neeley *et al* (2019) [22] and σ_{a_0} is the uncertainty in the calculated intercept from the fitting process. The uncertainty in brackets is the systematic uncertainty, defined as the uncertainties arising from the PL relation and given by equation 13,

$$\sigma_{\text{sys}} = \sqrt{\sigma_a^2 + (0.3\sigma_b)^2} \quad (13)$$

where σ_a is again the uncertainty in the calibrated intercept given in Neeley *et al* (2019) [22] and σ_b is the uncertainty in the gradient, for each respective IRAC channel. Distance moduli were also calculated to Pal 13, giving $\mu = 16.88 \pm 0.04 (\pm 0.12)$ mag in channel 1 and $\mu = 16.96 \pm 0.04 (\pm 0.12)$ mag in channel 2.

The distance d , in parsecs (pc) can be calculated from the distance modulus by,

$$\mu = 5 \log_{10} d - 5 \quad (14)$$

which can be rearranged to solve for the distance, shown in equation 15,

$$d = 10^{\left(\frac{\mu}{5} + 1\right)} \quad (15)$$

and so, distances to Pal 5 were calculated to be $d = 20.04 \pm 0.37 (\pm 1.11)$ kpc in channel 1 and $d = 20.69 \pm 0.35 (\pm 1.15)$ kpc in channel 2, as well as in Pal 13 to be $d = 23.80 \pm 0.38 (\pm 1.32)$ kpc in channel 1 and $d = 24.69 \pm 0.49 (\pm 1.37)$ kpc in channel 2. These uncertainties are the same as for the distance moduli but have been propagated through into kpc. A summary of the results from this project is provided in table 1.

4. Discussion

4.1. Comparing distance measurements to previous authors

The distances to Pal 5 for both IRAC channels in Table 1 are in agreement with each other when the random and systematic uncertainties are considered and so a mean distance to the Pal 5 globular cluster is $d_{Pal5} = 20.37 \pm 0.36 (\pm 1.13)$ kpc. This value can be compared to previous efforts to measure the distance to Pal 5, such as Price-Whelan *et al* in 2019 where an average distance to Pal 5 of 20.6 ± 0.2 kpc was achieved [24]. Therefore, the distances to Pal 5 obtained in this paper are consistent with these previous results. Another measurement to Pal 5 of $23.6^{+0.7}_{-0.6}$ kpc was obtained by Küpper *et al.* in 2015 [25]. The results presented in Table 1 and the average distance calculated in this paper is inconsistent with these results. The discrepancies between the values obtained could be due to considerations involving metallicity and use of a Period-Luminosity-Metallicity (PLZ) relation, which was not considered in the method outlined in section 2. Another discrepancy could be due to the consideration of dust extinction in both the results presented in Table 1 and [24], which as mentioned in Price-Whelan *et al.*, was not considered in Küpper *et al.* [24, 25].

The distances to Pal 13 in IRAC channel 1 and 2 in Table 1 are also consistent with each other and so an average distance to the cluster is determined to be $d_{Pal13} = 24.25 \pm 0.44 (\pm 1.35)$ kpc. A value of 24.3 ± 1.1 kpc was obtained by Cotê *et al.* [26] which is consistent with the value obtained in this paper. Furthermore, a distance of 23.6 ± 0.2 kpc was obtained by Shipp *et al.* [27], which is again consistent with the distance to Pal 13 obtained. Therefore $d_{Pal13} = 24.25 \pm 0.44 (\pm 1.35)$ kpc is consistent with recent efforts to calculate the distance to Pal 13.

4.2. Uncertainties and improvements

There are a number of sources of uncertainty associated with the measurements obtained for Pal 5 and Pal 13. Random sources of error arise from the least-squares fitting model used to determine the intercept point of the PL fit, alongside the uncertainty in the calibrated intercept given in [22]. The PL fit is ultimately based on the apparent magnitudes obtained during the photometry process and so is caused by photometric uncertainties due to background noise in the epoch data, blending or crowding of some RRLs, which was noticed for one RRL in Pal 5, or nearby bright sources causing contamination of the background, as in the case for another RRL identified in Pal 5. Systematic uncertainties are also considered and comes about from the calibration of the PL relations in [22].

The uncertainties for both globular clusters are around 7% of the value calculated which is considerably higher than the anticipated 1-2% in the SMHASH program. However, systematic uncertainty accounts for 76% and 75% of the uncertainty associated with the distance measurement in Pal 5 and Pal 13 respectively, and this is due to the uncertainty in the gradient of the PL fit obtained in [22], with uncertainties in IRAC channels 1 and 2 of $\sigma_b = 0.38$ and 0.39 respectively. If systematic uncertainty is ignored, and only random uncertainty considered then the uncertainties in the distance measurements to Pal 5 and Pal 13 reduces to 1.8% for both. As noted in [22], this large systematic uncertainty is primarily due to the large uncertainties associated with the parallax measurements in *Gaia*, which are used to calibrate the PL relations. Therefore, improvements in the calibrated PL relation are required to obtain an ideal accuracy of 2% in distance measurements and this may come with improvements in parallax measurements in future *Gaia* data releases [28].

Further improvements to the accuracy of the distance measurements could be made by considering the metallicities of the RRL populations in the globular clusters and instead using a PLZ relation which could reduce uncertainties by up to 1% [22]. Both Pal 5 and Pal 13 are in the process of tidal disruption [6, 7] and so consideration of some RRLs within these regions could improve the overall distance measurement by increasing the sample size as the small sample sizes in Pal 5 and Pal 13 could also be

a contributing factor. In Figure 1, the channel 1 PL relations for Pal 5 was fit using four RRLs, with seven in channel 2 and similarly in Figure 2, the Pal 13 PL relations were fit using three RRLs and four RRLs for channels 1 and 2 respectively.

4.3. The search for new candidate RR Lyrae variables

Identification of new possible RR Lyrae variables was attempted by analysing certain statistical methods and applying these to the measured apparent magnitude data for all detected point sources in Pal 5. These statistical methods are known as variability indices and three indices were chosen from Sokolovsky *et al* in 2017 [29] to test for signs of variability associated with any point sources detected that were not indicated in the CVS or in *Gaia* DR2.

The first test used was to calculate the standard deviation, σ , of the measured apparent magnitudes of each source detected in both IRAC channels. A variable star is expected to have a larger scatter about the mean magnitude of the light curve when compared to stars which do not vary.

The second test used was the inter-quartile range (IQR) which considers the difference between the upper and lower quartiles of the apparent magnitudes measured for a particular star, splitting the central 50% of the data in half to form two sets of data and then subtracting the median of each half from each other.

The final test considered was the Welch-Stetson index, I , which was calculated using equation 16 [29],

$$I = \sqrt{\frac{1}{n(n-1)}} \sum_{i=1}^n \left(\frac{m_{i,[3.6]} - \bar{m}_{[3.6]}}{\sigma_{i,[3.6]}} \right) \left(\frac{m_{i,[4.5]} - \bar{m}_{[4.5]}}{\sigma_{i,[4.5]}} \right) \quad (16)$$

where n is the number of pairs of measurements, $m_{i,[3.6]}$ is the measured apparent magnitudes in each IRAC channel, $\bar{m}_{[3.6]}$ is the mean apparent magnitude for that channel and $\sigma_{i,[3.6]}$ is the uncertainty. Equation 16 requires the star to be present in both channel 1 and 2 and also must be detected in all 12 epochs in both channels.

Analysis of the variability indices did not return the anticipated results and so no new possible variable stars were found in Pal 5. RRLs previously known in the CVS and *Gaia* did not necessarily return high values and thus there was no confidence that stars which did return large values for each of these tests were variable, which was indeed the case when analysing the light curves of these stars. While unfortunate, it is likely this is due to small sample of epoch data (12 data points per star) and the particular abundance of RRc stars in Pal 5, perhaps due to the smaller amplitudes associated with RRc stars.

5. Future work

The globular clusters analysed in this project contain a relatively small sample of RR Lyrae stars and attempts to avoid crowded fields has been made. The next step would be to apply this process to a more crowded globular cluster, such as IC4499, work has already started on this but is not yet complete. The IC4499 cluster contains at least 97 known RRLs from initial investigation of the catalogue of variable stars [10] and so this may require adjusting the PSF photometry code to further optimise it for crowded fields.

A further development on this work would include considering the metallicities of the RRLs in the cluster which could improve the precision of the distance measurements, as was touched upon in the discussion and so this would be considered going forwards.

In terms of the wider field, looking to the future, improvements in uncertainties would be a main consideration, particularly concerning systematic uncertainties brought about from the calibrated PL relations. Improvements in parallax measurements could significantly reduce the uncertainties in distance measurements [22, 28]. This would be particularly exciting as this could help work towards refining the value and uncertainty in Hubble's constant, as described in the introductory sections of this report, which would help improve the field's understanding of the scale and evolution of the universe.

6. Conclusion

Distances to the Pal 5 and Pal 13 globular clusters have been calculated using data collected from the *Spitzer* IRAC in the mid-infrared spectrum. An average distance of $d_{Pal5} = 20.37 \pm 0.36 (\pm 1.13)$ kpc to the Pal 5 globular cluster was obtained and found to be in agreement with the most recent measurement

using a similar method when the random and systematic uncertainties are considered. An average distance of $d_{Pal13} = 24.25 \pm 0.44 (\pm 1.35)$ kpc to the Pal 13 globular cluster was calculated and in agreement with previous measurements undertaken.

These distance measurements were obtained by performing PSF photometry on mosaic images of individual epochs across a duration of at least one period for the largest RRL period in the frame using software coded in python. Apparent magnitudes of the detected sources were calculated and the RRLs within the frames identified and their data extracted and mean apparent magnitudes calculated by way of a gaussian local estimation algorithm. Extinction effects caused by dust in the MWG was corrected for and the apparent magnitudes adjusted as a result. PL relations were then fit to the data for each IRAC channel for both clusters using a calibrated PL relation from Neeley *et al.* [22] to yield the intercept point of the fit using a least-squares model. All data points were within one standard deviation of the fit for both channels in Pal 5 and Pal 13 when considering uncertainty in the average apparent magnitude. The intercept was then used to calculate the distance moduli for each channel for both Pal 5 and Pal 13 and converted into distances in kpc.

Acknowledgements

I would like to thank my supervisor for this project, Dr Victoria Scowcroft, for providing the data used throughout the project and invaluable guidance, advice, and motivation for the duration of the project. I would also like to thank Miss Abigail Chown also for her excellent help and advice throughout. I would also like to thank my project partner, Mr Jake Bird, with whom I have worked closely with for the project. Furthermore, I would like to give credit to the Jet Propulsion Laboratory, California Institute of Technology and NASA as this work used data collected from the *Spitzer Space Telescope*; as well as credit to the European Space Agency (ESA) for their data from *Gaia*.

References

- [1] Garofalo A, *et al* 2018 *Mon. Not. R. Astron. Soc.* **481** 578-595
- [2] Beaton R L, *et al* 2016 *Astrophys. J.* **832** 210
- [3] Muraveva T, *et al* 2018 *Mon. Not. R. Astron. Soc.* **480** 4138-4153
- [4] Bono G, *et al* 2016 *Mem. S.A.It.* **87** 358
- [5] Scowcroft V, *et al* 2011 *Astrophys. J.* **743** 76
- [6] Erkal D, *et al* 2017 *Mon. Not. R. Astron. Soc.* **470** 60-84
- [7] Bradford J D, *et al* 2011 *Astrophys. J.* **743** 167
- [8] Reach W T, *et al* 2005 *Publ. Astron. Soc. Pac.* **117** 978-990
- [9] Reach W T, *et al* 2006 *Infrared Array Camera Data Handbook Version 3.0*, *Spitzer Science Centre*, Pasadena.
- [10] Clement, *et al* 2001 *Astron. J.* **122** 2587-2599
- [11] Brown A G A, *et al* 2018 *Astron. Astrophys.* **616** A1
- [12] Kinman T and Rosino L 1962 *Astron. J. Publ. Astron. Soc. Pac.* **74** 499-506
- [13] Anderson J and King I R 2000 *Publ. Astron. Soc. Pac.* **112** 1360-1382
- [14] Stetson P B 1987 *Publ. Astron. Soc. Pac.* **99** 191-222
- [15] Bradley L, *et al* 2020 *astropy/photutils: 1.0.1* URL <https://doi.org/10.5281/zenodo.4049061>
- [16] Price-Whelan A M, *et al* 2018 *Astron. J.* **156** 123
- [17] Persson A, *et al* 2004 *Astron. J.* **128** 2239
- [18] Monson A J, *et al* 2017 *Astron. J.* **156** 96
- [19] Ginsburg A, *et al* 2019 *Astron. J.* **157** 98
- [20] Schlafly E F and Finkbeiner D P 2011 *Astrophys. J.* **737** 103
- [21] Hendel D, *et al* 2018 *Mon. Not. R. Astron. Soc.* **479** 570-587
- [22] Neeley J R, *et al* 2019 *Mon. Not. R. Astron. Soc.* **490** 4254-4270
- [23] Iben I 1974 *Annu. Rev. Astron. Astrophys.* **12** 215-256
- [24] Price-Whelan A M, *et al* 2019 *Astron. J.* **158** 223
- [25] Küpper A H W, *et al* 2015 *Astrophys. J.* **803** 80
- [26] Coté P, *et al* 2002 *Astrophys. J.* **574** 783-804
- [27] Shipp N, *et al* 2020 *Astron. J.* **160** 224
- [28] Neeley J R, *et al* 2015 *Astrophys. J.* **808** 11

- [29] Sokolovsky K V, *et al* 2017 *Mon. Not. R. Astron. Soc.* **464** 274-292