A Mid-infrared Study of RR Lyrae Stars with the WISE All-Sky Data Release

Tatyana Gavrilchenko¹, Christopher R. Klein¹, Joshua S. Bloom¹, Joseph W. Richards²

ABSTRACT

We present a group of 3740 previously identified RR Lyrae variables well-observed with the Wide-field Infrared Survey Explorer (WISE). We explore how the shape of the generic RR Lyrae mid-infrared light curve evolves in period-space, comparing light curves in mid-infrared and optical bands. We find that optical light curves exhibit high amplitudes and a large spectrum of light curve shapes, while mid-infrared light curves have low amplitudes and uniform light curve shapes. From the period-space analysis, we hope to improve the classification methods of RR Lyrae variables and enable reliable discovery of these pulsators in the WISE catalog and future mid-infrared surveys such as the James Webb Space Telescope (JWST). We provide mid-infrared templates for typical RR Lyrae stars and demonstrate how these templates can be applied to improve estimates of mid-infrared RR Lyrae mean magnitude, which is used for distance measurement. This method of template fitting is particularly beneficial for improving observational efficiency. For example, using light curves with observational noise of 0.05 mag, we obtain the same level of accuracy in mean magnitude estimates for light curves randomly sampled at 12 data points with template fitting as with light curves randomly sampled at 20 data points with harmonic modelling.

Subject headings: infrared: stars; methods: statistical; RR Lyrae: distance scale

Submitted for publication in MNRAS, December 2013.

1. Introduction

RR Lyrae stars are pulsating variables that lie near the lower luminosity edge of the instability strip on the Hertzsprung-Russell diagram. At optical wavelengths their mean flux magnitude is nearly constant with a slight metallicity dependence (Smith 2004; Chaboyer 1999). At near- and mid-infrared wavelengths, however, RR Lyrae stars exhibit a period-luminosity relation (Longmore et al. 1986). Historically, RR Lyrae stars have been used as important distance indicators for Galactic globular clusters (Oosterhoff 1939; Suntzeff et al. 1991; Cacciari & Renzini 1976). More
recently, and with the calibration of their longer wavelength period-luminosity relations, RR Lyrae are being used to trace galactic structure near the Galactic Center (Soszyński et al. 2011) and in the halo (Sesar et al. 2010), as well as map the depth structure of the Magellanic Clouds (Soszyński et al. 2009, 2010). An overarching goal of this field is the improvement in RR Lyrae distance measurements through reduced uncertainty in period-luminosity relations. There have been recent advances in the calibration of period-luminosity of Cepheid variable stars (Riess et al. 2011; Freedman et al. 2012).

RR Lyrae stars are evolved low-mass (≈ 0.7 M⊙) Population II stars that inhabit the Galactic halo and bulge, and permeate the disk (Smith 2004). Their abundance and age make them, in principle, good tracers of Galactic structure. RR Lyrae stars are divided into subclasses based on the shapes of their optical light curves. RRab stars have periods mostly ranging from 0.3 to 0.9 day, $V$-band amplitudes of 0.8 to 1.5 magnitudes, and asymmetrical light curves, whereas RRc stars have periods mostly ranging from 0.18 to 0.5 day, $V$-band amplitudes of 0.3 to 0.8, and more sinusoidal light curves (Bailey 1902). The main difference between the two subclasses is that RRab stars pulsate in the fundamental mode and RRc stars pulsate in the first radial overtone (Percy 2007). Two rarer subclasses are RRd stars, which are double-mode pulsators, and RRe stars, which pulsate in the second radial overtone. The distribution of RR Lyrae subclasses in the LMC as measured by the OGLE-III survey is 71% RRab, 20% RRc, 4% RRd, and 5% RRe (Soszyński et al. 2009).

Probing the sky in infrared wavebands minimises the effects of interstellar dust and allows detection of faint but warm $^3$ sources such as asteroids. The Wide-field Infrared Survey Explorer (WISE) satellite (Wright et al. 2010) provides full-sky coverage in four mid-infrared wavebands. The survey mission was to cover 99% of the sky with at least 8 frames in a 6 month interval then to continue to survey until the cryogenic coolant was exhausted. The survey is up to 60 times more sensitive than its predecessor, IRAS (InfraRed Astronomical Satellite), launched in 1983. The Spitzer Space Telescope has a flux limit 20 times deeper than typical WISE sensitivity (Meixner et al. 2006; Wright et al. 2010). However, this sensitivity requires long exposure times and dedicated pointings, making Spitzer unsuitable for a complete survey of the sky. The W1 and W2 WISE passbands are nearly the same as the first two bands on Spitzer’s IRAC instrument, which allows WISE to supplement catalogs and follow up on observations made by Spitzer (Meixner et al. 2006).

The upcoming successor to some WISE work will be the James Webb Space Telescope (JWST), since all of the WISE bands fall into the 0.6 to 28 μm range of JWST. JWST will detect faint, distant RR Lyrae, which will be used as distance markers with near- and mid-IR period-luminosity relations, the latter of which have been derived using WISE data (Klein et al. 2011; Madore et al. 2013).

---

$^3$ A 725 K blackbody peaks at 4 μm.
While variable star science was not among the main objectives of the WISE mission, the survey’s scanning pattern enables it to detect variable stars with periods less than 2 days. Due to relatively large amplitudes and strong mid-IR emission, RR Lyrae stars are the most common periodic variables detected by WISE, along with eclipsing Algol variables (Hoffman et al. 2012). The negligible effects of dust in mid-IR bands is another source of motivation for this project (Rieke & Lebofsky 1985).

With multi-band studies of light curve shape and period-luminosity relations imminent, it will be important to obtain the light curve of a source in several bands. However, well-represented light curves of variable stars can be costly to obtain with respect to telescope time. A series of observations must be made, preferably spanning several period cycles. It is also desirable to sample at random positions in the phase of the object to avoid beating the cycle of pulsation with the cycle of observation. Template fitting is a potential method to recover the same amount of light curve information with fewer observations. Templates have been shown to make accurate mean magnitude and metallicity estimates for $V$-band light curves of RRab stars (Kovács & Kupi 2007). Accurate mean magnitude and period measurements are crucial for precision distance studies, especially since period-luminosity relations have been derived for RR Lyrae variables in WISE bands (Klein et al. 2011).

The primary goals of the present research are to acquire a representative sample of RR Lyrae light curves in the mid-IR, to contrast the light curve morphology with light curves available in optical bands, and to formulate a set of mid-IR light curve templates that will aid future studies. To begin, a database of RR Lyrae stars previously identified in visible bands was collated and cross-referenced with the WISE database to form a primary sample set. The final WISE RR Lyrae catalog was created from sources for which periods were derived. These periods were used to create mid-IR light curve models, which appear homogenous in shape, losing distinctive features seen in optical light curves. The models were grouped into typical light curve templates which are an intended tool for the recovery of mean magnitude values for poorly-sampled sources.

In Section 2, we describe in detail the construction of our catalog and our WISE data retrieval methods. We discuss the methodology for determining the period of each source in Section 3.1, the process by which we derive fitted model light curves in Section 3.2, and present statics on the sample in Section 4. Finally, we derive a set of typical light curve templates in Section 5. Our main results are discussed and summarised in Section 6.

2. Data

2.1. Building a Database of RR Lyrae

The first goal was to find a large sample of WISE-observed RR Lyrae stars. A database of previously-identified RR Lyrae stars was created, combining information from GCVS, All Sky
Automated Survey (ASAS), Simbad, VizieR, and individual papers. For many of the sources in this database the only available data were the coordinates and RR Lyrae classification. When provided, information about the period, distance, subclass, and magnitude for several different wavebands was also stored. If a single source appeared in multiple surveys or papers, information from all relevant surveys was included, with markers indicating contradicting measurements between surveys. The resulting database contains about 17,000 sources, of which about 5,000 sources have documented V-band periods. A random sample from the database is shown in Table 1 with references to the original sources given in Table 2. Columns not shown in the table are magnitude values for the following bands: U, B, R, I, J, H, K, and SDSS ugriz bands.

The database inherits the biases of the surveys used to create it, so it is neither complete nor unbiased. In addition, there may be contamination by sources incorrectly classified as RR Lyrae. The most likely contaminants are eclipsing binary systems in $\sim$0.2 day period orbits where the two bodies have similar masses, since these result in sinusoidal light curves similar to those of the RRc subclass. An alternative to this procedure is to select all stars with a variability flag from the WISE database, $\sim$460,000 objects, and to make further colour, period, and amplitude restrictions to form a sample of likely RR Lyrae candidates. However, since the bounds of such cuts are not well-studied in W1 and W2 bands, this procedure may result in a sample highly contaminated by Algol-type eclipsing binaries and W UMa binaries.

### 2.2. Retrieving Data from WISE

WISE performed an all-sky survey in four mid-infrared bands centred at 3.4, 4.6, 12, and 22 $\mu$m. The strategy of WISE was to scan in great circles with a centre located at the Sun. The result of this scanning pattern was a high number of exposures taken around the ecliptic poles and a low number of exposures taken near the ecliptic plane. Typically, points around the ecliptic poles were observed over 100 times and points around the ecliptic plane were observed 12 to 13 times. The full cryogenic mission phase began on 7 January 2010, completed full coverage of the sky on 17 July 2010, and surveyed approximately 20% of the sky a second time before the mission phase ended on 6 August 2010. A preliminary data release with coverage of 57% percent of the sky was released in April 2011, superseded by the all-sky data release in March 2012.

Coordinates from the compiled database of known RR Lyrae stars were cross-referenced with the WISE all-sky release catalog. A search radius of 10 arcseconds was used and data points with artefact contamination flags or photometric quality flags were rejected. The coordinate-linked cross-reference search returned $\sim$8,500 sources with ten or more measurements in either the W1 or W2 bands. Analysis of the W3 and W4 light curves is omitted, since both the photometric quality and the number of epochs in those longer-wavelength bands are greatly diminished compared to

4http://wise2.ipac.caltech.edu/docs/release/allsky/
Table 1. Sample of the RR Lyrae Catalog

| ID          | RA     | DEC     | Subclass | Period | V Magnitude | Reference |
|-------------|--------|---------|----------|--------|-------------|-----------|
| V* V532 Her | 272.98 | 28.07   | RRab     | 0.616  | 11,17       |           |
| 154584      | 110.16 | 6.68    | RRab     | 0.616  | 20,22       |           |
| NSV 06930   | 227.05 | -69.84  | RRab     | 0.616  | 1           |           |
| V2637 Sgr   | 273.17 | -36.72  | RRab     | 0.616  | 1           |           |
| OGLE BUL-SC45 245339 | 18.05 | -30.02 | RRab     | 0.616  | 1           |           |
| MACHO 176.18702.1147 | 269.95 | -27.2869 | RRab    | 0.616  | 1           |           |
| V* V4267 Sgr | 272.17 | -31.5431 | RRab    | 0.616  | 3,17,19     |           |
| V0344 Cam   | 66.063 | 71.6589 | RRb      | 0.616  | 11,17       |           |
| ZZ And      | 12.395 | 27.0222 | RRab     | 0.616  | 11,17       |           |
| V0344 Cam   | 66.063 | 71.6589 | RRb      | 0.616  | 11,17       |           |
| NSV 06930   | 227.05 | -69.84  | RRab     | 0.616  | 11,17       |           |
| MACHO 176.18702.1147 | 269.95 | -27.2869 | RRab    | 0.616  | 11,17       |           |
| V* V505 Lyr | 279.19 | 28.0694 | RRab     | 0.616  | 11,17       |           |
| V* V505 Lyr | 279.19 | 28.0694 | RRab     | 0.616  | 11,17       |           |
| 275.376     | 268.0971 | -29.7703 | RRab    | 0.616  | 3,17,19     |           |
| BLG 40925   | 80.07317 | -69.67780 | RRab   | 0.616  | 3,17,19     |           |
| MACHO 104.20383.612 | 270.9915 | -27.918 | RRab     | 0.616  | 11,17       |           |
| MACHO 134.33910.838 | 278.8755 | -27.43281 | RRab    | 0.616  | 11,17       |           |
| V* V505 Lyr | 279.19 | 28.0694 | RRab     | 0.616  | 11,17       |           |
| V0344 Cam   | 66.063 | 71.6589 | RRb      | 0.616  | 11,17       |           |
| ZZ And      | 12.395 | 27.0222 | RRab     | 0.616  | 11,17       |           |
| V0344 Cam   | 66.063 | 71.6589 | RRb      | 0.616  | 11,17       |           |
| NSV 06930   | 227.05 | -69.84  | RRab     | 0.616  | 11,17       |           |
| MACHO 176.18702.1147 | 269.95 | -27.2869 | RRab    | 0.616  | 11,17       |           |
| V* V505 Lyr | 279.19 | 28.0694 | RRab     | 0.616  | 11,17       |           |
| 275.376     | 268.0971 | -29.7703 | RRab    | 0.616  | 11,17       |           |
| V303 Sgr    | 274.7667 | -31.4717 | RRab    | 0.616  | 11,17       |           |
| V0759 Cep   | 7.3362  | 84.7625 | RRb      | 0.616  | 11,17       |           |
| BB97c 37    | 272.68462 | -31.97056 | RRab   | 0.616  | 11,17       |           |
| V* V1357 Sgr | 270.6475 | -29.9583 | RRab     | 0.616  | 11,17       |           |
| V2466 Sgr   | 269.9404 | -28.8722 | RRab     | 0.616  | 11,17       |           |
| MACHO 13932982.4416 | 278.3235 | -28.59539 | RRab   | 0.616  | 11,17       |           |
| 2MAS J00112104+0010138 | 2.83767 | 0.1705 | RRab     | 0.616  | 11,17       |           |
| HL Com      | 184.5054 | 30.1883 | RRab     | 0.616  | 11,17       |           |
| MACHO 101.20911.863 | 271.254 | -27.34639 | RRab   | 0.616  | 11,17       |           |
| T Hya       | 6.0663  | -76.2819 | RRab     | 0.616  | 11,17       |           |

Note. — Columns showing magnitude for the $U$, $B$, $R$, $I$, $J$, $H$, $K$, and SDSS $ugriz$ bands are omitted.

Table 2. References for the RR Lyrae Catalog

| Reference Number | Reference Number | Reference |
|------------------|------------------|-----------|
| 1                | Groenewegen et al. (2008) | 14        | Maintz (2005) |
| 2                | Blanco (1984)     | 15        | Collinge et al. (2006) |
| 3                | Blanco (1992)     | 16        | Zorotovic et al. (2010) |
| 4                | Blanco & Blanco (1997) | 17        | http://www.sai.msu.su/gcvs/gcvs/index.htm |
| 5                | Clementini et al. (1995) | 18        | Pigulski et al. (2009) |
| 6                | Kinman et al. (1982) | 19        | http://simbad.u-strasbg.fr/simbad/ |
| 7                | Kinman et al. (1995) | 20        | Richards et al. (2012) |
| 8                | Miceli et al. (2008) | 21        | Pojmanski (1997) |
| 9                | Saha (1983)       | 23        | Debosscher et al. (2007) |
| 10               | Suntzeff et al. (1991) | 24        | Pigulski et al. (2003) |
| 11               | Wetterer et al. (1999) | 25        | Debosscher et al. (2003) |
| 12               | Storm et al. (1994) | 26        | Debosscher et al. (2007) |
| 13               | Szekely et al. (2007) | 27        | Debosscher et al. (2007) |
W1 and W2. Figure[1] is a spatial plot of the RR Lyrae database, showing the sources with matched WISE light curves in red.

![Spatial plot of RR Lyrae database and WISE light curves](image)

Fig. 1.— The distribution of the $\sim$8,500 RR Lyrae sources found in the WISE catalog (red circles) superimposed on the $\sim$17,000 sources of the RR Lyrae database (blue crosses). The Aitoff plot coordinate system is in Galactic coordinates.

Despite choosing the highest quality photometric flags, some of the light curves contained outliers. The variable nature of these stars makes outlier rejection tricky. To exclude outlying points, any observations located 0.5 or more magnitudes away from the mean light curve magnitude were eliminated. This threshold was chosen because it is well above the mid-infrared peak-to-peak amplitude for RRab stars, which was computed to be 0.225 magnitudes in Section 4. Outlier rejection was done recursively by removing the point farthest from the mean magnitude if it was over 0.5 magnitudes away, then recalculating the mean magnitude and repeating the procedure.

3. Mid-IR Light Curves

3.1. Determining Periods

WISE light curves in the W1 and W2 bands were used to attempt to determine a period for each source. The two period-finding algorithms used were Lomb-Scargle (Lomb 1976, Barning 1963, Scargle 1982) and Super-Smoother (Reimann 1994). The Lomb-Scargle algorithm uses a
least-squares frequency analysis of a range of test frequencies to create a periodogram, which is an estimate of the power spectrum. A modified Lomb-Scargle algorithm developed by [Richards et al. (2011)] was used in the present work. While the original algorithm fixes the mean magnitude of the model light curve at zero, the modified algorithm allows a floating mean value. The reason for this alteration is that observations of a source which are not uniformly sampled in period space will result in an inaccurate mean magnitude measurement. Since equally-spaced measurements in phase space are not expected with WISE data, this modification results in more robust period estimates. The Lomb-Scargle algorithm outputs the period, the periodogram, and the harmonic model, using up to 8 harmonic components. The *Supersmoother* algorithm folds the light curve on a range of test periods and determines which test periods result in the smoothest folded light curve. The outputs are the period for the source, a goodness of fit of the model to the data, and the epoch of maximum brightness. Both algorithms require a light curve as the input, which is a list of the timestamp, magnitude, and the magnitude error for each observation of the source.

The preliminary list of candidate periods was comprised of the top three periods returned from each of the two period-finding algorithms, so six total candidate periods for each of the two bands. These lists of candidate periods were trimmed by removing periods outside the 0.2 to 1 day range of RR Lyrae variables and by averaging any two candidate periods for a given source which were less than 0.05 day apart. This threshold for unique periods was used because the scanning pattern WISE completes one orbit in 90 minutes, which corresponds to 0.0625 day. The final lists typically had between one and three candidate periods for each source per waveband.

If possible, a single period from the candidate list was chosen as the most likely true period, which means that the periodogram attains a maximum at that period. For a light curve with few magnitude measurements and high photometric errors, the periodogram may have several peaks, making it difficult to identify the true period. For these ambiguous cases, no true period can be determined and these sources are excluded from further analysis. A period-determining process that minimises such occurrences is desired. The schematic in Figure 2 is a representation of the process used, and the text in this section describes each step in detail.

A harmonic model for each of the candidate periods was created, once again using the Lomb-Scargle algorithm, but now restricting the interval of scanned periods to be centered at the candidate period and to have a width of 0.1 day (effectively boosting the periodogram resolution). In addition, points in the original light curve which were over four standard deviations away from the model light curve were considered outliers and eliminated from further investigation. Upon running the Lomb-Scargle process, several features of the harmonic model were measured to serve as quality markers of the model. These features are the period, the periodogram peak height, the residual determined by the mean of the squares of the errors between the data and the model, the number of observations in a light curve, the largest gap in phase space of the folded light curve, the ratio of the data amplitude and the model amplitude, the average photometric error divided by the amplitude of the model, and the number of critical points in the model. The initial quality cuts of the period candidate selection are:
Fig. 2.— A schematic of the process used for period determination from a list of candidate periods. See section 3.1 for more detailed explanation of the process.

- candidate period is between 0.2 and 1 day
- periodogram peak power spectral density is greater than 3 dB Hz$^{-1}$
- harmonic residual is less than 0.006 magnitudes
- maximum gap in phase space is less than 30% of a period cycle
- the ratio of data amplitude to model amplitude is greater than 0.8 and less than 2.5
- the ratio of average photometric error to model amplitude is less than 0.3
- the model has only two critical points (where the derivative changes sign)
The final condition requiring exactly two critical points is reasonable because an RR Lyrae light curve is expected to have only one local maximum and one local minimum in a single pulsation cycle. If more than two critical points are detected, it is an indication of a poor model fit or an incorrect period estimate. For example, if a light curve is folded on a period which is twice the value of the true period, then two pulsation cycles will be shown. To eliminate multiples of the true period from consideration, a candidate light curve does not pass through the quality cuts if it has more than two critical points. To allow room for regions of relative flatness in the light curve model, any neighbouring critical points whose difference is less than 20% of the total model amplitude are considered to be a single critical point. This condition is strong enough to reject models that are clearly not RR Lyrae stars, yet allows leeway for borderline cases.

Next, a second duplicate removal was performed on each candidate list: if two periods less than 0.05 days apart were found in a list, the period with the higher PSD peak was kept and the other period was removed. The next step was to check if a single period is present in both the W1 and W2 shortened candidate lists. This is motivated by the expectation of an RR Lyrae star to have a single period regardless of the bandpass used to observe it. Periods appearing in both candidate lists up to an uncertainty of 0.05 day were flagged as highly likely to be the true period. If there was exactly one pair of flagged periods, the average of the W1 and W2 periods was taken to be the true period for both bands. If there were two or more pairs of flagged periods, then the pair of periods with the highest W1 PSD peak was averaged and taken to be the true period. If no periods were flagged, then the lists are once again considered separately.

Next the period with the highest PSD peak for each band was marked as the final candidate. If both bands contained a final candidate, the two candidates must be farther than 0.05 day apart, since otherwise they would have been flagged in the previous step. Since it is forbidden for a single source to have two distinct periods, the source must be rejected from the database. If a final candidate exists for only one band, then that period is taken to be the true period for the source.

It is important to note that both period-finding algorithms returned a surprisingly high number of RR Lyrae stars with periods between 0.1 and 0.2 day. Oftentimes, a candidate list had a low period in this range along with a more typical period in the 0.4 to 0.6 day range. These two candidate periods resulted in harmonic models with very similar harmonic residuals. Since WISE completes one orbit in 90 minutes, which is 0.0625 day, the time between successive observations is most often that value. This results in an observation bias that may confuse the period-folding algorithms into favouring periods shorter than 0.2 days. To avoid this bias, the period range of RR Lyrae in the catalog was restricted to be between 0.2 and 1 day.

3.2. Creating Robust Models

To remove fluctuations in the light curve shape that are actually instrumental inconsistencies and random errors in photometry, a bootstrap method was used on each of the sources for which
a period was determined. In addition, this process better incorporates error into the model. For a
single source, the magnitude of each data point in the raw light curve was perturbed. The perturbed
magnitude was chosen randomly from a Gaussian distribution of possible magnitudes centred at
the original magnitude with a standard deviation equal to the photometric error of that point.
Once each data point was perturbed in this somewhat random way, a perturbed harmonic model
was produced once again using the Lomb-Scargle method, fixing the period to a range of 0.01 day
centred at the derived period. This process was repeated 500 times, and the average of these 500
perturbed models is referred to as the robust model. In addition, the standard deviation of the
perturbed models was calculated to obtain an error measurement for the robust model, which is
referred to as the robust model error. Four examples of robust models in both bands are shown in
Figure 3. This process provides an error measurement at any point in the model light curve, and
the robust model error serves as an assessment of the overall quality of the model.

If the ratio of the robust model error to the model amplitude for the W1- or W2-band light
curve is no greater than 0.02, the source is added to the final catalog of WISE RR Lyrae stars.
The resulting catalog contains 3740 sources with W1 light curve models and 1584 sources with W2
light curve models. Of these, 1424 sources have light curve models in both bands. The average
number of epochs is 18 for a source in W1 and 19 for a source in W2. The average mean magnitude
and average amplitude and 12.468 mag and 0.220 mag for W1 and 11.417 mag and 0.216 mag for
W2. Distances and distance uncertainties were calculated using mid–infrared period–luminosity
relations (Klein et al. in prep.). These period–luminosity relations are dependent on the RR Lyra
subclass. If a source had been previously classified, the previous classification was used. Otherwise,
sources with a period less than 0.4 day were tentatively classified as RRc and sources with periods
greater than 0.4 day were classified as RRab. If a source had a distance value in both the W1- and
W2-band, these two distances were averaged and the uncertainties propagated accordingly. The
final catalog of WISE RR Lyrae stars is available in electronic format as supplemental material.

4. Characterisation of RR Lyrae in the Mid-IR

To provide a comparison for mid-IR light curves, a sample of optical light curves of RR Lyrae
stars was obtained from the OGLE-III Catalog of Variable Stars (Soszyński et al. 2009). The
sample contains 23,920 RRab, RRc, and RRe stars in the LMC, with a distribution of 74% RRab,
21% RRc, and 5% RRe. Here, RRe stars are defined as low amplitude RR Lyrae stars with periods
of less that 0.3 day, and the distribution between RRc and RRe stars is more or less continuous
(see Figure 3 in Soszyński et al.). Since the distinction is difficult to make with WISE data, for
all further discussion stars classified as RRe by the OGLE-III Catalog are grouped in the RRc
subclass. Using this convention, the OGLE-III LMC sample has a distribution of 74% RRab and
26% RRc. The LMC yields a large, complete, and uniform sample and the effects of Galactic dust
are minimised. The two filters used are V-band centred at 551 nm and I-band centred at 806 nm,
the latter of which may be considered to probe near-infrared wavelengths.
Fig. 3.— Several sources from the catalog of WISE RR Lyrae stars are shown, with the $W_1$-band on the left and the $W_2$-band on the right. For the bootstrapping process, 500 perturbed models are generated by perturbing the magnitude of each data point and fitting a harmonic model with 8 harmonics to each new set of points. 50 of these models are drawn as green curves. The robust model, drawn as a black curve, is the average of the perturbed models. This model is used in all further analysis. The standard deviation of the perturbed models, $\sigma_b$, gives an error for the robust model as a function of phase. The average standard deviation over one cycle of the light curve, $\langle \sigma_b \rangle$, is an effective quality measure of the light curve.
Harmonic light curves for the OGLE-III sample were created using a methodology similar to the creation of the WISE harmonic models. First, the light curve for each source was perturbed 100 times as in Section 3.2. The Lomb-Scargle method was used to find the period with the highest power spectral density within a range of 0.01 day centred at the period provided by OGLE. A harmonic model with up to 8 harmonics was then computed for this period. The average of the 100 perturbed harmonic models was taken to be the robust model. To ensure that these models fit the light curve well, a model rejection process was run based on the percentage of outlying light curve points. Since the $V$-band light curves tended to have far fewer measurements and far smaller measurement errors than $I$-band light curves, different outlier conditions were used for the two bands. $V$-band light curves with over 10% outlying data points were rejected, where an outlying point was defined to be over 8 times its measurement error from the model. $I$-band light curves with over 5% outlying data points were rejected, where an outlying point was defined to be over 4 times its measurement error from the model. Note that this procedure eliminated most of the RR Lyrae undergoing Blazhko–modulation from the sample, since these light curves are characterised by a spread in the light curve shape, likely resulting in more outliers.

The final optical-comparison sample is comprised of $\sim$13,000 $V$-band light curve models, with 78% RRab and 22% RRc, and $\sim$18,500 $I$-band light curve models, with 75% RRab and 25% RRc. A histogram of the periods in the full WISE and OGLE catalogs is shown in Figure 4. In the figure, the secondary peak representing subclass RRc stars that occurs between 0.25 and 0.4 day in optical wavelengths is missing from the mid-infrared histogram. This occurs because subclass RRc stars have lower amplitudes than RRab stars but similar photometric errors, and the quality cuts in the light curve selection process restricted the ratio of average photometric error to model amplitude to be greater than 0.2. Thus, many RRc sources were omitted from the sample due to observational limitations.

One prominent difference between RR Lyrae stars in optical bands and mid-IR bands is that the mean amplitude decreases as the wavelength increases. Figure 5 depicts period–amplitude relations in the four wavebands: the OGLE $V$- and $I$-bands, and the $W_1$ and $W_2$-bands. Note that the 0.2 day cutoff in period in the two WISE plots is directly induced (see the end of Section 3.1), while the 0.1 amplitude cutoff can be explained as an artefact of the fixed mean error-to-amplitude ratio. This suggests that there exist RR Lyrae stars with even smaller $W_1$- and $W_2$-band amplitudes, but these are difficult to detect with WISE since the average photometric error is on the order of the amplitude of the light curve. In addition, the WISE plots only show the sources in the final WISE RR Lyrae database that have been classified into the RRab or RRc subclasses by previous surveys, which is 38% of the catalog. While all four plots present a distinction between the RRab and RRc subclasses, these subclasses are more distinguished in the $V$- and $I$-bands and they begin to blend together in the $W_1$- and $W_2$-bands. The average amplitude is 1.147 for RRab and 0.546 for RRc in the $V$-band, 0.595 for RRab and 0.279 for RRc in the $I$-band, 0.225 for RRab and 0.156 for RRc in the $W_1$-band, and 0.22 for RRab and 0.19 for RRc in the $W_2$-band. These tight period–amplitude relations in mid-IR bands may be useful for identification and classification purposes.
Fig. 4.— Histograms of periods for the full WISE and OGLE databases. For the WISE period–finding process, candidate periods were restricted to be between 0.2 and 1 day. The y-axis plots the probability density of period bin, which will be 1 when integrated over all period bins.

Figure 6 demonstrates the evolution of light curve shape as a function of period in the four wavebands. The space of possible periods for RR Lyrae stars, 0.2 to 1 day, was divided into six period bins: 0.2 to 0.3 day, 0.3 to 0.35 day, 0.35 to 0.4 day, 0.4 to 0.6 day, 0.6 to 0.7 day, and 0.7 to 0.9 day. Sources that fall into the first three bins are mostly subclass RRc stars and sources that fall into the last three bins are mostly subclass RRab, although this period boundary is not strictly defined. All of the light curves in the database are normalised to attain their minimum magnitude at 0 and maximum magnitude at 1. The normalised light curves are then grouped according to period bin. Each coloured curve represents the average of all light curves in the corresponding period bin. In the $V$- and $I$-bands, there are two distinct groups: the light curves in the first three period bins and in the last three period bins group together. However, in the $W1$- and $W2$-bands, the two groups vanish, and the average light curves for all of the period bins become undistinguishable.

Figure 7 adds a visual representation of the average amplitude of the period bins. Each coloured light curve, as presented in Figure 6 with a normalised flux amplitude of 1, is now scaled by the average amplitude of all the light curves in the period bin. While Figure 6 only provides information...
Fig. 5.— Contour plot period–amplitude relations for RR Lyrae stars in four wavebands. RRab subclass stars are denoted by the purple contours and RRc subclass stars are denoted by the orange contours. Contour lines are drawn independently for each subclass and have levels at the 99, 95, 90, 80, and 70th percentile. The data shown in the two upper panels are a subset of the OGLE-III Catalog of Variable Stars in the LMC. The $V$-band sample contains 10,101 RRab and 2,690 RRc and the $I$-band sample contains 13,761 RRab and 4,203 RRc sources (see section 4 for details on the sample selection process). The two lower panels are a subset of the final WISE RR Lyrae catalog, showing only sources which have subclass identification from previous surveys, which is 38% of the catalog. The $W_1$-band subset contains 2,362 RRab and 288 RRc sources and the $W_2$-band subset contains 1,229 RRab and 187 RRc sources.

about light curve shape, Figure 7 incorporates information about the light curve amplitude. Both of these features evolve from optical to mid-infrared wavelengths.

The information of Figure 7 is further summarised in Figure 8. Here, the number of period bins is increased to 12 in order to provide a meaningful plot, and the average amplitude of light curves in each period bin is plotted as a function of the average period for each bin. At shorter wavelengths, light curves with periods of around 0.45 day exhibit the largest amplitude, while in mid-IR wavelengths, all light curves express roughly the same amplitude regardless of period.

In optical bands, the RRab and RRc subclasses of RR Lyrae stars are defined by the amplitude of the light curve and the light curve shape. However, in the mid-IR, these two classes appear less
Fig. 6.— Comparison of light curve morphology in optical and infrared bands. Optical light curves are obtained from $I$-band and $V$-band photometry of subclass RRab and RRc RR Lyrae in the LMC from the OGLE-III Catalog of Variable Stars. Light curve models have been normalised to be 0 at maximum magnitude and 1 at minimum magnitude. Light curves have been grouped into 6 bins according to their period. All light curves in a period bin were averaged to create the coloured curves. In the $V$- and $I$-bands there is notable variation between light curves with periods less than 0.4 days and light curves with periods greater than 0.4 days. The former are sinusoidal, while the latter are triangular and asymmetric. In the $W_1$- and $W_2$-bands, the light curve shapes are homogeneous, with no appreciable difference between the two subclasses.

Fig. 7.— Average light curves in Figure 6 are shown scaled by the mean amplitude of all light curves in the period bin. In $I$- and $V$-bands, RRab and RRc subclasses are distinguished by their amplitudes, while in the $W_1$- and $W_2$-bands, there is no significant variation in amplitude within the subclasses.

distinct, and accurate subclass classification based solely on mid-IR photometry becomes difficult. The reason for this distinction in optical light curves is due to the pulsation mechanism of RR Lyrae stars. During pulsation three different properties of the star are changing: the bolometric magnitude, the radial velocity, and the temperature (Bono & Stellingwerf 1994). Observing a star
Fig. 8.— Binned period-amplitude plots for OGLE-III (V-band and I-band) and WISE (W1-band and W2-band) light curves. The period bins used are 0.2 to 0.3, 0.3 to 0.35, 0.35 to 0.4, 0.4 to 0.45, 0.45 to 0.5, 0.5 to 0.55, 0.55 to 0.6, 0.6 to 0.65, 0.65 to 0.7, 0.7 to 0.75, 0.75, 0.8, and 0.8 to 0.9 day. The average amplitude of all light curves in the period bin is drawn as a function of the average period of all light curves in each period bin. One standard deviation is drawn as the dashed coloured lines. The OGLE-III light curves have a clear maxima in the 0.45 to 0.5 bin, while the WISE plots remain relatively flat across the period bins.

at different wavelengths probes different stellar properties. The optical light curve from an RR Lyrae star is strongly affected by the temperature variation, whereas the mid-IR light curve is dominated by the variation in radius. This results in a triangular feature observed in longer-period optical light curves, corresponding with the RRab subclass. Since the mid-IR light curves are more strongly affected by the harmonic radial oscillation, the triangular feature is diminished at these wavelengths and the shape of these curves is more symmetric.

5. Mid-IR Templates

5.1. Procedure for Constructing Templates

The next goal was to create a set of templates that represents typical light curve shapes in the WISE RR Lyrae catalog. A visual investigation of entire catalog reveals that there are several key features that distinguish light curves. These features include the slope of the light curve before it
attained its maximum (in the region approximately 0.1 to 0.2 in phase space), the phase location
of the maximum, the full width at half maximum, and the amount of symmetry in the light curve.
The basic procedure was to group light curves based on similarity of shape, and then take the
average of each group of light curves as a template for that group.

A challenge to the application of this methodology is the possibility that light curve shapes
naturally form a continuous spectrum. For any feature, there is a maximum and minimum value,
but any value between these is allowed. For example, suppose that you want to make light curve
templates using only one feature: the phase location of the maximum brightness. Some light curves
are left leaning and some are right leaning, and it is not difficult to imagine that there exists a light
curve with any location of maximum brightness between the two extremes. Template clustering
on this single feature amounts to picking N values for the feature between the maximum and the
minimum, and grouping each light curve into a bin. The situation is complicated by the necessity
of constructing these clusters based on all the features simultaneously, not just a single feature.

The next challenge regards the optimal number of templates to produce. If a light curve
lies halfway between two bins, it will be a good candidate for either bin. Employing many bins
eliminates some ambiguity, but using fewer bins will make future analysis more meaningful. It was
found that 10 to 15 bins is optimal.

To ensure that the set of templates is representative of true RR Lyrae light curve shapes,
a sample of high quality light curves from the full WISE RR Lyrae catalog was selected for the
template construction process. The sample was created by selecting light curves with a minimum
of 16 observations, a maximum robust model error of 0.04 magnitudes (see Section 3.2), and a
maximum gap in phase coverage of 15%, which is 0.94 radians. A total of 1633 W1 light curves and
840 W2 light curves was used. The main feature of interest for this process was light curve shape,
since this is an intrinsic property of the source while the mean magnitude is dependent on the
distance at which observations are made. For this reason, light curves were normalised to achieve
maximum magnitude at 0 and minimum magnitude at 1.

Agglomerative hierarchical clustering (Gower & Ross 1969) was used to group the light curves
in the sample. First, a distance between each pair of light curves was computed using the Euclidean
metric. Hierarchical clustering was used to create a dendrogram from the condensed distance
matrix. In other words, each light curve was initially grouped into its own cluster and the algorithm
repeatedly combined the two closest clusters, eventually resulting in a single cluster containing the
entire set of elements. The linkage criteria used in this step was average linkage clustering, which
is to say that distance between any two clusters was determined by the average of all distances
between pairs of objects in the clusters. The clustering algorithm was set to stop merging clusters
once 23 nodes had been formed. Next, any cluster in the W1- band containing less than 15 sources
and any cluster in the W2- band containing less than 10 sources was discarded. This procedure was
used on both WISE bands and resulted in 10 clusters for each band. The average of each cluster
was taken to be one light curve template.
Figure 9 shows the 10 templates created for each of the W1- and W2-bands. For each band, four sample light curves are displayed, colour-coded by their associated template. Each set of templates appears consistent with a continuous distribution of light curve shapes. The main feature that varies throughout the set is the phase location of the maximum brightness, but this is not the only feature useful for distinguishing light curve shapes. In the W1-band, templates 8 and 9 have a very close location of maxima (at 0.494 and 0.504 in phase space, respectively) but they have a different value of the full width at half maximum (0.698 and 0.553, respectively). The complexity of the light curve shape provides justification for the method of hierarchical clustering on the whole template shape, as opposed to a single feature. The templates are provided in electronic format along with the other supplemental material.

Fig. 9.— W1- and W2-band light curve templates resulting from hierarchical clustering methods on light curve shape. The template colour is based on the relative location of maxima, with red indicating the template with the leftmost maximum and purple indicating template with the rightmost maximum. The embedded plots show real light curves from the corresponding template. For these small plots, the light curve spans one full phase cycle and the amplitude is scaled to be 0 at maximum magnitude and 1 at minimum magnitude.

5.2. Using Templates for Mean Magnitude Estimation

A practical application of light curve templates is the recovery of light curve information for poorly-sampled sources. Suppose a source with a known period is sampled a handful of times in the W1- or W2-band. The objective is to estimate the mean magnitude, which is needed to apply period–luminosity relations for the distance measurement. The most straightforward way
to estimate the mean magnitude is to average all of the magnitude measurements in the light curve. However, this result may be inaccurate due to sampling biases and asymmetries in the light curve. A more precise method is to perform a Fourier analysis and fit a harmonic model to the light curve, then take the mean magnitude of the model. This method may result in over-fitting, or fitting extra bumps and dips that are not representative of the true light curve. Over-fitting is especially likely if the light curve is poorly sampled, has gaps in phase coverage, or has a low signal-to-noise ratio. An alternative method is template fitting: for each normalised template in a class- and bandpass-specific catalog, scale the period, amplitude, and minimum magnitude in order to minimise the residual difference between the light curve and the template. The template with the lowest residual is chosen as the best match and the mean magnitude of that template fit is taken to be the mean magnitude of the source.

To demonstrate the effectiveness of the three mean magnitude estimation methods (straight mean, harmonic model mean, and template mean), each method was used on a sample of artificial light curves, varying the number of epochs of observation (N) and the photometric sensitivity ($\sigma$). First, a random model light curve was selected from the full WISE RR Lyrae catalog and resampled randomly at N positions. Noise was added to the light curve by taking a Gaussian sample of each magnitude datapoint with a standard deviation of $\sigma$, similar to the bootstrap perturbations used in Section 3.2. An artificial light curve was created for the following parameter values: N = 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 40, 50, 60 and $\sigma$ = 0.025, 0.05, 0.075, 0.1, 0.15 magnitudes. The mean magnitude of each artificial light curve was estimated using the three methods described above. For each value of N and $\sigma$, 500 artificial light curves were generated.

For each of the three methods, for each value of N and $\sigma$, a mean magnitude residual was calculated by taking the absolute value of the difference between the recovered mean and the actual mean of the model. The method that most accurately predicts the mean magnitude of the light curve will have the lowest overall mean magnitude residual. Figure 5.2 shows the median and median absolute deviation for each mean magnitude estimation method, keeping $\sigma$ fixed at 0.05 magnitude.

As seen in Figure 5.2, template fitting generally provides more accurate mean magnitude values than harmonic modelling. When the number of data points is larger than 30, the difference between the two methods becomes negligible. The straight mean has the largest scatter and the lowest accuracy and should not be used for mean magnitude estimates, even with a high number of data points. These findings are consistent with Kovács & Kupi (2007), despite significant differences in methodology used for template creation.

The templates can be used as a direct application for mean magnitude estimates for light curves with less than 30 points observed with JWST. If measuring the distance to an RR Lyrae is the goal of observing the source, then the only information needed is the mean magnitude of the source. Templates make it possible to recover this measurement with the same accuracy as harmonic modelling, but requiring fewer observations, essentially reducing the necessary telescope
Fig. 10.— A comparison between the accuracy of three methods for deriving the mean magnitude of a light curve. For each number of light curve data points N, 500 random light curves from the WISE RR Lyrae database were selected and sampled randomly at N points. An artificial light curve was created by adding noise to the magnitude value, taking a Gaussian magnitude sample with the magnitude value for the mean and $\sigma$ for the standard deviation. The mean magnitude of the artificial light curve was recovered with each of the three methods, and a mean magnitude residual was calculated for each method. Each filled circle represents the median values of each set of mean magnitude residuals, with dashed lines drawn one median absolute deviation.

Time. When the given templates are not directly applicable because of the bandpass used is not the W1- or W2-band, the methodology is transferrable. While the process of generating a complete set of light curves requires an initial sample of well-observed light curves, it will cut down observation time for the rest of the light curves. It may be economical to invest observation time to generate a templates set and then use the set to obtain accurate mean magnitude estimates for all subsequent sources with fewer epochs of observation.
6. Conclusions

We have presented a catalog of 3740 RR Lyrae stars in WISE with derived periods and distances. This catalog can be used for creating dust maps of the galaxy up to distances of 14 kiloparsecs. Most certainly the catalog does not contain every RR Lyra star observed by WISE. However, the sample is large enough that it summarises the typical shapes of mid-infrared RR Lyrae light curves. The catalog may be used as a training set in a machine learning initiative to identify all RR Lyrae stars in the full WISE database.

Harmonic light curve models were created for the sources in the catalog. To study light curve morphology, these models were compared with similarly-derived harmonic models from the OGLE-III catalog of variable stars. There are two main distinctions between the models in OGLE-III bands and those in WISE bands. First, light curve amplitudes are higher for visible wavelengths. For the V-, I-, W1- and W2-bands, the mean RRab amplitudes are 1.147, 0.595, 0.228, and 0.222 magnitudes and the mean RRc amplitudes are 0.546, 0.279, 0.162, and 0.186 magnitudes. The second difference is that models in visible wavelengths display significant variation in light curve shape depending on the source period, while models in the mid-infrared have rather uniform light curve shapes. This trend in light curve morphology is depicted in Figure 6. Evolution in shape as a function of bandpass arises because different bandpasses probe different stellar layers. Observations at long wavelengths yield information about the relationship between the period and the radius of the star, which undergoes harmonic oscillation. Observations at shorter wavelengths are sensitive to temperature variation, which exhibits sawtooth–like patterns.

We have created 10 templates for typical light curve shapes in the W1- and W2-bands, employing statistical clustering methods. We have investigated the application of these templates as a method for the estimation of mean magnitude. Template fitting generally improved on mean magnitude estimates from harmonic modelling methods, but the level of improvement was negligible when used on sources with over 30 data points. With a simulated photometric sensitivity of 0.05 magnitudes, harmonic modelling on a light curve with 25 observations will typically estimate the mean magnitude to within 0.006 magnitudes of the actual value. The same accuracy can be achieved using template fitting on a light curve with 18 observations. This improvement may become significant during large RR Lyrae observation campaigns.

Acknowledgements: We thank Nathaniel R. Butler for providing the implementation of the Lomb-Scargle algorithm that was used in this paper. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.
REFERENCES

Bailey, S. I. 1902, Annals of Harvard College Observatory, 38, 1
Barning, F. J. M. 1963, Bull. Astron. Inst. Netherlands, 17, 22
Blanco, B. M. 1984, AJ, 89, 1836
—. 1992, AJ, 103, 1872
Blanco, B. M., & Blanco, V. M. 1997, AJ, 114, 2596
Bono, G., & Stellingwerf, R. F. 1994, ApJS, 93, 233
Cacciari, C., & Renzini, A. 1976, A&AS, 25, 303
Chaboyer, B. 1999, Post-Hipparcos Cosmic Candles, 1 No. 111 (Dordrecht: Kluwer)
Clementini, G., Tosi, M., Bragaglia, A., Merighi, R., & Maceroni, C. 1995, MNRAS, 275, 929
Collinge, M. J., Sumi, T., & Fabrycky, D. 2006, ApJ, 651, 197
Debosscher, J., Sarro, L. M., Aerts, C., Cuypers, J., Vandenbussche, B., Garrido, R., & Solano, E. 2007, A&A, 475, 1159
Derue, F., et al. 2002, A&A, 389, 149
Freedman, W. L., Madore, B. F., Scowcroft, V., Burns, C., Monson, A., Persson, S. E., Seibert, M., & Rigby, J. 2012, ApJ, 758, 24
Gower, J. C., & Ross, G. J. 1999, Journal of the Royal Statistical Society. Series C (Applied Statistics), 18, 54
Groenewegen, M. A. T., Udalski, A., & Bono, G. 2008, A&A, 481, 441
Hoffman, D. I., Cutri, R. M., Masci, F. J., Fowler, J. W., Marsh, K. A., & Jarrett, T. H. 2012, AJ, 143, 118
Kinman, T. D., Mahaffey, C. T., & Wirtanen, C. A. 1982, AJ, 87, 314
Kinman, T. D., Wirtanen, C. A., & Janes, K. A. 1965, ApJS, 11, 223
Klein, C. R., Richards, J. W., Butler, N. R., & Bloom, J. S. 2011, ApJ, 738, 185
—. in prep., ApJ to appear
Kovács, G., & Kupi, G. 2007, A&A, 462, 1007
Lomb, N. R. 1976, Ap&SS, 39, 447
Longmore, A. J., Fernley, J. A., & Jameson, R. F. 1986, MNRAS, 220, 279
Madore, B. F., et al. 2013, ApJ, 776, 135
Maintz, G. 2005, A&A, 442, 381
Meixner, M., et al. 2006, AJ, 132, 2268
Miceli, A., et al. 2008, ApJ, 678, 865
Oosterhoff, P. T. 1939, The Observatory, 62, 104
Percy, J. R. 2007, Understanding Variable Stars (Cambridge University Press)
Pigulski, A., Kołaczkowski, Z., & Kopacki, G. 2003, Acta Astron., 53, 27
Pigulski, A., Pojmanski, G., Pilecki, B., & Szczygieł, D. M. 2009, Acta Astron., 59, 33
Pojmanski, G. 1997, Acta Astron., 47, 467
Reimann, J. D. 1994, PhD thesis, Univ. California, Berkeley
Richards, J. W., Starr, D. L., Miller, A. A., Bloom, J. S., Butler, N. R., Brink, H., & Crellin-Quick, A. 2012, ApJS, 203, 32
Richards, J. W., et al. 2011, ApJ, 733, 10
Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 288, 618
Riess, A. G., et al. 2011, ApJ, 730, 119
Saha, A. 1984, ApJ, 283, 580
Scargle, J. D. 1982, ApJ, 263, 835
Sesar, B., et al. 2010, ApJ, 708, 717
Smith, H. A. 2004, RR Lyrae Stars (Cambridge University Press)
Soszyński, I., Udalski, A., Szymański, M. K., Kubiak, J., Pietrzyński, G., Wyrzykowski, Ł., Ulaczyk, K., & Poleski, R. 2010, Acta Astron., 60, 165
Soszyński, I., et al. 2009, Acta Astron., 59, 1
—. 2011, Acta Astron., 61, 1
Storm, J., Nordström, B., Carney, B. W., & Anderson, J. 1994, A&A, 291, 121
Suntzeff, N. B., Kinman, T. D., & Kraft, R. P. 1991, ApJ, 367, 528
Szekely, P., Kiss, L. L., Jackson, R., Derekas, A., Csak, B., & Szatmary, K. 2007, A&A, 463, 589
Wetterer, C. J., McGraw, J. T., Hess, T. R., & Grashuis, R. 1996, AJ, 112, 742
Wright, E. L., et al. 2010, AJ, 140, 1868
Zorotovic, M., et al. 2010, AJ, 139, 357