A kinematic check of the distance scale of Galactic Cepheids

Brian Thorsbro

Lund Observatory
Lund University

Abstract

A selection of well known period luminosity relations for Cepheid variables are studied under the assumption that they have a mean vertical velocity of zero in the Galactic plane. Proper motion data from Hipparcos and Tycho-2 catalogues are coupled with the distances obtained from applying the period luminosity relations on a database of known classical Cepheids to calculate their velocities. A statistical analysis is performed on the velocities to identify the best sample of Cepheids to perform the kinematic check on. Two samples are found depending on whether the period luminosity relation has a colour term or not. Examining the samples leads to rejecting one sample and the other sample shows that in order to agree with the above assumption the solar vertical velocity $w_\odot$ is unlikely to be around the conventional $7 \ \mathrm{km \, s^{-1}}$, but has to be lowered to around $6.5 \ \mathrm{km \, s^{-1}}$ or less; or the period luminosity relations have to be corrected on the order of -0.2 mag, i.e. a little brighter than previously assumed. It is further argued that the correction to the period luminosity relations is more in agreement with other studies than lowering the solar velocity.

Populärvetenskaplig beskrivning

I den här uppsatsen undersöker jag en ny metod för att testa period-luminositetsrelationen för cepheider. Detta har stort intresse eftersom denna relation är en viktig faktor vid avståndsbestämning på kosmologisk skala.

Avståndet till en viss stjärna kan bestämmas om vi känner den ljusmängd stjärnan producerar, alltså dess luminositet, så att vi kan jämföra den med den mängd ljus våra teleskop uppfångar.

En cepheid är en stjärna vars luminositet varierar över en bestämd period och för vilken denna period är direkt kopplad till stjärnans luminositet. Genom att tålmodigt observera en cepheid under allt från några dagar till flera månader kan perioden bestämmas. Med hjälp av period-luminositetsrelationen kan man så beräkna luminositeten, vilken i sin tur ger oss avståndet till stjärnan såsom beskrivet ovan.

Vi vet inte det exakta förhållandet mellan period och luminositet från teoretiska beräkningar, så det måste bestämmas empiriskt. Det görs genom att mäta perioden för cepheider i vår närhet, till vilka vi kan bestämma avståndet på annat sätt. När vi vet avståndet till en cepheid och dess skenbara ljusstyrka kan vi beräkna dess luminositet. På så sätt är det möjligt att bestämma luminositeten för cepheider i vår närhet och tillsammans med observationer av perioden kan vi komma fram till relationen mellan period och luminositet. Denna process innehåller ett antal osäkerheter vilket medför att många olika versioner av relationen har presenterats genom åren.

Den nya metod jag undersöker i den här uppsatsen bygger på hypotesen att cepheider i vår närhet rör sig slumpmässigt vinkelrätt (vertikalt) mot Vintergatans plan. Det innebär att deras genomsnittliga vertikala hastighet bör vara noll. För att kunna beräkna cepheidernas hastighet krävs dels avståndet till dem, dels observationer av hur de rör sig relativt oss. Den första delen är information vi har från de olika versionerna av period-luminositetsrelationen, den andra delen kommer från observationer gjorda av Hipparcos-satelliten. Tillsammans gör de det möjligt att beräkna varje cepheids hastighet relativt oss. Genom att därefter applicera solens hastighet kan vi komma fram till cepheidens hastighet vinkelrätt mot det galaktiska planet. Den ursprungliga hypotesen är att denna är noll, om den inte är det beror det antingen på att hypotesen är inkorrekt eller på att den undersökta period-luminositetsrelation är felaktig.

Resultatet är därmed en bedömning av hur väl de olika versionerna av relationen mellan period och luminositet stämmer överens med den enkla hypotesen att cepheidernas genomsnittliga vertikala hastighet är noll.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cepheids are stars with a peculiar property of regularly changing their brightness with a consistent period. It turns out that this property can be used to determine the distance to such Cepheids. Given an estimate of about 20000 Cepheids in the Milky Way (Windmark et al. 2011) as well as the more than 3000 known Cepheids in the Large Magellanic Cloud (Joshi & Joshi 2014) it is no surprise that Cepheids are found in other galaxies. This makes Cepheids particularly useful for determining extra-galactic distances. Cepheids have historically played and still play an important role in astronomy with regards to the effort of learning more about the distance scale of the universe we live in (Hubble 1925; Marconi 2009).

However, there are disagreements on how good the distances are that we can deduce from this property of Cepheids. So this study looks into several results from previous studies on Cepheids and examines how well these results can be applied to classical Cepheids that have been observed within a distance of a few kpc from the Sun. This is done by making the assumption that these Cepheids on average has zero vertical velocity with respect to the Galactic plane. Kinematics together with statistics is used to analyse the data and reach a result.

In Ch. 2 the property that makes Cepheids interesting is introduced and described. In Ch. 3 the thesis assumption is presented and it is explained how the kinematics and statistics are applied. The data selected as foundation for this study are presented in Ch. 4 and the results of the study follow in Ch. 5. Finally the results are discussed in Ch. 6.

This thesis assumes that the reader is familiar with physics on the level of a third year undergraduate student. Please notice that a basic introduction to some of the astrophysical concepts used throughout the thesis has been provided and can be found in Appendix A.

Chapter 2

Cepheid variables

Cepheids are bright variable stars that can be seen from far away. There is a relation between how bright they are and how their light varies which makes Cepheids very important for distance determination on intergalactic scales. This property of Cepheids is central to this thesis, and thus this chapter is dedicated to introducing Cepheids and explaining how they can be used for determining distances.

First the concept of variable stars is introduced. Cepheids are pulsating variable stars, and what that entails is described next. The pulsation mechanism leads to the so-called period-luminosity-colour (PLC) relation, which is the key part to using Cepheids for distance determinations, so this relation is explained in some detail at the end of the chapter.

2.1   Variable stars

Stars whose brightness changes as we observe them over time are called variable stars. A supernova that appears like a new star and then fades away is a variable star. A star that interchangeably appears dim at some times and bright at other times is also a variable star and perhaps the type of variable star that is most often referred to. The change in brightness can be caused by many different phenomena, for instance it could be an eclipsing binary where one star at times hides behind another star making their combined light less bright. Or it could be a star that in some way is unstable causing it to change its brightness regularly. Unstable stars are not that uncommon since most stars go through an instability period towards the end of their life.

In Fig. 2.1 many of the conventional types of variable stars have been mapped on a Hertsprung-Russel diagram. The Cepheids important for this study are marked as classical Cepheids in the figure.

2.2   Pulsating variable stars

Stars that regularly change their brightness due to some instability are called pulsating stars. By observing how the Doppler shift of the light from pulsating stars changes over
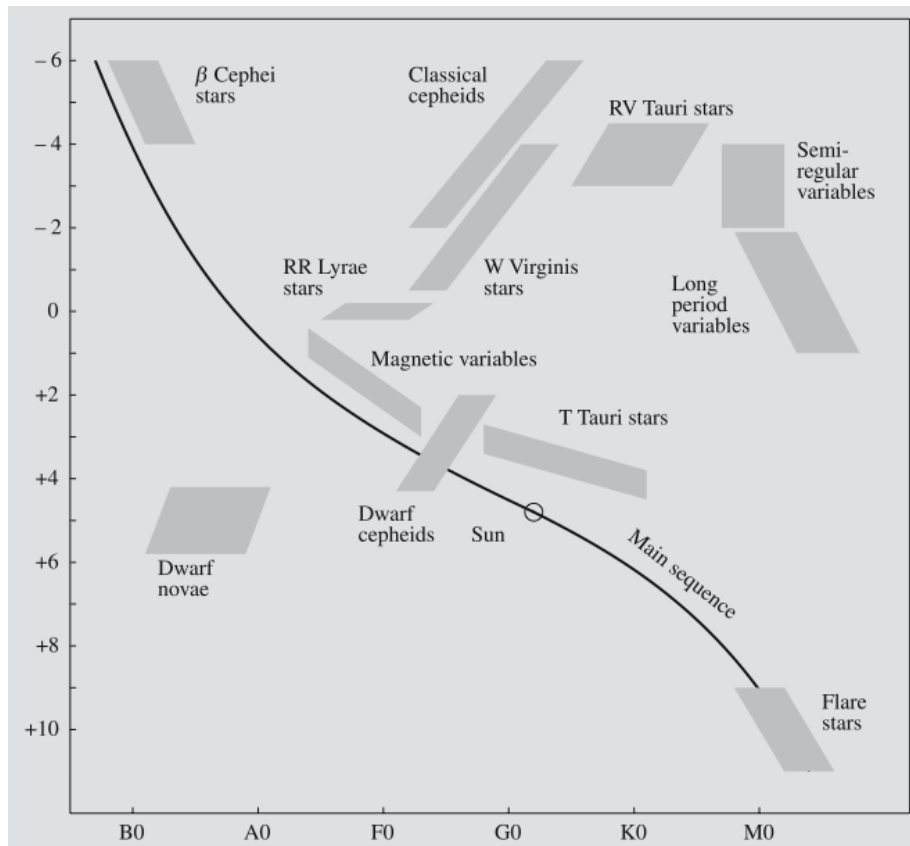
Figure 2.1: The location of variables in a Hertsprung-Russel diagram. (Karttunen et al. 2007, Fig. 13.2)

time it is apparent that the stars change their size rather rapidly during one period of pulsation. A doubling of the star's radius is not unusual. Apart from magnetic variables, flare stars and dwarf novae, all of the types of variables shown in Fig. 2.1 are pulsating variable stars.

The pulsation mechanism can be explained by thinking of the outer layer of a pulsating star as a control mechanism that regulates if the radiation produced by the nuclear fusion in the centre of the star is allowed to escape. When the star is contracted the outer layer consists of doubly ionized helium which is opaque and thus prevents radiation from escaping, but the continued production of radiation will cause the pressure to rise forcing the star to expand in size. As the star expands the outer layer cools and go from consisting of doubly ionized helium to consists of singly ionized helium. Singly ionized helium is transparent to the radiation which can then escape, causing the radiation pressure to drop and thus the star begins to contract. The contraction will heat up the helium causing it to be doubly ionized again and then the cycle begins anew. The inertia of the outer layer will cause this to become a self oscillatory phenomenon driven by the energy coming from the fusion process at the core, which means that it can be quite a stable oscillation for a long time. This description with ionized helium as the control mechanism has been shown
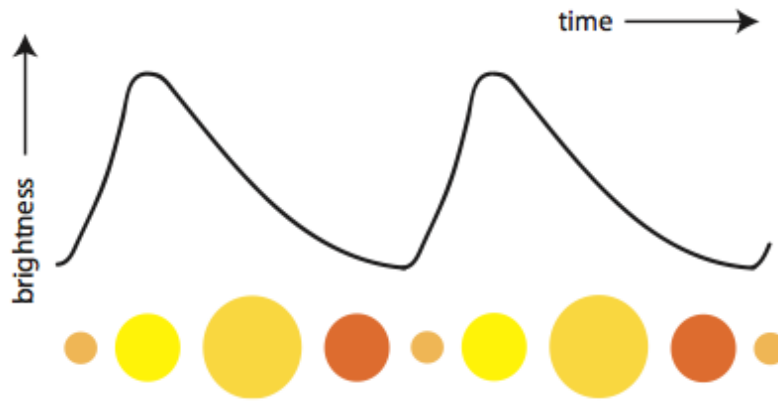
Figure 2.2: The luminosity cycle of a pulsating star. The disks illustrate how the Cepheid changes its colour and size during a period of pulsation. (Jenkins 2013, Fig. 31)

to be a good model for classical Cepheids (Zhevakin 1963), but it is fair to say that the mechanism is not understood in full detail (Madore & Freedman 1991). For other pulsating stars the pulsation mechanism may differ from classical Cepheids. And to complicate the picture even more, some pulsating stars are oscillating in higher harmonics, called overtone pulsators.

The luminosity of the pulsating star is strongest just as the trapped light contained behind the opaque layer is allowed to escape, giving a luminosity curve as illustrated in Fig. 2.2.

Arguably the most interesting aspect of pulsating stars is the Ritter relation (Ritter 1879),

$$P\sqrt{\rho} = Q(P), \tag{2.1}$$

where P is the period of the pulsation (assuming it is not an overtone pulsation), ρ the mean density of the star and Q is a weak function of the period.

It is not an isolated phenomena that the time of the oscillation process is proportional to the inverse square root of the mean density. Infall to a point source of gravity for a homogeneous sphere shares this result. For the homogeneous sphere, where the mass enclosed inside radius r is $M(r) = \frac{4}{3}\pi r^3 \rho$, a test particle is subject to a gravitational acceleration which is

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = -\frac{GM(r)}{r^2} = -\frac{4\pi G\rho}{3} r, \tag{2.2}$$

where ρ is the mean density, t is time, r radius of the sphere and G the gravitational constant. This is a homogeneous second order linear differential equation that can be identified as the equation of motion for a harmonic oscillator (Binney & Tremaine 2008, p. 64). The general solution to this differential equation with the initial condition of zero

velocity is

$$r(t) = k \cos\left(\sqrt{\frac{4\pi G\rho}{3}}\, t\right), \tag{2.3}$$

where k is some constant that requires a second initial condition for it to be determined. The infall time to the centre is a quarter period, i.e. $r(t = P/4) = 0$, which gives

$$\sqrt{\frac{4\pi G\rho}{3}}\, \frac{P}{4} = \frac{\pi}{2} \quad \Leftrightarrow \quad P\sqrt{\rho} = \sqrt{\frac{3\pi}{G}}, \tag{2.4}$$

and a relation similar to the Ritter relation is found.

All stars are subject to gravitational infall and a star like the Sun would collapse in about 20 to 30 minutes if allowed. Fortunately the radiation pressure from the fusion processes in the core prevents this from happening, and the Sun and other main sequence stars are in a so-called hydrostatic equilibrium where these forces keep each other in check. If someone compressed the Sun a little it would resist and expand again, but it would not begin oscillating as friction would quickly dampen it back to hydrostatic equilibrium.

However, for a Cepheid, the hydrostatic equilibrium is confused since the pressure force depends on whether or not the outer helium layer is opaque or transparent. As a consequence of this variation the size of the Cepheids oscillate. Given that the force controlling the infall is gravity it means that inertia governs the time scale of the oscillation, and thus it is not surprising that the Ritter relation is similar to the relation found for gravitational infall.

With stellar models that uses hydrostatic equilibrium (and spherical symmetry) as requirements it is possible to show that the mean density is related to the radius of the star. So for a Cepheid we have a relation between its period of oscillation and its mean radius. This will be useful when looking into the PLC relation.

2.3   The period-luminosity-colour relation

Returning to Fig. 2.1 and looking on the "instability strip" that marks the classical Cepheids, it is clear that the strip is not infinitely thin, so there is not a bijective relation between colour and luminosity, i.e. for a given luminosity there could be several choices of colour. While it was contested for a while whether or not colour should be included in the relation, it has been shown empirically that it is necessary to take colour into account (Martin et al. 1979).

It is possible to get a general outlook on the PLC relation by looking at the Stefan-Boltzmann law for black body radiation (Madore & Freedman 1991). The Stefan-Boltzmann law applies to Cepheids because their radiation curves follows that of a black body radiation curve, which mean that the luminosity can be described as function of radius and effective temperature:

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4. \tag{2.5}$$

L is the luminosity of the Cepheid, R is the radius of the Cepheid, T_{eff} is the effective temperature and σ is the Stefan-Boltzmann constant. The effective temperature is often called the surface temperature even though stars don't have well defined surfaces. Using the definition of absolute bolometric magnitude, M_{bol}, the Stefan-Boltzmann's law can be rewritten to

$$M_{\mathrm{bol}} = -2.5 \log_{10}(k_1 L) = -5 \log_{10}(R) - 10 \log_{10}(T_{\mathrm{eff}}) + k_2, \qquad (2.6)$$

where k_1 is a constant with unit time per energy (like W^{-1}), which wraps up all the constants coming from converting the flux to the luminosity (as well as taking the magnitude difference with the star Vega as described in appendix A), and $k_2 = -2.5 \log_{10}(k_1 4\pi\sigma)$. Physically M_{bol} is unitless, but to keep the equation simple it is not shown how the units cancel out.

In the previous section it was shown that the period of a Cepheid is related to its radius, so the "$-5 \log_{10}(R)$" expression can be converted to a "$\alpha \log_{10} P$" expression. Further it is known that a colour index can be used to describe the effective temperature, which leads to converting "$-10 \log_{10}(T_{\mathrm{eff}})$" to "$\beta(\langle B_0 \rangle - \langle V_0 \rangle)$". Together this leads to the conventional way of expressing the PLC relation, which here is described as the mean absolute visual magnitude taken over a full pulsation period:

$$\langle M_V \rangle = \alpha \log_{10} P + \beta(\langle B_0 \rangle - \langle V_0 \rangle) + \gamma, \qquad (2.7)$$

where $\langle M_V \rangle$ is the mean absolute magnitude of the Cepheid, P the period of pulsation, $\langle B_0 \rangle$ and $\langle V_0 \rangle$ the mean intrinsic colours of the star as per the UBV photometric system. α, β and γ are constants that cover the underlying physics described above. The reason for this format is that the period as well as the intrinsic colours are obtainable through observations, and thus it is possible to obtain observational data and analyse it to empirically determine the constants through a fitting process.

The challenge is to get these constants right. Different studies have been done into the PLC relation and they disagree on exactly what values should be assigned to these constants. In fact many studies do not include a colour term as the authors find it unclear how big a contribution the colour term adds compared to the statistical noise. This will be examined in more detail later in this thesis in Sect. 4.2.

The usage of colour indices is not without problems. There is a phenomenon called reddening, which exists because interstellar dust scatters blue light more than red. So if light from a star travels through a significant amount of dust it will appear both redder and less intense at the observer's end. To solve this problem it is possible to determine empirically what the colour index should be for the observed period and assume that the difference in result is due to reddening. Another possibility is to use the "Wesenheit function", which provides a reddening-free magnitude (see Appendix A).

The constant α is often referred to as the slope and γ as the zero-crossing, since historically the PLC relation did not include the colour term (Leavitt & Pickering 1912). Leavitt & Pickering (1912) found a relation based on the apparent visual magnitude, V, which is

$$\langle V \rangle = \alpha_V \log_{10} P + \gamma_V, \qquad (2.8)$$

with subscripts on the constants to distinguish them from the constants in above PLC relation. Given the distance modulus (μ_V), described in Appendix A as $\mu_V = V - A_V - M_V$, the PLC relation without a colour term is quickly recovered:

$$\langle M_V \rangle = \alpha_V \log_{10} P + \gamma_V - \langle \mu_V \rangle - A_V. \tag{2.9}$$

The A_V term is called the visual extinction and describes the reddening effect, which can be considered to be a constant in the time scale of pulsation. Similarly μ_V can be considered a constant. Thus the three constants in Eq. 2.9 can be combined to the γ constant of the PLC relation. All the uncertainties that arise due to the difficulty of determining reddening is therefore contained in the γ constant. In contrast the period is rather simple to observe and the errors are negligible in comparison. This makes the zero crossing the constant that is most likely to be incorrect when analysing a PLC relation.

Chapter 3

Theory of the kinematic check

The title of this thesis is "A kinematic check of the distance scale of Galactic Cepheids" and this chapter covers the meaning behind the "kinematic check". The foundation of the kinematic check is an assumption about the motions of the Galactic Cepheids, and this is described first. The kinematic check utilizes the velocities of the Galactic Cepheids and how the velocities of the Cepheids are calculated follows next. Finally the theory and implementation of the kinematic check is discussed in detail.

3.1 The thesis assumption

The central assumption for the kinematic check is the assumption that the expectation value of the vertical velocity (perpendicular to the Galactic plane) of the classical Galactic Cepheids is equal to zero, that is:

$$\langle w \rangle = 0, \tag{3.1}$$

where w is the vertical velocity of the Cepheids with respect to the Galactic plane and the angular brackets denote the mean value or expectation. Furthermore it is understood that the distribution of velocities will be symmetric around the expectation value.

The classical Galactic Cepheids are important for this thesis because they are heavy and thus short lived stars. The short life span means that these Cepheids do not have enough time to experience many encounters that change their orbital path in the Galaxy and are therefore expected to follow circular-like orbits around the Galactic centre. Other types of Cepheids, like population II Cepheids and dwarf Cepheids, have long life spans which makes them less ideal for this study. Further, any perturbation that may cause a Cepheid to have a velocity perpendicular to the Galactic plane is equally likely to be "up" or "down" making the distribution symmetric. In other words, the thesis assumption is reasonable.

3.2 Calculating the velocities of the Cepheids

With instruments like the Hipparcos satellite it is possible to measure the proper motion of nearby stars. Proper motion of a star is the motion a star has on the sky (usually expressed in mas yr^{-1}) and can be measured without knowing the distance to the star. Knowing the distance to the star will enable calculation of the vertical velocity (usually expressed in km s^{-1}). However, the Sun's motion has to be taken into account, as the desired velocity needs to be independent of the Sun's motion, but proper motion is measured from our perspective. With a little geometry the calculation is

$$w = \mu d + w_\odot, \tag{3.2}$$

where $w_\odot$ is the vertical velocity of the Sun's with respect to the galactic plane, μ is the vertical proper motion of the observed star, d the distance to the star and w the star's vertical velocity with respect to the galactic plane.

Proper motion measured by the Hipparcos satellite are in equatorial coordinates, which is a spherical coordinate system aligned with the celestial sphere as it was oriented at some specific time, called the epoch of the equatorial coordinate system. However, given the need of this study to calculate with velocities perpendicular to the Galactic plane it is natural to convert the proper motion measurements to the Galactic coordinate system, which is also a spherical coordinate system, but with the Sun at its centre and oriented towards the Galactic centre (Blaauw et al. 1960).

Converting equatorial coordinates to Galactic coordinates is somewhat complicated, but algorithms describing this in detail exist (van Altena 2013, Ch. 4). After such a transformation μ in Eq. (3.2) becomes the proper motion in Galactic latitude.

3.3 The kinematic check

The thesis assumption is that the expectation value of the vertical velocities should be zero with a distribution that is symmetric around this expectation value. The symmetry implies that if a random Cepheid is examined there is a 50% chance that it would have negative velocity and a 50% chance that it would have a positive velocity. Given a sample of n Cepheids, the number of Cepheids that has negative velocity, say it lies in the "negative bin", follows a binomial distribution.

The binomial distribution is described by its probability function, P_{bin}, which gives the probability that k Cepheids out of a sample of n Cepheids all lie in the same bin when there is p probability for one Cepheid to be in that bin. The probability function is defined by

$$\text{P}_{\text{bin}}(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, ..., n. \tag{3.3}$$

As mentioned above the chance to be in any of the bins is 50% so $p = 1/2$, which simplifies the distribution expression a bit to

$$P_{\text{bin}}(k) = \binom{n}{k} \frac{1}{2^n}, \quad k = 0, 1, ..., n. \tag{3.4}$$

A sample of n Cepheids taken out of the population of all Cepheids is descriptive of the full population, but only to a certain degree. The uncertainty is described by the variance of the distribution, which is conventionally[1] written as σ^2. The variance of the binomial distribution is $\sigma^2 = np(1 - p)$, where n is the sample size and p is the probability to be in one of the bins. For $p = 1/2$ it reduces to $\sigma^2 = n/4$. The square root of the variance gives the standard deviation which is then

$$\sigma = \frac{\sqrt{n}}{2}. \tag{3.5}$$

In order to compare samples of different sizes it is more useful to express the distribution of stars in fractions. Such a frequency description is attained simply by dividing by the sample size, which leads to a standard deviation that is

$$\sigma = \frac{1}{2\sqrt{n}}. \tag{3.6}$$

The expectation value is $\frac{1}{2}$, since ideally half of the stars should be in a bin. Together with the standard deviation it is possible to set up an interval that covers the distance of one standard deviation to each side of the expectation value, which is called a (68%) confidence interval. For the above distribution the confidence interval for the fraction k/n is

$$I_{\text{conf}} = \left[\frac{1}{2} - \frac{1}{2\sqrt{n}}, \frac{1}{2} + \frac{1}{2\sqrt{n}} \right]. \tag{3.7}$$

A binomial distribution is a discrete distribution, but a binomial distribution with $np(1 - p) > 10$, that is $n > 40$ with $p = 1/2$, can be approximated with a normal distribution. The larger the n the better the approximation. For a normal distribution the probability that the fraction (of stars in a bin) is in the confidence interval is 68%.

For this study there are three points of interest, the median and the two boundary points of the confidence interval. The median should preferably be zero in order to match the thesis assumption, but the confidence interval gives lower and upper boundary points that should be examined as well. So for example for a sample with 16 Cepheids the boundaries for the confidence interval in terms of fractions is

$$\left\{ \frac{1}{2} \pm \frac{1}{2\sqrt{16}} \right\} = \left\{ \frac{3}{8}, \frac{5}{8} \right\}. \tag{3.8}$$

---

[1]Since the dimension of the variance is the square of the dimension of the mean

Examining the velocity that splits the 3/8 most negative velocities from the rest and the velocity that splits 5/8 most negative velocities from the rest gives two velocities with which a confidence interval in terms of velocties can be constructed. The smaller this interval is the more it is possible to confidently express what the median of the velocity distribution is. Being able to confidently determine if the velocity distribution matches the thesis assumption is obviously very important.

As the sample size increases the confidence interval in terms of fractions will inevitable creep closer to the median giving the impression that the more Cepheids added to the sample the merrier. However, it is important to keep in mind that the velocity distribution may change as Cepheids are added to the sample. If a Cepheid with a large measurement error is added to the sample then the quality of data may be diluted to such a degree that the confidence interval in terms of velocity grows even though the fractions interval shrinks.

A good method to find the best sample is to take all the available Cepheids and sort them based on the error associated to each of them. Such an error could for instance be in the measurements of proper motion measurements which will translate directly into an error in the vertical velocity calculation. The confidence interval in terms of velocity is examined as the Cepheids are added to the sample. As long as adding a new Cepheid to the sample causes the confidence interval in terms of velocties to shrink more Cepheids can be added. Once the confidence interval in terms of velocities starts to grow due to reaching the Cepheids with large errors no more Cepheids should be added.

The elegance of this method is that it is possible to find the sample with the smallest error without the need to know exactly how any errors in proper motion or other sources propagates through the calculations.

The results obtained with this method are in the unit of velocity, which is not very useful when looking at a period-luminosity-colour (PLC) relation which is in the unit of magnitudes. However, it is possible to keep changing the constants in the PLC relation until the found median in the selected sample is zero, i.e. changing the PLC relation until it matches the thesis assumption. As the γ constant is the constant associated with most uncertainties this process can be restricted to changing that constant only. The same can be done for the boundary points of the confidence interval. The result is then a measure of change, with confidence interval, needed for a given PLC relation to match the thesis assumption.

Chapter 4

Data used in this study

176 classical Cepheids in the Galactic neighbourhood have been selected for this study. These Cepheids forms the base on which a kinematic check is performed on ten different period-luminosity-colour (PLC) relations. An important part of the kinematic check is the solar velocity. This chapter describes where the data is obtained from and the post processing that is performed on the data.

4.1 Selected Cepheids

Berdnikov et al. (2000) have compiled a list of 455 Galactic classical Cepheids which is one of the largest catalogues of Cepheids available with good homogeneous BVI colour data. Tammann et al. (2003a) have worked with this catalogue and both removed known overtone pulsators as well as included and corrected colour excess information by Fernie et al. (1995). In this process the data set has been reduced to 321 fundamental mode Galactic classical Cepheids.

I have correlated the 321 Cepheids with proper motion information from the Hipparcos and Tycho-2 catalogues. A total of 176 Cepheids have been found to have a complete dataset needed for the kinematic check, of these 156 have proper motion from the Hipparcos catalogue. The 176 Cepheids are listed in Appendix B.

Technically the Cepheid catalogue of Tammann et al. (2003a) has been downloaded from the VizieR star catalogue service (Ochsenbein et al. 2000). The SIMBAD database (operated at CDS, Strasbourg, France) has been queried with a query script containing all the stars from the Cepheid catalogue and the returned proper motion data from the Hipparcos and Tycho-2 catalogues have been merged with the Cepheid catalogue.

As the kinematic check requires that the Cepheids are ordered by error, an error value has been assigned to each Cepheid. The error value is based on the error ellipse of the proper motions as obtained from Hipparcos and Tycho-2 catalogues and is defined by multiplying the two axis of the error ellipses associated with the proper motion measurements. The data from the Tycho-2 catalogue is classified as having a lower quality than the data from the Hipparcos catalogue and to include this information in the ordering of the Cepheids,

(a) Product of the proper motion error ellipse axes
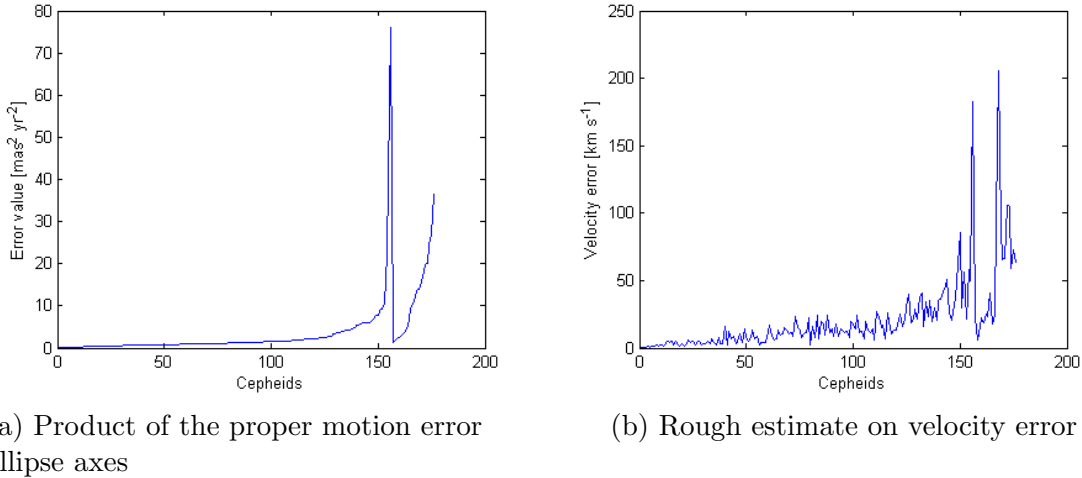
(b) Rough estimate on velocity error

Figure 4.1: Error values of the selected Cepheids.

the Cepheids are ordered by the combined index (quality, error value). This means that the Cepheids associated with proper motion data from the Hipparcos catalogue will appear first in the ordered list, internally ordered by error value, and when that list is exhausted the Cepheids associated with proper motion data from the Tycho-2 catalogue will follow. To visualize this the error values are displayed in Fig. 4.1a ordered by the algorithm just described.

While the kinematic check works as long as the Cepheids are sorted by the magnitude of the error, it is nice to have a rough estimate on how the measurement errors in the proper motions translates into errors in the velocities of the Cepheids. This estimate is shown in Fig. 4.1b. The estimate has been calculated by taking the square root of the error value, as it is the product of the error ellipse axes, and multiplied with the distance to the given Cepheid. The distances needed have been obtained by using the 1991-MF PLC relation, which is described below. Note that this estimate does not translate the error ellipse into Galactic coordinates and hence should not be used for anything other than getting an impression of the scale of the error.

The absorption coefficient, which is the ratio between the visual extinction and the $B-V$ colour excess, is near constant for most stars and is often set to 3.1, as mentioned in Eq. A.8 in Appendix A. Tammann et al. (2003a) have evaluated this absorption coefficient by examining studies of Cepheids in open clusters and associations and Cepheids with Baade-Becker-Wesselink distances and determined the coefficient to be

$$R = 3.17 \pm 0.13, \tag{4.1}$$

which is adopted in this study.

The stars from the database are shown in Fig. 4.2 which is a map of the stars as they appear when projected onto the Galactic plane (in Galactic coordinates) relative to the Sun. The Sun is located at coordinates (0,0) in the map and has a velocity vector pointing

along the y-axis in positive direction (straight up). The black line marks the direction to the Galactic centre. The database contains right ascension and declination information for the stars and the distance scale is found by using the 1991-MF PLC relation, which is described below.
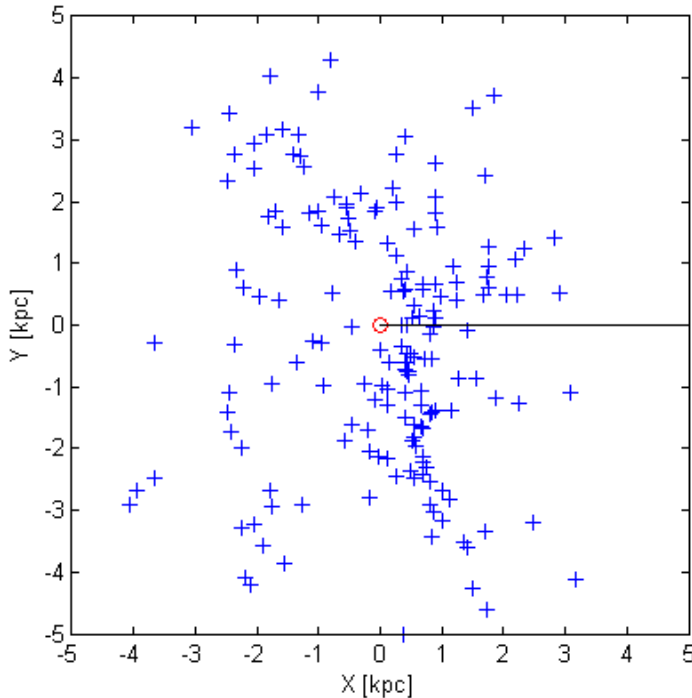
Figure 4.2: Galaxy map of the positions of the selected Cepheids projected onto the Galactic plane. The red circle in the centre marks the Sun's location and the Sun is moving in positive Y direction. The black line is the direction towards the Galactic centre.

4.2 Selected period-luminosity-colour relations

Recall the PLC relation from Eq. (2.7) in Sect. 2.3, repeated here:

$$\langle M_V \rangle = \alpha \log_{10} P + \beta_{BV}(\langle B_0 \rangle - \langle V_0 \rangle) + \gamma. \tag{4.2}$$

Some PLC relations use the $(V - I)_0$ colour index instead with the following format:

$$\langle M_V \rangle = \alpha \log_{10} P + \beta_{VI}(\langle V_0 \rangle - \langle I_0 \rangle) + \gamma. \tag{4.3}$$

As mentioned in Sect. 2.3 there are different determinations that not all agree on what the constants in the PLC relation should be. Perryman (2009) has provided an overview of

Table 4.1: Summary of selected period-luminosity-colour (PLC) relations. The subscript of the β constant governs the format of the PLC relation. The M_V magnitude is provided to give a quick comparison between the relations and it is calculated using a Cepheid from the dataset with a typical period, having the following data $\log_{10} P = 0.94$ days, $\langle B_0 \rangle - \langle V_0 \rangle = 0.62$ and $\langle V_0 \rangle - \langle I_0 \rangle = 0.70$.

PLC	α	β_{BV}	β_{VI}	γ	Typical $\langle M_V \rangle$
					[mag]
1991-MF	-2.76	0	0	-1.40	-3.99
1997-FC	-2.81	0	0	-1.43	-4.07
1997-FW	-3.80	2.70	0	-2.38	-4.28
1998-Lu	-2.12	0	0	-1.73	-3.72
1999-La	-2.77	0	0	-1.44	-4.04
2001-Fr	-3.255	0	2.45	-2.644	-3.99
2003-Ta	-3.141	0	0	-0.826	-3.78
2007-vL	-3.29	0	2.45	-2.576	-3.95
2013-An	-3.08	0	0	-0.94	-3.84
2013-Ma	-3.245	0	2.45	-2.545	-3.88

PLC relations released up to 2009, which provides a base set of PLC that can be examined with the kinematic check described in Ch. 3. I have further found and added two PLC relations, which are more recent.

The selected PLC relations for this study are listed in chronological order by publication date in the following. For all relations the period P is in the unit of days. A summary of all selected PLC relations is found in Table 4.1. Please take note of the magnitudes calculated for a typical star in the $\langle M_V \rangle$ column. The values varies over a range of 0.56 mag. By definition of magnitudes this translates to a distance factor of ca. 1.3. I.e. the distance to the Cepheid is 30% greater using 1997-FW compared to using 2003-Ta.

PLC 1991-MF

Madore & Freedman (1991) uses what was at the time new measurements in the near infrared to evaluate reddening and generate PLC relations for more than just the visual colour band. The relevant PLC relation is:

$$\langle M_V \rangle = -2.76(\pm 0.11)(\log_{10} P - 1.00) - 4.16(\pm 0.05). \tag{4.4}$$

PLC 1997-FC

Feast & Catchpole (1997) adopt a slope from a previous study of Cepheids in the Large Magellanic Cloud, but estimate a zero crossing by using Hipparcos parallax data. They

find the PLC relation to be

$$\langle M_V \rangle = -2.81 \log_{10} P - 1.43(\pm 0.10). \tag{4.5}$$

PLC 1997-FW

Feast & Whitelock (1997) adopt the α and the β_{BV} from the PLC relation determined by Feast & Walker (1987) and recalculate the γ using Hipparcos parallax data. They find

$$\langle M_V \rangle = -3.80 \log_{10} P + 2.70(\langle B_0 \rangle - \langle V_0 \rangle) - 2.38(\pm 0.10). \tag{4.6}$$

PLC 1998-Lu

Luri et al. (1998) use a maximum likelihood approach to estimate the PLC relation using 219 Cepheids from the Hipparcos catalogue and find

$$\langle M_V \rangle = -2.12(\pm 0.08) \log_{10} P - 1.73(\pm 0.20). \tag{4.7}$$

PLC 1999-La

Lanoix et al. (1999) build on the previous result by Feast & Catchpole (1997), but argue that the result can be improved by using Monte Carlo simulations and find

$$\langle M_V \rangle = -2.77 \log_{10} P - 1.44(\pm 0.05). \tag{4.8}$$

PLC 2001-Fr

Freedman et al. (2001) calibrate a PLC relation based on a multitude of secondary distance determination methods and averaging them out. The PLC relation is in its original form expressed using the Wesenheit function

$$\langle M_W \rangle = -3.255(\pm 0.01)(\log_{10} P - 1) - 5.899(\pm 0.01), \tag{4.9}$$

where $W = V - 2.45(V - I)$. Translated into an absolute visual magnitude this gives

$$\langle M_V \rangle = -3.255(\pm 0.01)(\log_{10} P - 1) + 2.45(\langle V_0 \rangle - \langle I_0 \rangle) - 5.899(\pm 0.01). \tag{4.10}$$

PLC 2005-Ta

Tammann et al. (2003b) put together a database (which is the database used for this project) to estimate a PLC relation and they find

$$\langle M_V \rangle = -3.141(\pm 0.100) \log_{10} P - 0.826(\pm 0.119). \tag{4.11}$$

PLC 2007-vL

van Leeuwen et al. (2007) revise Hipparcos parallaxes and as part of that process they update the PLC relation from the work of Freedman et al. (2001) and find

$$\langle M_V \rangle = -3.29 \log_{10} P + 2.45(\langle V_0 \rangle - \langle I_0 \rangle) - 2.576(\pm 0.030). \tag{4.12}$$

PLC 2013-An

Anderson et al. (2013) use Galactic Cepheids in open clusters and use alternative distance determinations to the clusters to calibrate the PLC relation and find

$$\langle M_V \rangle = -3.08(\pm 0.50) \log_{10} P - 0.94(\pm 0.42). \tag{4.13}$$

PLC 2013-Ma

Mager et al. (2013) team up with Madore and Freedman to determine metallicity dependencies of PLC relations. For that study they use an analysis on Cepheids in the Large Magellanic Cloud (LMC) done by Soszynski et al. (2008) and couple it with a selected distance modulus to the LMC and find

$$\langle M_V \rangle = -3.245(\pm 0.035)(\log_{10} P - 1) + 2.45(\langle V_0 \rangle - \langle I_0 \rangle) - 5.790(\pm 0.042). \tag{4.14}$$

4.3 Selected solar velocity

In his overview of results obtained by using Hipparcos data Perryman (2009) has provided an extensive list over different solar velocities with respect to the Galactic plane. As only the vertical component of the Sun's velocity vector ($w_\odot$) is of interest for this study the following discussion will be limited to this part of the velocity vector.

Perryman recommends using $w_\odot = 7.17 \pm 0.38$ km s^{-1}, which is a result from Dehnen & Binney (1998), because that study used a kinematically-unbiased sample and further is in good agreement with work done before the availability of data from the Hipparcos satellite. Binney has on several occasions revisited the result from 1998, and the latest estimate is found to be $w_\odot = 7.25^{+0.37}_{-0.36}$ km s^{-1} (Schönrich et al. 2010). Subsequently a study that builds on the measurements done by the Radial Velocity Experiment (RAVE) survey finds $w_\odot = 6.57 \pm 0.21$ km s^{-1} (Coşkunoğlu et al. 2011). A recent study on young objects where the differential rotation of the galaxy and the effects of the spiral arm density wave is taken into account finds $w_\odot = 6.5 \pm 0.3$ km s^{-1} (Bobylev & Bajkova 2014). However, part of that study depends on Cepheids, and if they are excluded from the data used, to get an estimate independent of Cepheids, $w_\odot = 7.06 \pm 0.65$ km s^{-1} is found by examining the data in their Table 1 or 2.

This study will thus focus on examining the PLC relations with respect to a selection of solar velocities in the range between 6.5 km s^{-1} and 7.2 km s^{-1}.

Chapter 5

Results

The kinematic check has been implemented as a function in the software package MATLAB using the Cepheid database described in Sect. 4.1. The implementation can be found in Appendix C. An example script illustrating how the function has been used is shown in Appendix D.

First the results from analysing the median velocity of the Cepheids is shown and it discussed how this relates to the choice of solar velocity. This is followed by the results obtained when the selected Cepheid sample is forced to a given size in order to examine how the period-luminosity-colour (PLC) relations relate relatively to each other.
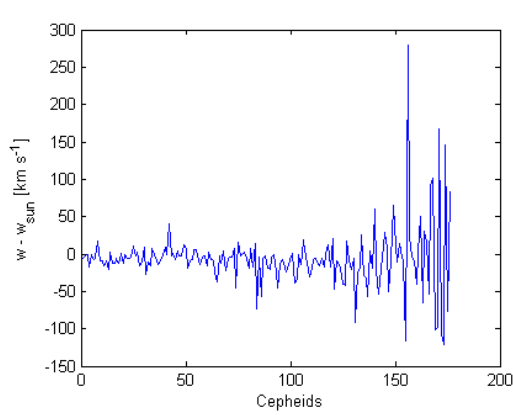
5.1 Velocity distribution

The Cepheid velocities have been calculated using the various PLC relations and Table 5.1 shows the found medians with confidence interval. The preferred sample in the table is the sample size where the confidence interval in terms of velocities is smallest. The length has been divided by two to give a better impression of the distance to each side of the median, even though the median may not be entirely in the middle of the interval. To show the result independently of the choice of solar velocity the median of $w' = w - w_\odot$ is used instead of the median of w discussion in Sect. 3.3. The median of w' should then be $-w_\odot$ instead of zero if the thesis assumption is true and the used PLC relation is correct. Using Eq. (3.2) and the definitions of magnitudes in Appendix A the required γ correction of the used PLC relation is thus expressed as

$$\gamma_{\mathrm{corr}} = 5 \cdot \log_{10}\left(\frac{w'}{-w_\odot}\right).$$
(5.1)

All the plots in Fig. 5.1 are based on the 1991-MF PLC relation. In Fig. 5.1a the velocity distribution obtained using the PLC relation is shown. Notice that the Cepheids are ordered by the error value along the x-axis. In Fig. 5.1b the cumulated density function of the velocity distribution of the preferred sample of Cepheids have been plotted such that if the velocities have a normal distribution the cumulated density function should follow a

Table 5.1: Result of kinematic check in terms of $w - w_\odot$.

PLC	Pref. sample	Median$(w - w_\odot)$ $[\mathrm{km\,s^{-1}}]$	$I_{\mathrm{conf}}(w - w_\odot)$ $[\mathrm{km\,s^{-1}}]$	Length/2 $[\mathrm{km\,s^{-1}}]$
1991-MF	116	-6.29	$[-7.02, -5.53]$	0.75
1997-FC	116	-6.57	$[-7.22, -5.73]$	0.75
1997-FW	88	-4.59	$[-5.32, -3.94]$	0.69
1998-Lu	85	-4.23	$[-4.80, -3.26]$	0.77
1999-La	116	-6.44	$[-7.17, -5.65]$	0.76
2001-Fr	88	-4.56	$[-5.37, -3.59]$	0.89
2003-Ta	116	-5.88	$[-6.26, -4.97]$	0.65
2007-vL	88	-4.48	$[-5.30, -3.54]$	0.88
2013-An	116	-5.97	$[-6.36, -5.11]$	0.63
2013-Ma	87	-4.38	$[-5.22, -3.53]$	0.84

(a) Velocity distribution
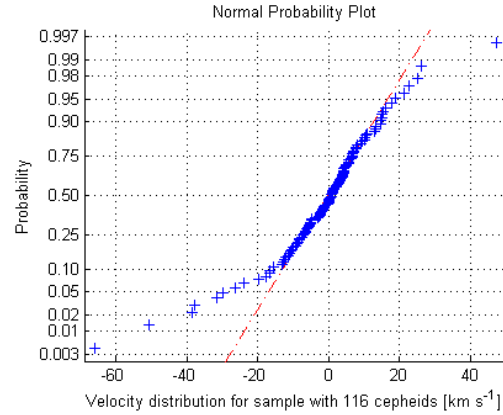
(b) Normal plot of preferred sample
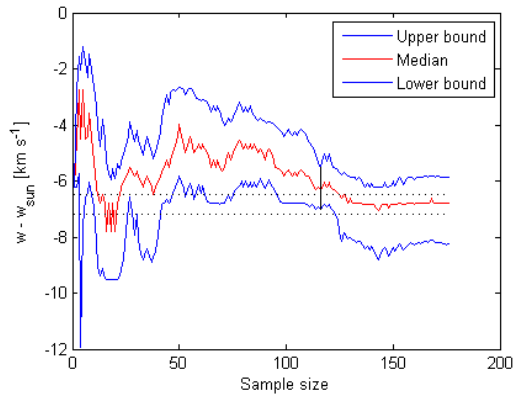
(c) Confidence interval in terms of velocities
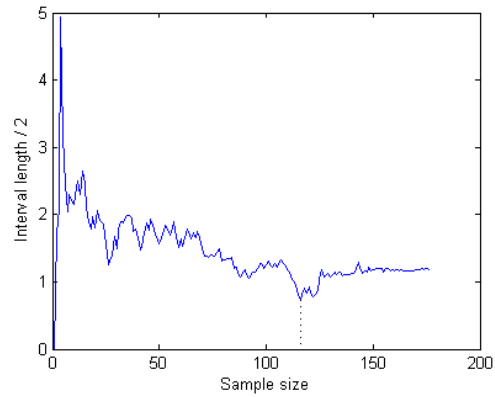
(d) Half length of confidence interval

Figure 5.1: Velocity distributions and confidence intervals for 1991-MF. The plots for the other PLC relations can be found in Appendix E.

Figure 5.2: $w - w_\odot$ distributions using preferred sample of each PLC relation. The dotted lines marks the range of possible solar velocities adopted in this study.

straight line. Fig. 5.1c shows how the confidence intervals in terms of velocity develop as Cepheids are added to the sample. The dotted straight lines in Fig. 5.1c indicate the range of values the median should take to agree with the thesis assumption based on the selection of solar velocity. Fig. 5.1d shows the half length of the intervals to better illustrate the minimum.

The other PLC relations produces plots that are very similar and are thus omitted here. The interested reader can find these plots in Appendix E.

The result from Table 5.1 is plotted in Fig. 5.2. The selection range of the solar velocity is marked to show how the PLC relations agree with this range.

5.2 Forced sample size

It is clear from Table 5.1 that there are two preferred samples that are found by several PLC relations. To examine how the PLC relations perform relatively to each other it has been examined what the results are, if the sample selection is forced. The forced samples used are 88 and 116 Cepheids, which are plotted respectively in in Fig. 5.3 and in Fig. 5.4.

23
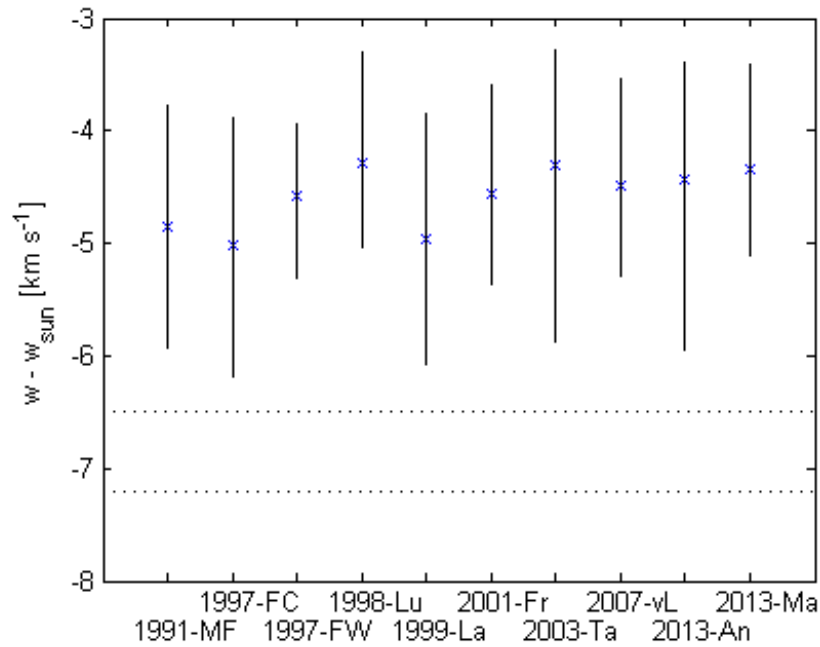
Figure 5.3: $w - w_\odot$ distributions using sample with 88 Cepheids. The dotted lines marks the range of possible solar velocities adopted in this study.

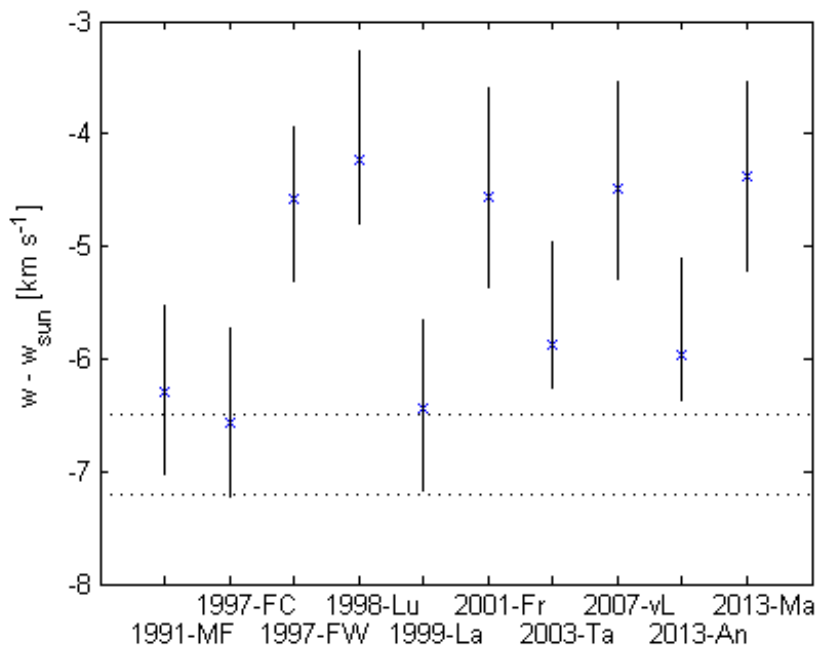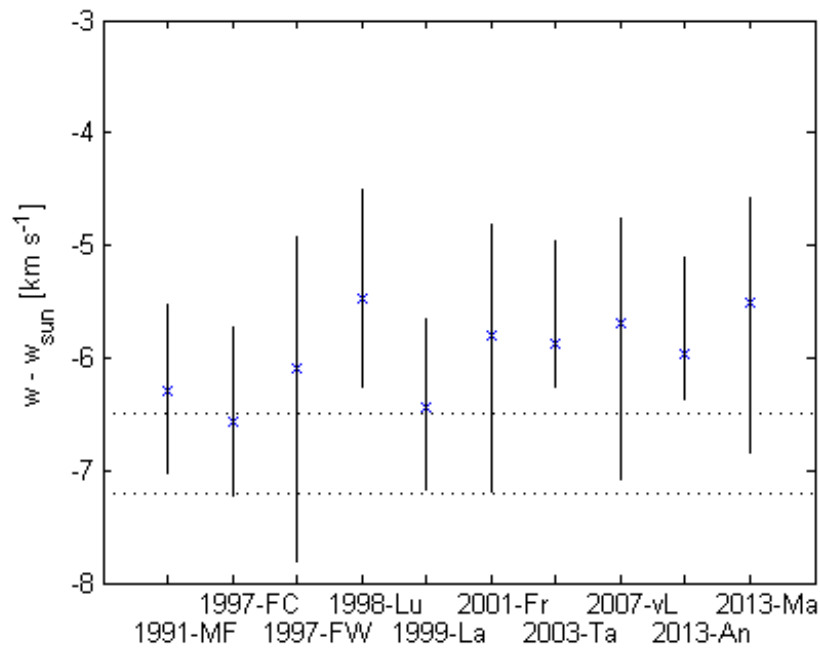
Figure 5.4: $w - w_\odot$ distributions using sample with 116 Cepheids. The dotted lines marks the range of possible solar velocities adopted in this study.

(a) $R = 2.8$ (b) $R = 3.5$
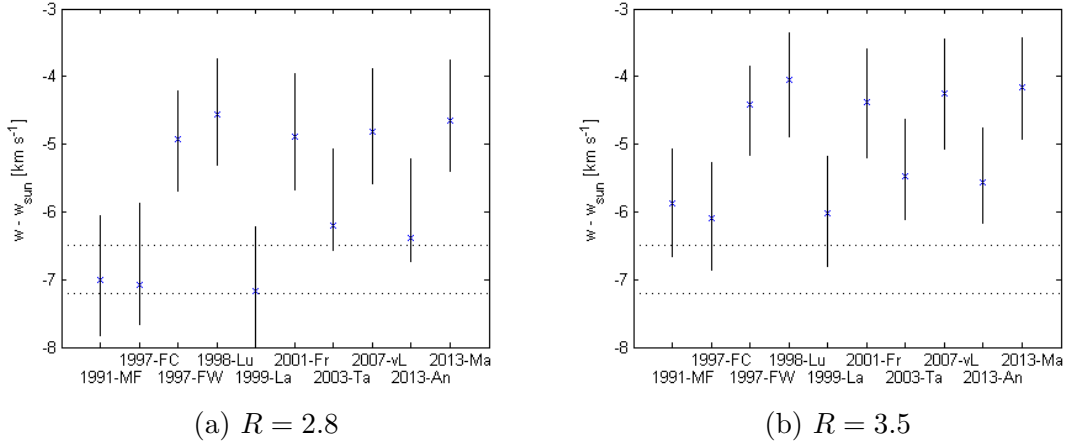
Figure 5.5: $w - w_\odot$ distributions using different R value. The dotted lines marks the range of possible solar velocities adopted in this study.

5.3 Choice of R

Examining the effects of the choice of the absorption coefficient R, which is the ratio between the visual extinction and the $B - V$ colour excess as explained in Sect. 4.1, yields the results that can be seen in Fig. 5.5. The PLC relations that preferred the sample with 116 Cepheids with $R = 3.17$ still prefers the sample with 116 Cepheids independent of the choice of R. The PLC relations that previously preferred the sample with 88 Cepheids now seem to vary a bit between samples with 90 Cepheids when choosing $R = 3.5$ and sample with 92 Cepheids when choosing $R = 2.7$. Thus the PLC relations still divides into two distinct groups.

Chapter 6

Discussion

In this final chapter the results are discussed followed by the conclusions I draw from these results.

6.1 Discussion

There are a few trends in the results that are worth discussing. Firstly it is a good thing that the velocity distributions indeed seem to be ordered by error, as the large velocities confined at the end of the plot in Fig. 5.1a are clearly linked to the large errors estimated in Fig. 4.1b. It is also worth noticing that the preferred samples, apart from the tails, appear to have normal distributed velocities, as shown in Fig. 5.1b.

From Fig. 5.1c it may seem like the median velocity is not very stable. This is not an issue, as the median velocity changes in the order of a few $\mathrm{km\,s^{-1}}$, while the velocities in Fig. 5.1a vary in the tens of $\mathrm{km\,s^{-1}}$ to hundreds of $\mathrm{km\,s^{-1}}$. It is interesting to see that local minima appear in Fig. 5.1d, which suggests that the amount of dilution added to the sample does not grow consistently with each Cepheid added to the sample. However, all but one period-luminosity-colour (PLC) relation have a clear global minimum. That PLC relation is the 1998-Lu PLC relation, which can be seen in Appendix E in Fig. E.7. This means that the choice of preferred sample for this relation is somewhat arbitrary and thus not to be trusted.

For the PLC relations without a colour term there is a very clear preference for the sample with 116 Cepheids, as seen in Table 5.1. The only exception is for 1998-Lu, but as mentioned above this relation does not have a clear preferred sample. Like the PLC relations without a colour term, the PLC relations with a colour term agree on a common sample size, namely the sample with 88 Cepheids.

Forcing the PLC relations without a colour term to use the preferred sample of the PLC relations with a colour term and vice versa shows that the corrections found are quite similar across the board when using the same sample, as illustrated by Fig. 5.3 and Fig. 5.4. It thus becomes clear that the biggest impact on the result lies in the choice of sample.

Looking at the results in Sect. 5.3 it is also clear that the PLC relation's choice of

preferred sample seems to be quite insensitive to the choice of the absorption coefficient, R.

It is not obvious if the sample with 116 Cepheids should be preferred over the sample with 88 Cepheids. Looking at the Fig. 5.1c the mean of the sample reaches a general maximum in the sample size 50 to 100 region after which it declines. That sample with 88 Cepheids is thus at an extreme, which could be considered to just be a chance result. However, the errors does not really start to grow big until sample size 130 to 150 is reached as illustrated in Fig. 4.1b, so arguably the sample with 116 Cepheids could be preferred over the sample with 88 Cepheids simply because it is larger.

Assuming that the sample with 116 Cepheids is somewhat more trustworthy, Fig. 5.4 shows that the median of $w - w_\odot$ is relatively close to -6.5 km s^{-1}. Thus with a solar velocity of 6.5 km s^{-1} and the sample with 116 Cepheids little to no correction is needed to the γ constants in the PLC relations in order for the result to agree with the thesis assumption. This is shown in Fig. 6.1a. Insisting on a solar velocity of 7.0 km s^{-1} (using 7.0 km s^{-1} since it is a reliable results from kinematics studies using Hipparcos data), the result would instead be a correction on about -0.2 mag to the γ constant in the PLC relations. This is shown in Fig. 6.1b.

However, going to the other end of the spectrum with a choice of solar velocity between 6.5 km s^{-1} and 7.0 km s^{-1} and using the sample with 88 Cepheids the corrections are more in the region of -0.7 mag to -1.0 mag as shown in Fig. 6.1c and Fig. 6.1d.

A negative correction means that the stars are actually brighter than previously believed, meaning that they must be further away than initially anticipated. For example, a correction of -0.2 mag means that the distances should be increased by about 10%. This change of distance scaling will naturally have consequences for our cosmology as the Hubble constant would decrease and the age of the universe would be greater than previously believed. But fortunately there are other ways to determine the Hubble constant, such as using distance determinations to supernovae or galaxy surface brightness functions (Freedman et al. 2001), so it is possible to argue whether the suggested corrections are acceptable.

One such possibility is to compare how well these results affect the distance to the Large Magellanic Cloud (LMC). The distance to the LMC is usually expressed in distance modulus and given the definition of distance modulus a negative γ correction will result in a positive adjustment of the distance modulus. There are several good measurements of the distance to the LMC independent of Cepheids, and although there has been discussion about publication bias, the consensus is a distance modulus of 18.49(± 0.09) mag (de Grijs et al. 2014). Note that de Grijs et al. (2014) have included distances determined with the use of Cepheids in this distance modulus. However, their Fig. 1 to 3 shows that modern distance determinations using RR Lyrae and CMD (colour-magnitude-diagram) gives values for the distance modulus between 18.40 and 18.50 mag, and distance determination methods using eclipsing binaries gives a distance modulus between 18.50 and 18.55 mag. So the exclusion of cepheids will not change the estimate significantly.

Luri et al. (1998) has examined the distance to the LMC with both his own 1999-Lu PLC relation, but also with the 1997-FC PLC relation. The distances found are respectively

(a) Sample 116, $w_\odot = 6.5$ km s^{-1}

(b) Sample 116, $w_\odot = 7.0$ km s^{-1}

(c) Sample 88, $w_\odot = 6.5$ km s^{-1}

(d) Sample 88, $w_\odot = 7.0$ km s^{-1}
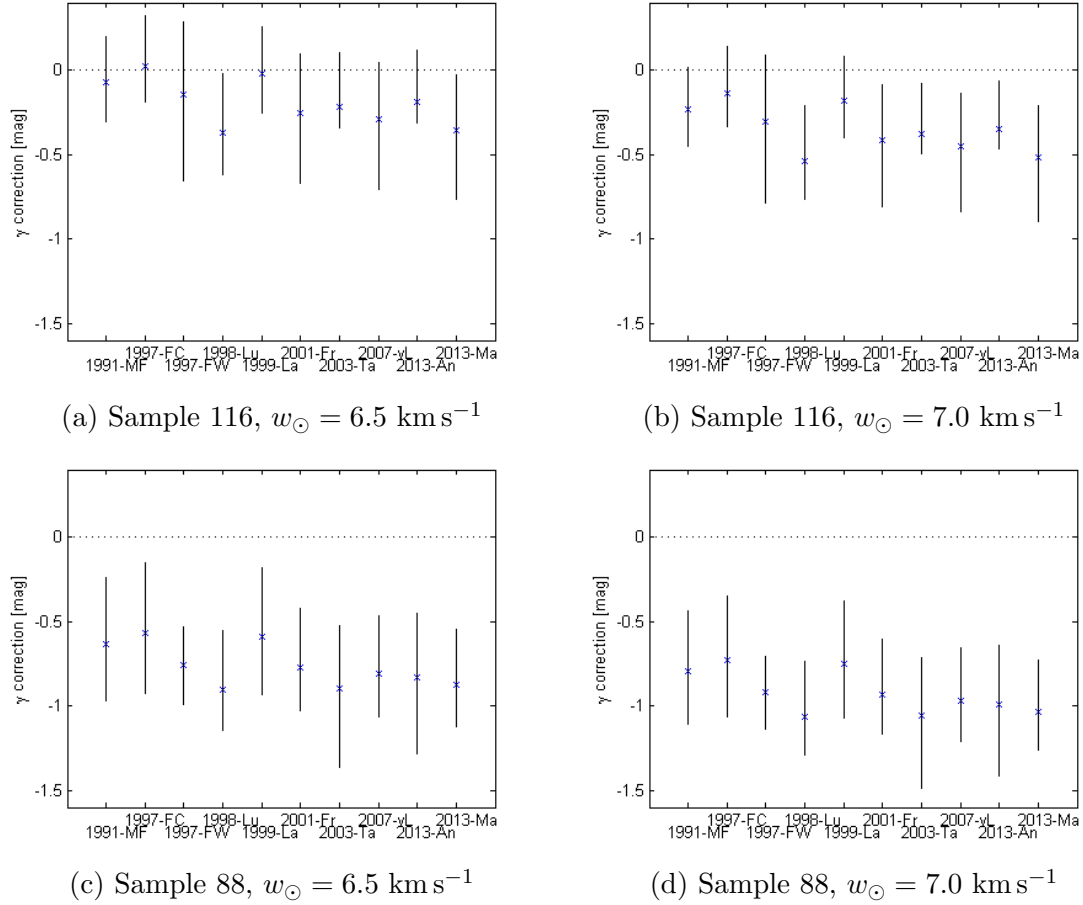
Figure 6.1: γ corrections as consequence of choice of forced sample and selected solar velocity.

18.32($\pm$0.17) mag and 18.21($\pm$0.20). These distances are slightly lower than the consensus by 0.2 to 0.3 mag, which coincide quite interestingly with the γ correction from using sample 116 and a solar velocity of 7.0 km s^{-1}. However, the correction from using the sample with 88 Cepheids is in the order of 0.7 to 1.0 mag depending on the choice of solar velocity, which give a distance modulus to the LMC of 19 mag or more. This correction is in strong disagreement with the consensus, so accepting this correction is not very reasonable.

There is of course the possibility that the thesis assumption is not correct, for instance if the Cepheids are not distributed evenly in the galaxy and thus could be affected by some local velocity irregularity, for example caused by a warp in the Galactic disk. Galaxy maps for each of the two discussed samples are shown in Fig. 6.2. There seems to be a slight concentration of Cepheids along a line going from north-north-west to south-south-east, but in general the Cepheids are fairly well spread out, and rather similar in the samples. Thus the difference between the results for the two samples cannot simply be explained by their different positions, but it is not possible to rule out a bias caused by other effects.

(a) Sample with 88 Cepheids   (b) Sample with 116 Cepheids
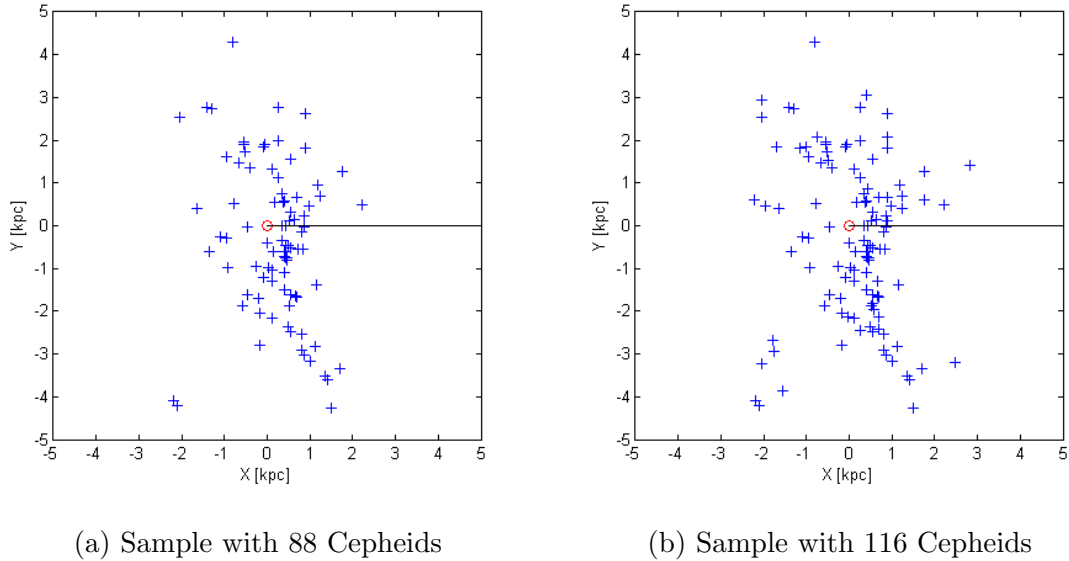
Figure 6.2: Galaxy maps filtered by sample selection.

6.2 Conclusion

The conclusions I draw from this study are that

- the choice of sample has a significant impact on the result,

- the sample with 116 Cepheids is probably more reliable due to simply being larger than the sample with 88 Cepheids; also the 88 sample gives an unreasonably large correction,

- by accepting the 116 sample, a discrepancy is found which can be explained by selecting a solar velocity of 6.5 $\mathrm{km\,s^{-1}}$, even though this is in disagreement with Hipparcos $w_\odot = 7.0$ $\mathrm{km\,s^{-1}}$ result and

- by accepting the 116 sample and requiring $w_\odot = 7.0$ $\mathrm{km\,s^{-1}}$, a -0.2 correction to the γ constants in the PLC relations is found (corresponding to 10% larger distances), which in turn is in good agreement with LMC distance modulus studies.

It is interesting to see that the result of this study can suggest a correction to the PLC relations such that they get a better agreement with the conventional distance modulus to the LMC, however it is clear that further investigation should be conducted into why the result is so sensitive to the selection of the sample.

With the launch of the Gaia satellite it is quite likely that the database of Cepheids can be increased in size, as Gaia is estimated to observe around 9000 Cepheids (Windmark et al. 2011). A larger database with smaller errors in the proper motion observations would undoubtedly increase the quality of the kinematic check.

Bibliography

Anderson, R. I., Mowlavi, N., & Eyer, L. 2013, in IAU Symposium, Vol. 289, IAU Symposium, ed. R. de Grijs, 153–156

Berdnikov, L. N., Dambis, A. K., & Vozyakova, O. V. 2000, A&AS, 143, 211

Binney, J. & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

Blaauw, A., Gum, C. S., Pawsey, J. L., & Westerhout, G. 1960, MNRAS, 121, 123

Bobylev, V. V. & Bajkova, A. T. 2014, MNRAS, 441, 142

Coşkunoğlu, B., Ak, S., Bilir, S., et al. 2011, MNRAS, 412, 1237

de Grijs, R., Wicker, J. E., & Bono, G. 2014, AJ, 147, 122

Dehnen, W. & Binney, J. J. 1998, MNRAS, 298, 387

Feast, M. & Whitelock, P. 1997, MNRAS, 291, 683

Feast, M. W. & Catchpole, R. M. 1997, MNRAS, 286, L1

Feast, M. W. & Walker, A. R. 1987, ARA&A, 25, 345

Fernie, J. D., Evans, N. R., Beattie, B., & Seager, S. 1995, Information Bulletin on Variable Stars, 4148, 1

Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47

Hubble, E. P. 1925, The Observatory, 48, 139

Jenkins, A. 2013, Phys. Rep., 525, 167

Joshi, Y. C. & Joshi, S. 2014, New A, 28, 27

Karttunen, H., Krüger, P., Oja, H., Poutanen, M., & Donner, K. J., eds. 2007, Fundamental Astronomy

Lanoix, P., Paturel, G., & Garnier, R. 1999, MNRAS, 308, 969

Leavitt, H. S. & Pickering, E. C. 1912, Harvard College Observatory Circular, 173, 1

Luri, X., Gomez, A. E., Torra, J., Figueras, F., & Mennessier, M. O. 1998, A&A, 335, L81

Madore, B. F. & Freedman, W. L. 1991, PASP, 103, 933

Mager, V. A., Madore, B. F., & Freedman, W. L. 2013, ApJ, 777, 79

Marconi, M. 2009, Mem. Soc. Astron. Italiana, 80, 141

Martin, W. L., Warren, P. R., & Feast, M. W. 1979, MNRAS, 188, 139

Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23

Perryman, M. 2009, Astronomical Applications of Astrometry: Ten Years of Exploitation of the Hipparcos Satellite Data (Cambridge University Press)

Ritter, A. 1879, Annalen der Physik und Chemie, Neue Folge, 8, 157

Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829

Soszynski, I., Poleski, R., Udalski, A., et al. 2008, Acta Astron., 58, 163

Tammann, G. A., Sandage, A., & Reindl, B. 2003a, VizieR Online Data Catalog, 340, 40423

Tammann, G. A., Sandage, A., & Reindl, B. 2003b, A&A, 404, 423

van Altena, W. F. 2013, Astrometry for Astrophysics

van Leeuwen, F., Feast, M. W., Whitelock, P. A., & Laney, C. D. 2007, MNRAS, 379, 723

Windmark, F., Lindegren, L., & Hobbs, D. 2011, A&A, 530, A76

Zhevakin, S. A. 1963, ARA&A, 1, 367

Appendix A

Basic Astrophysics

This appendix gives a brief introduction to the astrophysics used in this thesis with the aim to make the content available to a reader without a background in astrophysics.

Luminosity

In astrophysics the amount of energy radiated out of a star per unit of time is called the luminosity of the star. In essence it is a measure of how bright the star is. Under normal circumstances it is not possible to measure the luminosity of a star by observing it, instead it is possible to measure the flux of energy at the observers position. Assuming that the star is radiating the energy uniformly in all directions and no light is absorbed between the star and the observer, the flux at distance d then becomes

$$F = \frac{L}{4\pi d^2},$$ (A.1)

where F is the flux and L is the luminosity. If there is a loss of light between the star and the observer the measured flux will be reduced, this will be covered in the section below under the term reddening.

Magnitude

The difference in brightness between two stars can be described as the difference in magnitude of the two stars and is defined by

$$M_1 - M_2 = -2.5 \cdot \log_{10} \left(\frac{F_1}{F_2} \right),$$ (A.2)

where $M_1 - M_2$ is the difference in magnitude of two stars and F_i is the flux of star i. It is possible to talk about the magnitude of a star, which is then by convention the magnitude difference between the star and the star named Vega. Magnitude is physically unitless,
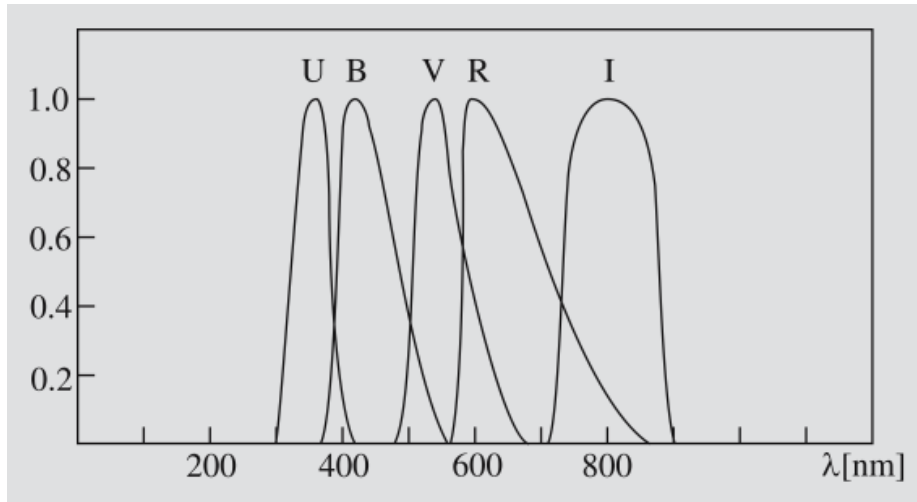
Figure A.1: The UBVRI filters showing the fraction of radiation flux being measured at a given wavelength, 1.0 being 100%. (Karttunen et al. 2007, Fig. 4.6)

but is often denoted with the unit "mag". The -2.5 factor is purely a historical remnant to make the magnitude definition somewhat compatible with ancient ways of classifying the brightness of stars, where a class 1 star is brighter than a class 2 star.

Two types of magnitudes are used regularly: Apparent and absolute magnitude. Apparent magnitude means that the fluxes used in the relation are the fluxes measured at the Earth; this then becomes a measure of how bright a star is on the sky (compared to Vega). Absolute magnitude means that the fluxes used are the fluxes as they would have been measured at a fixed distance of 10 pc from the star with no loss of light between the star and the observer; thus the absolute magnitude is directly linked to the luminosity of the star.

Colour

Measuring the energy flux for all types of radiation can be quite difficult, because the atmosphere of the Earth blocks many wavelengths of non-visible light. But it is possible to use a colour filter and only measure the flux for a certain wavelength range, for instance the flux of blue light. In order to compare the flux measurements between different observers the set of colour filters have been standardized. The most common standard of colour filters used is the UBVRI photometric system (Ultraviolet, Blue, Visual, Red, Infrared), which is illustrated in Fig. A.1.

Given that fluxes are often measured through a colour filter, it is possible to talk about magnitudes for a specific colour. For instance the apparent visual magnitude of a star is then the magnitude difference between the star and Vega with fluxes measured at the Earth using a visual light filter (the V peak with centre at 550 nm in Fig. A.1). A magnitude value that is based on fluxes without any filters is called a bolometric magnitude.

Colour indices

Subtracting one colour magnitude from another gives a colour index, for instance $B - V$ is the colour index where B is the apparent blue magnitude and V is the apparent visual magnitude. A colour index is used to give an indication of the colour of the star, i.e. how much blue light there is compared to the visual spectrum. The spectral class categorization system (OBAFGKM—O is blue and M is red) is based on the colours of the stars, and thus there is a direct link between $B - V$ colour index and spectral class. The Sun has colour index $B - V = 0.656$ and belongs to spectral class G.

Reddening

When light travels from a star to the Earth the light may encounter "dust" on the way, which may scatter the light away from the line of sight to the observer. This process has varying efficiency which depends on the wavelength of the light and is more efficient for blue light than red light, so that more blue light than red light is "lost"; hence this extinction phenomena is often called reddening.

The amount of extinction that happens as the light travels through this "interstellar medium" depends on the so-called optical thickness of the medium, which is denoted with a τ. The influence that optical thickness have on the flux can be described with

$$F_{\text{reddened}}(d) = F(d) \cdot e^{-\tau}. \tag{A.3}$$

Since the absorption is dependent on the wavelength of the light each colour filter gives different values for τ.

Apparent magnitude is based on the reddened flux, while the absolute magnitude is based on the intrinsic flux, thus the relation between apparent (V) and absolute (M_V) magnitude can be described using (A.2) as

$$V - M_V = -2.5 \cdot \log_{10} \left(\frac{F_{\text{reddened}}(d)}{F(10 \text{ pc})} \right) = -2.5 \cdot \log_{10} \left(\frac{L \cdot e^{-\tau}}{4\pi \cdot d^2} \cdot \frac{4\pi \cdot (10 \text{ pc})^2}{L} \right)$$

$$= 5 \cdot \log_{10} \left(\frac{d}{10 \text{ pc}} \right) + 2.5\tau \log_{10}(e) = 5 \cdot \log_{10} \left(\frac{d}{10 \text{ pc}} \right) + A_V, \tag{A.4}$$

where $A_V = 2.5\tau \log_{10}(e)$ is called the visual extinction (and similarly the blue extinction for A_B etc.). This can be rewritten as

$$V = V_0 + A_V, \quad V_0 = M_V + 5 \cdot \log_{10} \left(\frac{d}{10 \text{ pc}} \right), \tag{A.5}$$

where V_0 is called the intrinsic apparent visual magnitude, as it is the apparent visual magnitude without any reddening. An interesting consequence of this is that

$$(B - V)_0 = B_0 - V_0 = M_B - M_V. \tag{A.6}$$

Analysing a colour index gives

$$B - V = (M_B + 5 \cdot \log_{10}\left(\frac{d}{10\text{ pc}}\right) + A_B) - (M_V + 5 \cdot \log_{10}\left(\frac{d}{10\text{ pc}}\right) + A_V)$$
$$= (M_B - M_V) + (A_B - A_V) = (B-V)_0 + E_{B-V}, \tag{A.7}$$

where $E_{B-V} = (B-V) - (B-V)_0$ is called the colour excess.

It turns out that the ratio (R) between the visual extinction and the $B-V$ colour excess is near constant for most stars, i.e.

$$R = \frac{A_V}{A_B - A_V} = \frac{A_V}{E_{B-V}} \approx 3.1. \tag{A.8}$$

Distance modulus

The logarithmic expression shown in Eq. A.5 is per definition called the distance modulus (μ), and is often used as an expression of distance with the unit of mag, i.e.

$$\mu_V = V_0 - M_V = 5 \cdot \log_{10}\left(\frac{d}{10\text{ pc}}\right). \tag{A.9}$$

The Wesenheit function

It is possible to define a magnitude that is free of reddening. The Wesenheit function defines a W magnitude as

$$W = V - R_{VI} \cdot (V - I), \tag{A.10}$$

where V is apparent visual magnitude, I is apparent infrared magnitude and R_{VI} is defined, using (A.8) from above, as

$$R_{VI} = \frac{A_V}{E_{V-I}} = \frac{R \cdot E_{B-V}}{E_{V-I}} \approx \frac{3.1 \cdot E_{B-V}}{E_{V-I}}. \tag{A.11}$$

Using $V = V_0 + A_V$ it is quickly proven that $A_W = 0$, i.e. the W magnitude is free of reddening:

$$W = (V_0 + A_V) - R_{VI} \cdot ((V_0 + A_V) - (I_0 + A_I))$$
$$= V_0 + A_V - R_{VI} \cdot (V_0 - I_0 + E_{V-I}) = V_0 + A_V - R_{VI} \cdot (V_0 - I_0) - A_V = W_0.$$

Given that the Wesenheit magnitude is free of reddening the distance modulus in terms of Wesenheit magnitudes are often called the true distance modulus (μ_0):

$$\mu_W = W - M_W = W_0 - M_W = \mu_0. \tag{A.12}$$

Appendix B

Cepheid database

Table B.1: Cepheid database sorted by quality and error value. The error value is the product of the two axis of the error ellipse of the proper motion data. Quality 1 means proper motion data from Hipparcos catalogue, while quality 2 means proper motion data from Tycho-2 catalogue.

Star	$\log_{10} P$	V	E_{B-V}	$(B-V)_0$	$(V-I)_0$	μ_b	ErrVal	Qlty
		[mag]	[mag]	[mag]	[mag]	[mas yr^{-1}]	[mas^2 yr^{-2}]	
l car	1.551	3.737	0.160	1.103	0.968	-2.167	0.070	1
x sgr	0.846	4.562	0.201	0.549	0.654	-1.710	0.074	1
y sgr	0.761	5.743	0.188	0.667	0.722	-0.536	0.126	1
t vul	0.647	5.753	0.067	0.577	0.592	-6.686	0.134	1
bet dor	0.993	3.754	0.069	0.734	0.724	-0.003	0.167	1
t cru	0.828	6.564	0.178	0.752	0.728	-1.652	0.168	1
s tra	0.801	6.391	0.082	0.664	0.697	-1.046	0.182	1
x cyg	1.215	6.392	0.261	0.875	0.821	$+3.276$	0.193	1
r tra	0.530	6.660	0.134	0.582	0.641	-3.035	0.224	1
r mus	0.876	6.319	0.134	0.639	0.650	-2.136	0.231	1
s sge	0.923	5.610	0.112	0.694	0.690	-4.754	0.235	1
r cru	0.766	6.765	0.150	0.638	0.669	-2.201	0.265	1
sv vul	1.653	7.209	0.518	0.944	0.853	-2.068	0.312	1
t mon	1.432	6.125	0.195	0.972	0.895	$+0.303$	0.314	1
rs pup	1.617	7.034	0.453	0.975	0.962	-1.376	0.319	1
bg vel	0.840	7.643	0.439	0.741	0.751	-2.589	0.326	1
u car	1.589	6.282	0.287	0.896	0.862	-0.718	0.336	1
s cru	0.671	6.597	0.162	0.602	0.638	-3.383	0.360	1
av cir	0.487	7.402	0.368	0.000	0.580	$+0.367$	0.360	1
ap pup	0.706	7.381	0.241	0.622	0.620	-2.500	0.370	1
ax cir	0.722	5.889	0.262	0.471	0.557	-0.107	0.398	1
er car	0.888	6.819	0.099	0.771	0.744	-0.988	0.410	1

Star	$\log_{10} P$	V [mag]	E_{B-V} [mag]	$(B-V)_0$ [mag]	$(V-I)_0$ [mag]	μ_b [mas yr^{-1}]	ErrVal [mas^2 yr^{-2}]	Qlty
it car	0.877	8.094	0.209	0.770	0.760	-0.593	0.427	1
v car	0.826	7.359	0.157	0.713	0.736	-0.582	0.448	1
rz vel	1.310	7.082	0.293	0.834	0.851	$+1.334$	0.451	1
u vul	0.903	7.129	0.593	0.683	0.760	-0.649	0.455	1
x lac	0.736	8.406	0.339	0.563	0.618	$+0.482$	0.462	1
at pup	0.824	7.979	0.167	0.629	0.692	-2.058	0.469	1
su cyg	0.585	6.863	0.088	0.482	0.554	-2.150	0.476	1
v vel	0.640	7.575	0.212	0.571	0.616	$+1.703$	0.490	1
rv sco	0.782	7.041	0.338	0.628	0.707	-7.048	0.589	1
ry cma	0.670	8.106	0.223	0.624	0.687	-1.574	0.589	1
v cen	0.740	6.820	0.264	0.610	0.676	-4.199	0.614	1
xx cen	1.039	7.819	0.258	0.718	0.752	$+0.926$	0.656	1
s nor	0.989	6.429	0.178	0.766	0.778	-0.021	0.672	1
u aql	0.846	6.430	0.371	0.676	0.676	-3.168	0.681	1
vz cyg	0.687	8.957	0.274	0.604	0.640	-1.594	0.689	1
bb sgr	0.822	6.934	0.276	0.710	0.747	-1.639	0.690	1
bf oph	0.610	7.340	0.247	0.614	0.663	-0.182	0.713	1
xx car	1.196	9.322	0.343	0.722	0.765	$+0.492$	0.736	1
u sgr	0.829	6.695	0.403	0.694	0.730	-0.151	0.742	1
xy car	1.094	9.290	0.408	0.804	0.818	$+2.805$	0.757	1
rx aur	1.065	7.673	0.276	0.677	0.658	-0.320	0.759	1
sx vel	0.980	8.262	0.250	0.657	0.691	$+0.078$	0.765	1
s mus	0.985	6.128	0.220	0.626	0.657	-1.767	0.766	1
st tau	0.605	8.220	0.339	0.510	0.641	$+0.743$	0.773	1
ry vel	1.449	8.372	0.554	0.809	0.835	-0.224	0.783	1
ss sct	0.565	8.204	0.317	0.635	0.681	-0.622	0.806	1
hw car	0.964	9.128	0.184	0.813	0.879	$+0.750$	0.810	1
bg lac	0.727	8.887	0.316	0.643	0.669	$+0.871$	0.815	1
t vel	0.667	8.035	0.271	0.663	0.729	-3.376	0.815	1
dl cas	0.903	8.969	0.479	0.671	0.700	-1.004	0.821	1
yz car	1.259	8.711	0.372	0.754	0.787	-0.508	0.851	1
mw cyg	0.775	9.490	0.615	0.716	0.786	$+0.779$	0.854	1
uw car	0.728	9.418	0.439	0.583	0.649	$+0.701$	0.856	1
w sgr	0.881	4.670	0.112	0.632	0.679	-4.856	0.858	1
rx cam	0.898	7.678	0.536	0.665	0.735	-3.041	0.879	1
fm aql	0.786	8.275	0.617	0.686	0.700	-1.043	0.920	1
w gem	0.898	6.952	0.266	0.648	0.641	-0.660	0.924	1
sz cyg	1.179	9.430	0.587	0.916	0.881	-1.220	0.931	1
fr car	1.030	9.668	0.322	0.811	0.826	$+0.131$	0.940	1
vy cyg	0.895	9.584	0.615	0.626	0.669	-0.858	0.942	1

APPENDIX B. CEPHEID DATABASE

Star	$\log_{10} P$	V [mag]	E_{B-V} [mag]	$(B-V)_0$ [mag]	$(V-I)_0$ [mag]	μ_b [mas yr^{-1}]	ErrVal [mas^2 yr^{-2}]	Qlty
tx cyg	1.168	9.512	1.111	0.679	0.859	-1.403	0.946	1
fn aql	0.977	8.377	0.490	0.752	0.752	-3.395	0.969	1
sw vel	1.370	8.120	0.337	0.815	0.853	-2.780	0.979	1
z lac	1.037	8.416	0.378	0.718	0.735	-0.297	0.980	1
sz aql	1.234	8.631	0.552	0.881	0.858	-1.275	1.006	1
dd cas	0.992	9.879	0.493	0.709	0.666	$+0.229$	1.006	1
cd cyg	1.232	8.949	0.486	0.817	0.824	-1.721	1.050	1
dr vel	1.049	9.523	0.680	0.845	0.825	-0.623	1.053	1
gx car	0.857	9.341	0.379	0.668	0.720	-0.366	1.060	1
v lac	0.697	8.941	0.315	0.558	0.648	-0.451	1.072	1
vz pup	1.365	9.626	0.452	0.704	0.746	$+0.367$	1.075	1
ry cas	1.084	9.948	0.613	0.760	0.758	-3.193	1.080	1
uz car	0.716	9.331	0.184	0.696	0.729	$+1.237$	1.081	1
rs cas	0.799	9.942	0.784	0.701	0.719	-0.349	1.100	1
rr lac	0.808	8.847	0.296	0.589	0.651	-0.128	1.110	1
wz sgr	1.339	8.027	0.428	0.973	0.950	$+0.285$	1.148	1
cp cep	1.252	10.576	0.702	0.954	0.903	-0.270	1.155	1
rt aur	0.572	5.448	0.049	0.542	0.575	-8.505	1.173	1
aq car	0.990	8.853	0.158	0.768	0.782	$+0.509$	1.186	1
ux car	0.566	8.285	0.091	0.553	0.619	-2.315	1.195	1
ad pup	1.133	9.897	0.343	0.700	0.748	$+0.654$	1.216	1
v1162 aql	0.731	7.808	0.187	0.694	0.723	-11.521	1.265	1
uu mus	1.066	9.783	0.400	0.750	0.780	-0.221	1.299	1
rw cas	1.170	9.226	0.409	0.795	0.813	-3.735	1.299	1
ay cen	0.725	8.813	0.295	0.763	0.742	-0.687	1.311	1
wz car	1.362	9.259	0.362	0.788	0.822	-0.149	1.330	1
st vel	0.768	9.699	0.496	0.730	0.776	-1.895	1.357	1
ae vel	0.853	10.242	0.639	0.631	0.711	-1.704	1.379	1
cr cep	0.794	9.656	0.697	0.727	0.784	-0.525	1.382	1
z sct	1.111	9.586	0.491	0.855	0.827	-0.025	1.404	1
rt mus	0.490	8.990	0.292	0.547	0.654	-3.422	1.449	1
vx per	1.037	9.306	0.496	0.659	0.674	-3.760	1.458	1
gh cyg	0.893	9.897	0.629	0.651	0.660	-1.149	1.459	1
y aur	0.587	9.627	0.356	0.553	0.599	-0.738	1.465	1
sy cas	0.610	9.889	0.430	0.553	0.585	-1.252	1.476	1
yz sgr	0.980	7.341	0.285	0.749	0.761	-3.855	1.486	1
ss cma	1.092	9.941	0.533	0.690	0.763	-0.909	1.488	1
vx cyg	1.304	10.078	0.830	0.890	0.852	-0.136	1.504	1
sy aur	1.006	9.062	0.453	0.596	0.610	$+0.053$	1.598	1
kn cen	1.532	9.853	0.774	0.839	0.871	-2.013	1.600	1

Star	$\log_{10} P$	V	E_{B-V}	$(B-V)_0$	$(V-I)_0$	μ_b	ErrVal	Qlty
		[mag]	[mag]	[mag]	[mag]	[mas yr^{-1}]	[mas^2 yr^{-2}]	
x sct	0.623	10.010	0.557	0.617	0.675	-3.766	1.613	1
vy car	1.277	7.465	0.260	0.905	0.861	-0.146	1.663	1
tt aql	1.139	7.129	0.462	0.843	0.815	-2.565	1.680	1
x pup	1.414	8.526	0.409	0.798	0.830	$+1.260$	1.702	1
xy cas	0.653	9.976	0.480	0.657	0.624	$+0.607$	1.706	1
sx car	0.687	9.090	0.310	0.608	0.656	-1.371	1.713	1
sw cas	0.736	9.706	0.449	0.635	0.683	-2.961	1.732	1
ap sgr	0.704	6.951	0.174	0.647	0.687	-4.389	1.840	1
bn pup	1.136	9.890	0.417	0.778	0.798	-0.415	1.899	1
vw cas	0.777	10.756	0.485	0.730	0.742	-0.383	1.972	1
ww car	0.670	9.748	0.392	0.507	0.603	-1.302	1.993	1
sv vel	1.149	8.587	0.365	0.723	0.782	-1.276	2.017	1
x vul	0.801	8.843	0.790	0.606	0.633	-1.270	2.017	1
ww pup	0.742	10.613	0.362	0.491	0.643	-0.951	2.069	1
xz car	1.221	8.596	0.341	0.926	0.919	$+0.028$	2.071	1
u nor	1.102	9.228	0.862	0.754	0.773	$+1.793$	2.202	1
y sct	1.015	9.627	0.767	0.773	0.801	-1.924	2.249	1
br vul	0.716	10.686	0.866	0.601	0.553	$+2.451$	2.264	1
bk aur	0.903	9.438	0.424	0.647	0.645	-3.927	2.273	1
cn car	0.693	10.683	0.395	0.715	0.825	-0.589	2.324	1
y lac	0.635	9.147	0.202	0.530	0.592	-1.449	2.345	1
cd cas	0.892	10.788	0.748	0.708	0.793	-1.438	2.443	1
wy pup	0.720	10.579	0.292	0.565	0.576	-2.090	2.512	1
vw cen	1.177	10.250	0.417	0.941	0.962	-1.735	2.605	1
ey car	0.459	10.322	0.335	0.518	0.653	$+1.432$	2.664	1
sv mon	1.183	8.268	0.250	0.796	0.815	-0.628	2.891	1
ch cas	1.179	11.000	0.955	0.690	0.748	-1.265	3.255	1
rz gem	0.743	10.021	0.503	0.532	0.629	-0.554	3.505	1
uz cas	0.629	11.370	0.434	0.640	0.686	-4.582	3.647	1
tz mon	0.871	10.789	0.431	0.701	0.769	-1.132	3.745	1
cr ser	0.724	10.858	0.961	0.661	0.729	-2.261	3.888	1
aa gem	1.053	9.728	0.380	0.692	0.659	$+1.483$	3.958	1
ck sct	0.870	10.613	0.784	0.767	0.830	-2.532	4.248	1
cf cas	0.688	11.136	0.531	0.668	0.701	-1.945	4.295	1
cv mon	0.731	10.304	0.702	0.601	0.757	-5.954	4.369	1
ac mon	0.904	10.096	0.507	0.672	0.738	$+0.258$	4.469	1
ru sct	1.294	9.463	0.930	0.742	0.798	-2.093	4.716	1
aq pup	1.479	8.704	0.531	0.832	0.878	$+4.058$	5.429	1
kq sco	1.458	9.810	0.839	1.098	1.076	-1.465	5.452	1
vv cas	0.793	10.744	0.482	0.653	0.699	-3.146	5.757	1

APPENDIX B. CEPHEID DATABASE

Star	$\log_{10} P$	V [mag]	E_{B-V} [mag]	$(B-V)_0$ [mag]	$(V-I)_0$ [mag]	μ_b [mas yr^{-1}]	ErrVal [mas^2 yr^{-2}]	Qlty
vw pup	0.632	11.383	0.483	0.632	0.684	-0.988	6.000	1
cy cas	1.158	11.623	0.947	0.720	0.784	-0.194	6.035	1
sy nor	1.102	9.502	0.696	0.661	0.705	$+2.789$	6.037	1
gu nor	0.538	10.355	0.629	0.658	0.753	$+1.059$	6.132	1
ty sct	1.043	10.815	0.937	0.774	0.788	-4.044	6.468	1
ek mon	0.598	11.069	0.551	0.655	0.752	-0.632	6.872	1
tx mon	0.940	10.967	0.492	0.619	0.702	$+2.898$	7.687	1
kk cen	1.086	11.461	0.572	0.734	0.782	$+0.351$	7.814	1
ay sgr	0.818	10.620	0.841	0.582	0.762	-1.012	8.917	1
ak cep	0.859	11.202	0.635	0.693	0.731	$+0.757$	9.091	1
ry sco	1.308	8.012	0.714	0.751	0.820	-0.576	10.651	1
tw nor	1.033	11.667	1.214	0.791	0.822	-0.686	22.464	1
su cru	1.109	9.782	0.998	0.774	0.845	-14.849	40.347	1
ux per	0.660	11.633	0.512	0.525	0.502	$+13.300$	75.928	1
s vul	1.836	8.962	0.737	1.152	1.076	$+1.217$	1.320	2
v1344 aql	0.874	7.772	0.569	0.788	0.731	-1.152	1.820	2
x cru	0.794	8.395	0.284	0.709	0.739	-1.062	2.100	2
gy sge	1.713	10.150	1.236	1.049	1.064	-2.880	2.560	2
bz cyg	1.006	10.218	0.839	0.762	0.817	$+0.411$	3.060	2
bv mon	0.479	11.381	0.582	0.513	0.580	$+3.702$	3.610	2
uy car	0.744	8.949	0.177	0.660	0.725	-5.835	4.370	2
ez cyg	1.067	11.052	0.784	0.688	0.597	$+1.669$	5.060	2
ag cru	0.584	8.211	0.226	0.523	0.570	$+2.585$	9.300	2
vw cru	0.722	9.604	0.603	0.716	0.853	-6.952	10.730	2
fi car	1.129	11.613	0.684	0.881	0.907	$+3.095$	10.880	2
ct car	1.257	12.234	0.553	0.831	0.834	$+1.817$	13.680	2
xx mon	0.737	11.915	0.586	0.597	0.662	-4.292	14.040	2
ap cas	0.836	11.565	0.787	0.620	0.549	-5.801	14.820	2
dw per	0.562	11.579	0.591	0.577	0.561	$+10.366$	16.800	2
ex vel	1.122	11.578	0.809	0.780	0.773	-4.433	19.780	2
fo car	1.015	10.735	0.466	0.804	0.853	-5.174	20.210	2
uy per	0.730	11.343	0.869	0.606	0.739	$+12.293$	24.500	2
vy sgr	1.132	11.448	1.143	0.865	0.792	-5.473	27.000	2
vy per	0.743	11.255	0.935	0.604	0.754	$+7.934$	36.580	2

Appendix C

Matlab function: kinematiccheck.m

```
function [ returnData ] = kinematiccheck( plcMatrix , wSun , includedistanceinerrorvalue ,
    requireIcolour )
%KINEMATICCHECK takes a list of PLCs and do a kinematic check on them, based on a database
    of 215 galactic cepheids.
%
% Input parameters:
%   plcMatrix should contain a PLC per row that defines an absolute visual
%   magnitude and must have four coloumns for the constants:
%       alpha, beta_BV, beta_VI, gamma
%   wSun is the Sun's vertical speed in km/s, usually around 7 km/s
%   includedistanceinerrorvalue == 1 means that the PLC distance is included in the error
    value
%   requireIcolour == 1 means that all PLC relations will require I colour info even if it
    is not used
%
% returnData structure:
%   one row for each PLC with the following columns:
%   N = maxium sample size
%   n = optimal sample size based on velocity confidence interval
%   w_quants = the quantiles of the velocity distribution based on binomial distribution
%   w_diff = the length of the confidence interval
%   distributions = all the velocity distributions for each sample
%   n = optimal sample size based on the magnitude confidence interval
%   mag_quants = the quantiles translated to magnitudes
%   mag_diff = the length of the interval in magnitudes
%   error_value = sorted list of error values used

    % constants in use
    Rv = 3.17; % +/-0.13 reddening factor
    Ri = 1.89; % +/-0.12 reddening factor
    Rb = 4.17; % +/-0.15 reddening factor
    K = 4.7405; % conversion from mas/yr to km/s/kpc
    GtrnspE = [ -0.054875560416215  -0.873437090234885  -0.483835015548713 ;
        0.494109427875584  -0.444829629960011   0.746982244497219 ;
        -0.867666149019005  -0.198076373431202   0.455983776175067 ];
    alphaG = 192.85948; % [deg]
    deltaG = 27.1282; % [deg]
    lOmega = 32.93192; % [deg

    % read in the database
    cephdata = dlmread('cepheiddata.txt',';');
    cephdata(cephdata(:,18)==0,:) = []; % delete entries with no proper motion

    % add error values as 19th coloumn and sort matrix by it
    ErrVal = cephdata(:,15).*cephdata(:,16) + (cephdata(:,18)-1)*1000;
```

41

```matlab
cephdata = [ cephdata , ErrVal ];
cephdata = sortrows(cephdata, size(cephdata,2)); % sort on last column


ID       = cephdata(:, 1); %              - ID number, alternative list for name lookup
logP     = cephdata(:, 2); %              - log_10 of period in days
Bmag     = cephdata(:, 3); % [mag]        - reddened B magnitude
Vmag     = cephdata(:, 4); % [mag]        - reddened V magnitude
Imag     = cephdata(:, 5); % [mag]        - reddened I magnitude (0 if no data)
EBVf     = cephdata(:, 6); % [mag]        - colour excess E(B-V) fernie system
EBV      = cephdata(:, 7); % [mag]        - colour excess E(B-V) corrected
EVI      = cephdata(:, 8); % [mag]        - colour excess E(V-I) corrected
BV0      = cephdata(:, 9); % [mag]        - intrinsic (B-V)_0 colour
VI0      = cephdata(:,10); % [mag]        - intrinsic (V-I)_0 colour (0 if no data)
RA       = cephdata(:,11); % [deg]        - position, right ascension
DE       = cephdata(:,12); % [deg]        - position, declination
PMra     = cephdata(:,13); % [mas/yr]     - proper motion, right ascension (0 if no data)
PMde     = cephdata(:,14); % [mas/yr]     - proper motion, declination (0 if no data)
Ema      = cephdata(:,15); % [mas/yr]     - error elipse, major axis (0 if no data)
Emi      = cephdata(:,16); % [mas/yr]     - error elipse, minor axis (0 if no data)
Eang     = cephdata(:,17); % [deg]        - error elipse, angle (0 if no data)
Qlty     = cephdata(:,18); %              - 1=good quality, 2=not as good (0 if no data)
ErrVal   = cephdata(:,19); %              - redefined to make sure it is in sorted order


% apparent visual magnitude reddening corrected, i.e. intrinsic apparent visual
%     magnitude
Vmag0 = Vmag - Rv*EBV;

% calculate proper motion
C1 = sind(deltaG).*cosd(DE)-cosd(deltaG).*sind(DE).*cosd(RA-alphaG);
C2 = cosd(deltaG).*sind(RA-alphaG);
mu_b = (1./sqrt(C1.^2 + C2.^2)) .* (-C2 .* PMra + C1 .* PMde);




% analyse each PLC in turn and store results
returnData = cell(size(plcMatrix,1),9); % prepare data structure
for plcIdx = 1:size(plcMatrix,1)
    alpha = plcMatrix(plcIdx,1);
    beta_BV = plcMatrix(plcIdx,2);
    beta_VI = plcMatrix(plcIdx,3);
    gamma = plcMatrix(plcIdx,4);

    selection = ID > 0; % as default select all cepheids
    if beta_VI > 0 || requireIcolour > 0
        selection = selection & Imag > 0; % if plc requires Imag info, reduce
%             selection
    end
    N = sum(selection);
    returnData{plcIdx,1} = N; % store number of cepheids for given selection

    % defined the selected data sets
    sel_logP  = logP(selection);
    sel_BV0   = BV0(selection);
    sel_VI0   = VI0(selection);
    sel_Vmag0 = Vmag0(selection);
    sel_mu_b  = mu_b(selection);

    % optional choice to modify the error value sorting by including the distance
%         obtained via the PLC
    if includedistanceinerrorvalue == 1
        tmp_dist = findDistance( sel_Vmag0, alpha*sel_logP + beta_BV*sel_BV0 + beta_VI
%             *sel_VI0 + gamma );
        tmp_ErrVal = ErrVal(selection);
        tmp_ErrVal(tmp_ErrVal>1000) = tmp_ErrVal(tmp_ErrVal>1000)-1000; % undo quality
```

42

```
                adjustment
        tmp_ErrVal = tmp_ErrVal .* tmp_dist + (Qlty(selection)-1)*1000; % and then
                reintroduce it
        [ tmp_ErrVal, sortIndex ] = sort(tmp_ErrVal);
        sel_logP = sel_logP(sortIndex);
        sel_BV0 = sel_BV0(sortIndex);
        sel_VI0 = sel_VI0(sortIndex);
        sel_Vmag0 = sel_Vmag0(sortIndex);
        sel_mu_b = sel_mu_b(sortIndex);
        returnData{plcIdx,9} = tmp_ErrVal;
else
        returnData{plcIdx,9} = ErrVal;
end


% define the PLC
% - gammacorr is the correction the PLC is run with
% - n reduces the dataset so the PLC can be analysed for different cepheid samples
plc = @(gammacorr,n) alpha * sel_logP(1:n) + beta_BV * sel_BV0(1:n) + beta_VI *
    sel_VI0(1:n) + gamma + gammacorr;

% find the w distribution
w_dist = @(gammacorr,n) findVertSpeed( sel_mu_b(1:n) , sel_Vmag0(1:n) , plc(
    gammacorr,n) , K , wSun );

% analyze the PLC for each sample of cepheids n = 1..N
w_distributions = cell(N,1);
w_quants = zeros(N,3);
plc_corr = zeros(N,3);
for n = 1:N
    % determine the speed distribution without doing any correction
    dist = w_dist(0,n);
    w_distributions{n,1} = dist;

    % find the quantiles of above speed distribution
        speeds = quantile( dist , [0.5 * (1 - 1/sqrt(n)), 0.5,  0.5 * (1 + 1/sqrt(
            n))] );
    w_quants(n,:) = speeds;

    % translate speed quantiles into magnitude quantiles
    % first by defining functions that should be zero given a correction
    neg_diff = speeds(2) - speeds(1);
    pos_diff = speeds(3) - speeds(2);
    neg_func = @(corr) median(w_dist(corr,n)+neg_diff);
    mid_func = @(corr) median(w_dist(corr,n));
    pos_func = @(corr) median(w_dist(corr,n)-pos_diff);
    % then find the corrections needed to make above functions zero, initial guess
            no correction
    [ neg_corr , ~, exitvalue ] = fzero(neg_func, 0);
    if exitvalue ~= 1
        display(['error [neg] with sample = ',num2str(n)]);
    end
    [ mid_corr , ~, exitvalue ] = fzero(mid_func, 0);
    if exitvalue ~= 1
        display(['error [mid] with sample = ',num2str(n)]);
    end
    [ pos_corr , ~, exitvalue ] = fzero(pos_func, 0);
    if exitvalue ~= 1
        display(['error [pos] with sample = ',num2str(n)]);
    end
    plc_corr(n,:) = [ neg_corr , mid_corr , pos_corr ];

end
% find sample with smallest diff
diff = w_quants(:,3)-w_quants(:,1);
```

```matlab
            diff(1:10) = max(diff); % make sure the first 10 entries are not found as best
            returnData{plcIdx,2} = find(diff == min(diff), 1);

            % store the distributions along with the quants and difference for plotting
            returnData{plcIdx,3} = w_quants;
            returnData{plcIdx,4} = w_quants(:,3)-w_quants(:,1); % store the diff
            returnData{plcIdx,5} = w_distributions;

            % also find optimal sample for plc correction, to check if same
            diff = plc_corr(:,3)-plc_corr(:,1);
            diff(1:10) = max(diff); % make sure the first 10 entries are not found as best
            returnData{plcIdx,6} = find(diff == min(diff), 1);

            % store plc correction
            returnData{plcIdx,7} = plc_corr;
            returnData{plcIdx,8} = plc_corr(:,3)-plc_corr(:,1); % store the diff

    end

end


% define a function that takes an absolute visual magnitude, turns it into
% a distance modulus and find the distance in kilo parsecs
function [ local_d ] = findDistance( local_V0 , local_Mv ) % 'local_' to avoid scope
    issues
    local_d = (10 * 10.^((local_V0-local_Mv)/5)) / 1000; % [kpc]
end

% define a function that takes an absolute visual magnitude, turns it into a
% distance modulus and based on proper motion find the vertical speed in km/s
function [ local_w ] = findVertSpeed( local_mu_b , local_V0 , local_Mv , K , wSun ) % '
    local_' to avoid scope issues
    local_w = local_mu_b .* findDistance( local_V0 , local_Mv ) .* K + wSun; % [km/s]
end
```

Appendix D

Matlab example script

```matlab
% script that analyses the selected PLCs
clear all;
close all;

plcMatrix = [ -2.76, 0, 0, -1.40;          % 1991-MF
    -2.81, 0, 0, -1.43 ;                    % 1997-FC
    -3.80, 2.70, 0, -2.38;                  % 1997-FW
    -2.12, 0, 0, -1.73;                     % 1998-Lu
    -2.77, 0, 0, -1.44;                     % 1999-La
    -3.255, 0, 2.45, -2.644;                % 2001-Fr
    -3.141, 0, 0, -0.826;                   % 2003-Ta
    -3.29, 0, 2.45, -2.576;                 % 2007-vL
    -3.08, 0, 0, -0.94;                     % 2013-An
    -3.245, 0, 2.45, -2.545];               % 2013-Ma

wSun = 7.0;


data = analyseplcs(plcMatrix, wSun, 0, 0);

display('════════════════════════════════════════════════════');
display(['wSun_=_',num2str(wSun),'_km/s']);


for plcIdx = 1:size(plcMatrix,1)
    N = data{plcIdx,1};
    sample = data{plcIdx,2};
    w_quants = data{plcIdx,3};
    w_diff = data{plcIdx,4};
    w_dists = data{plcIdx,5};
    sample2 = data{plcIdx,6};
    mag_quants = data{plcIdx,7};
    mag_diff = data{plcIdx,8};

    display('_');
    display(['PLC_',num2str(plcIdx)]);
    display(['Prefered_sample_is_n_=_',num2str(sample),'_(check_vs:',num2str(sample2),')'
        ]);
    display(['Magnitude_correction_=_',num2str(mag_quants(sample2,2)),',_interval:_[',
        num2str(mag_quants(sample2,1)),',',num2str(mag_quants(sample2,3)),'],_length_=_',
        num2str(mag_diff(sample2))]);
end
```

45

Appendix E

Velocity distribution plots
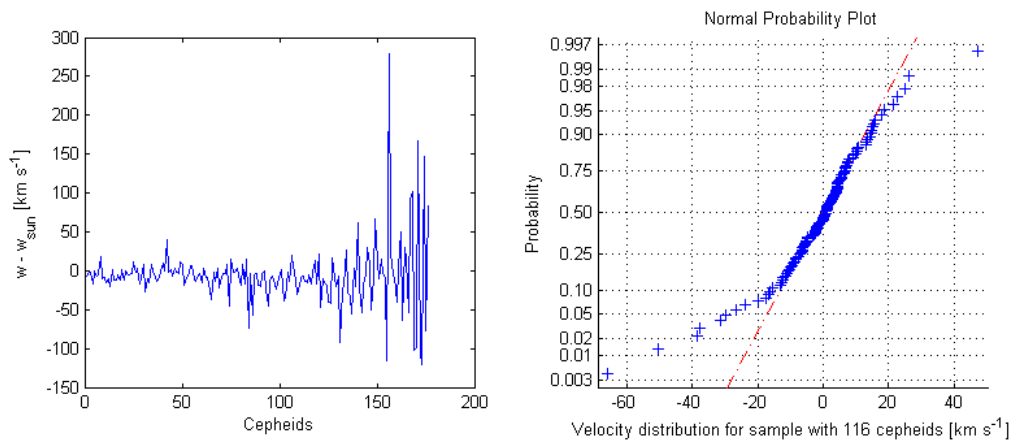
Figure E.1: Velocity distributions for 1991-MF.
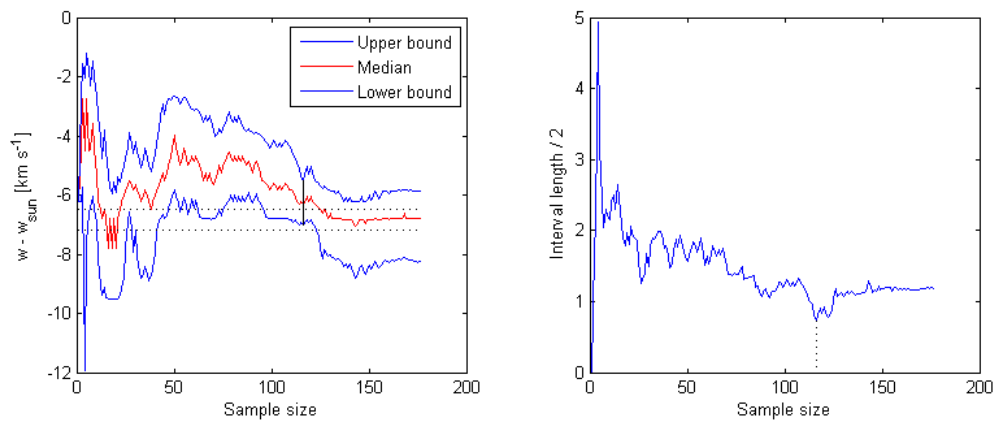
Figure E.2: Confidence intervals in terms of velocity for 1991-MF.
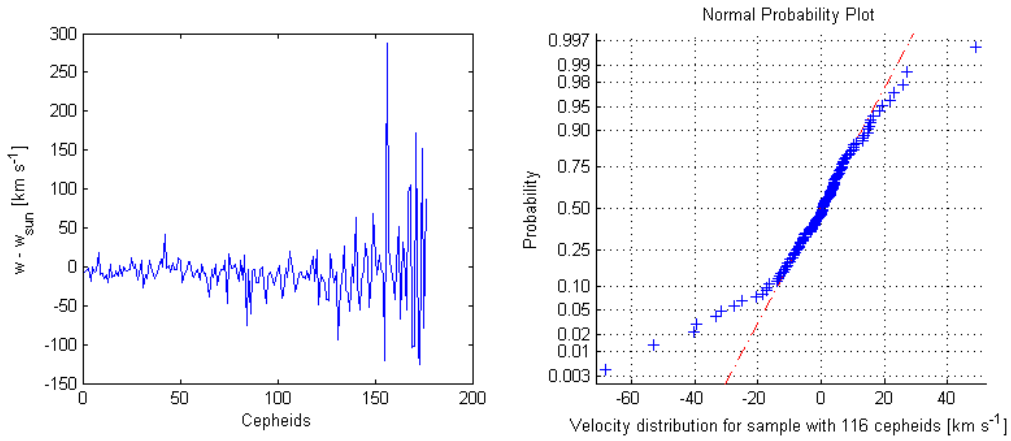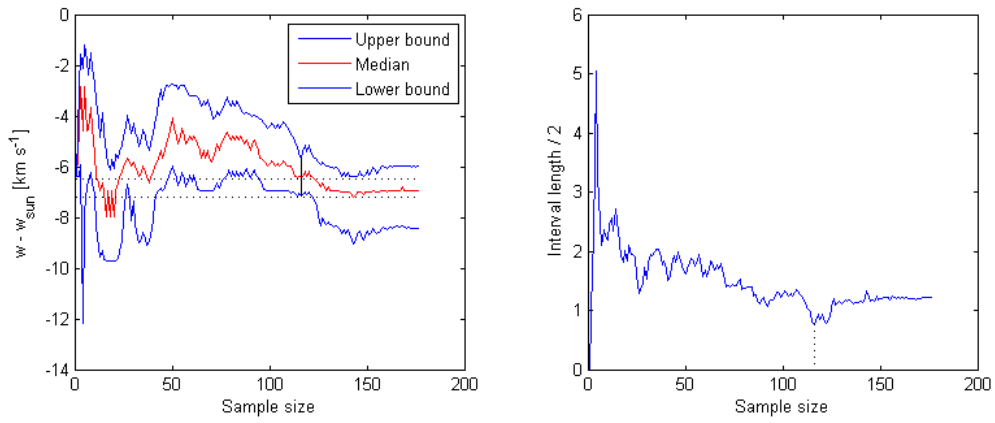
Figure E.3: Velocity distributions for 1997-FC.
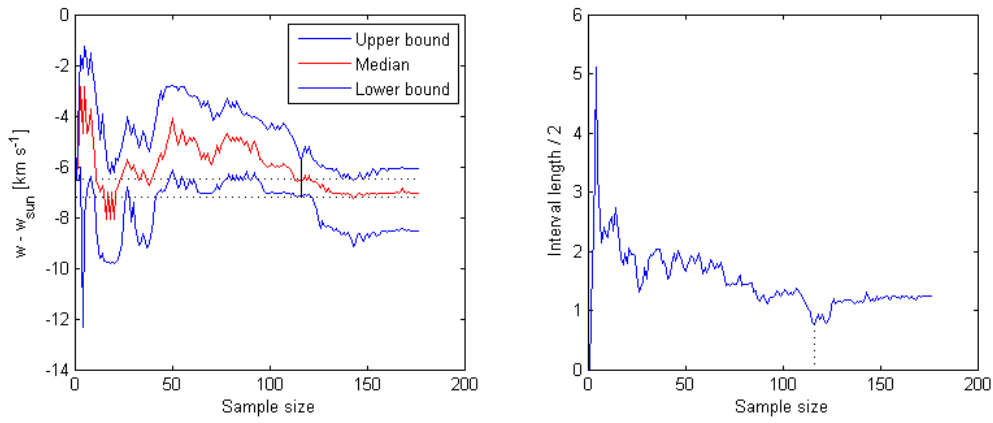
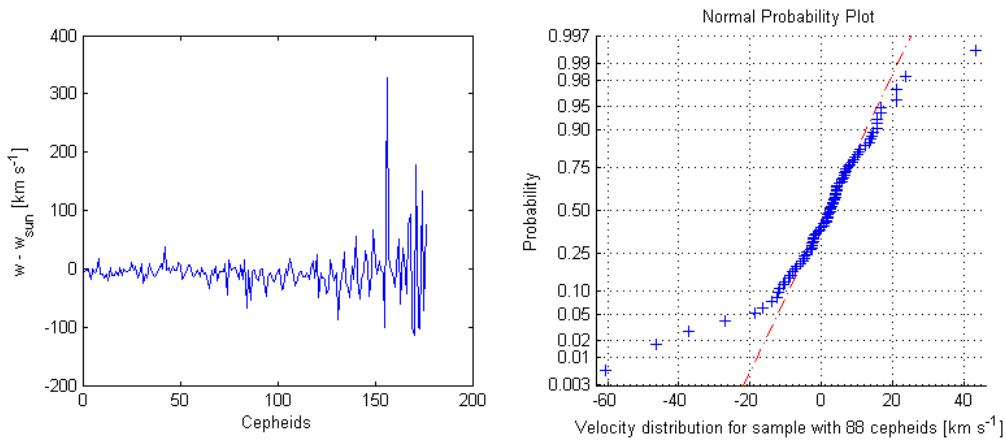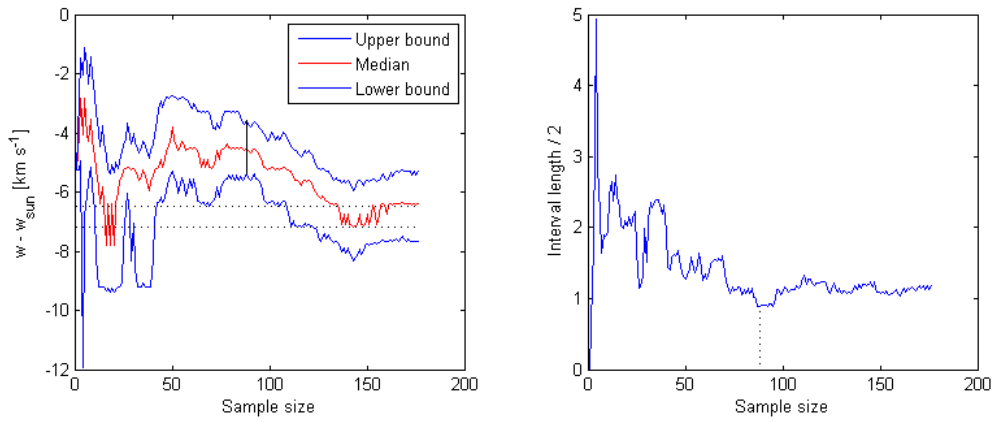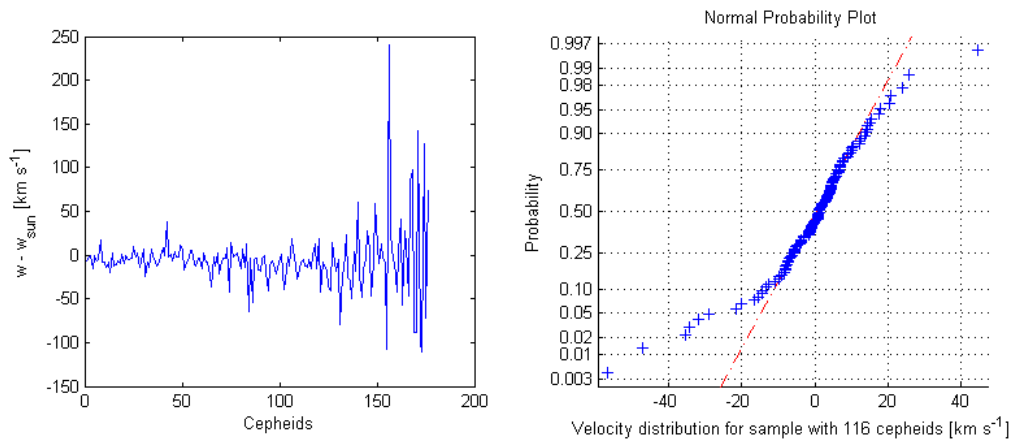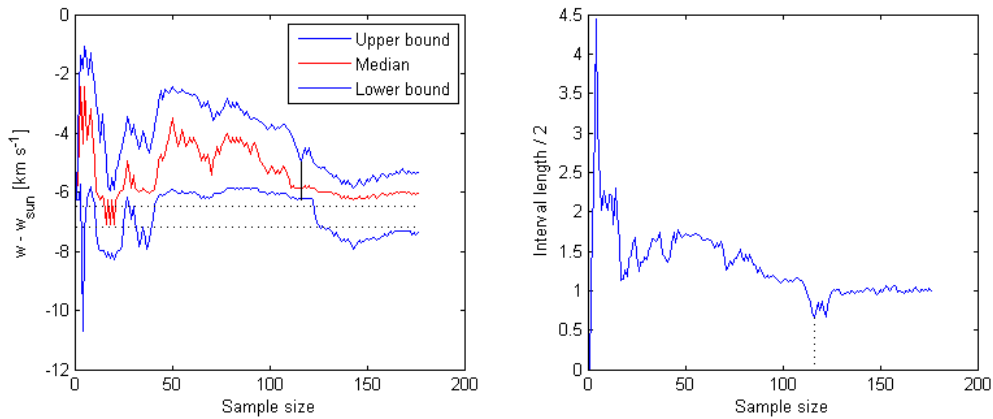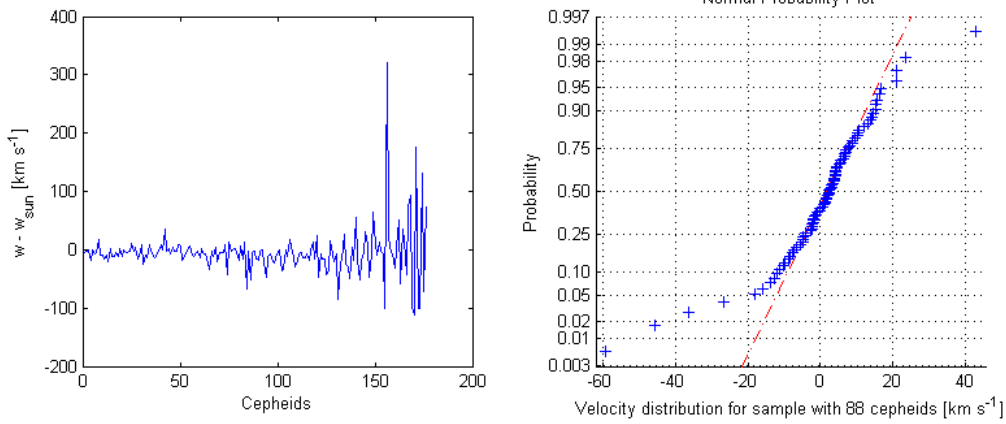Figure E.4: Confidence intervals in terms of velocity for 1997-FC.
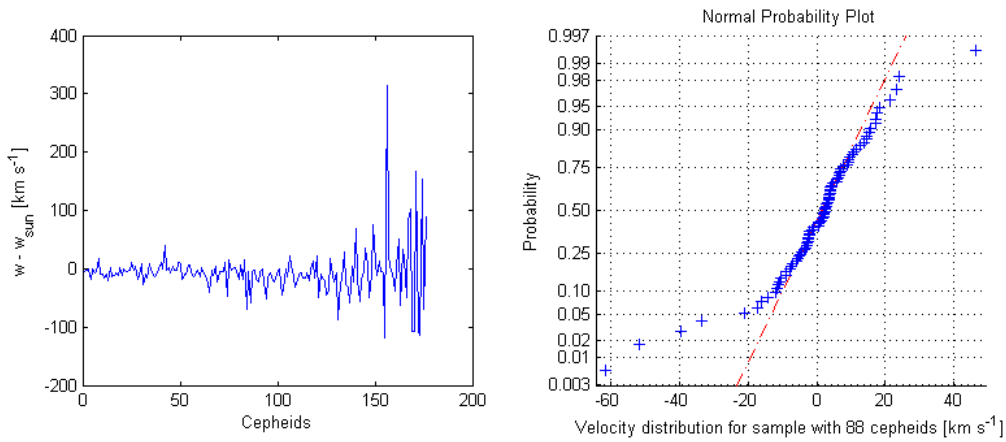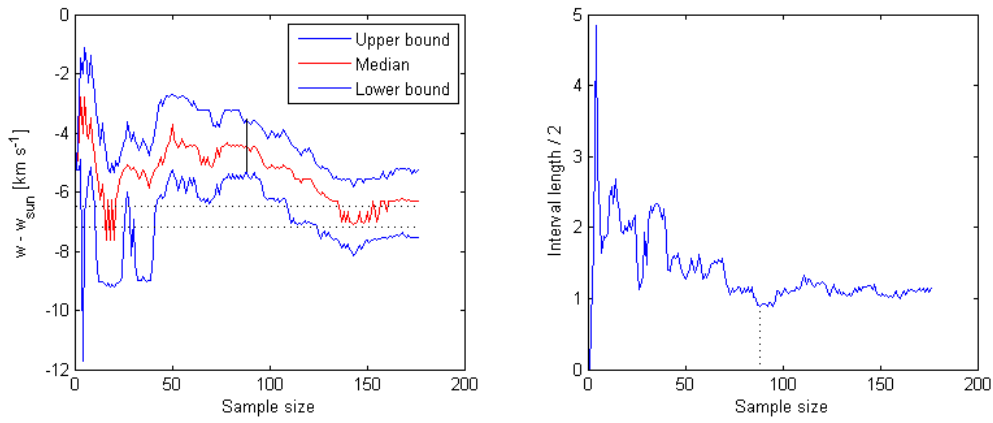

Figure E.5: Velocity distributions for 1997-FW.
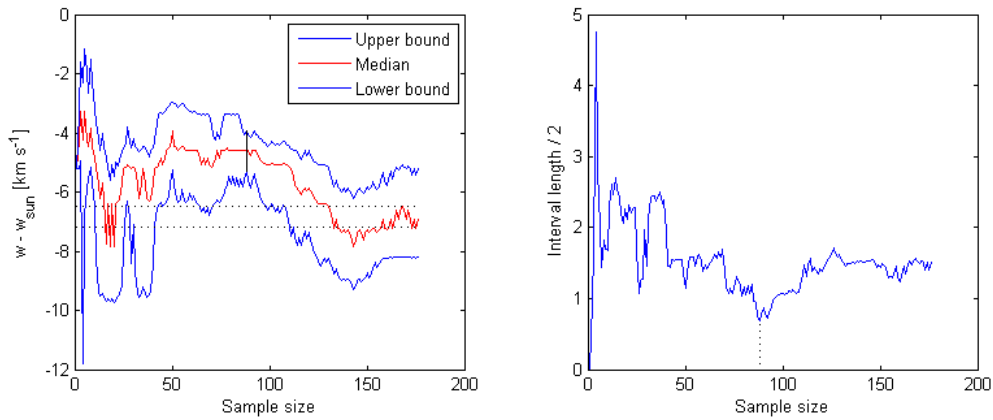
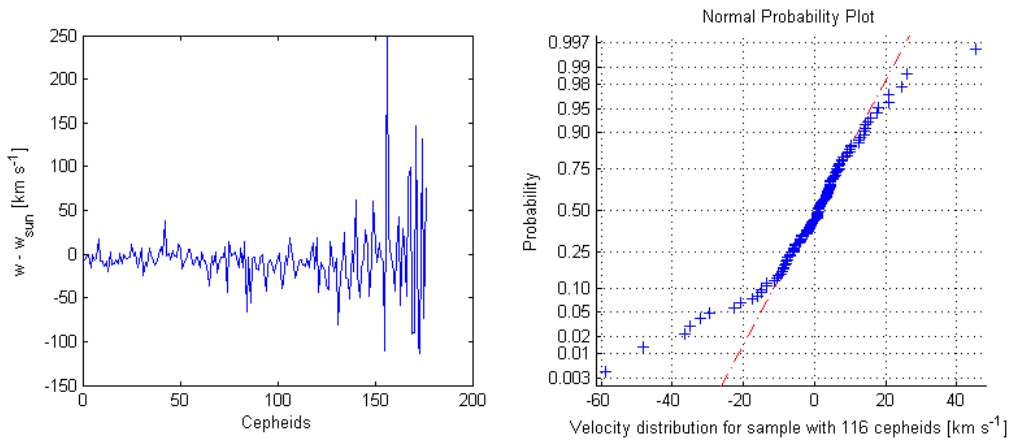Figure E.6: Confidence intervals in terms of velocity for 1997-FW.
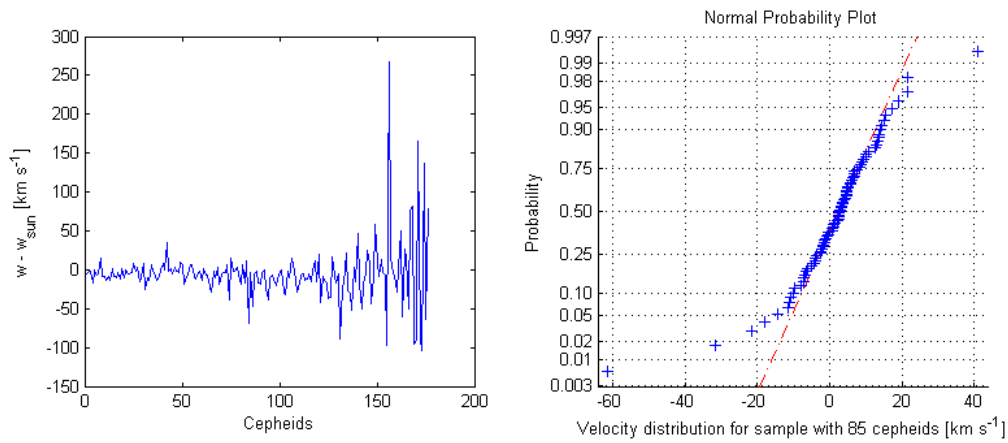
Figure E.7: Velocity distributions for 1998-Lu.
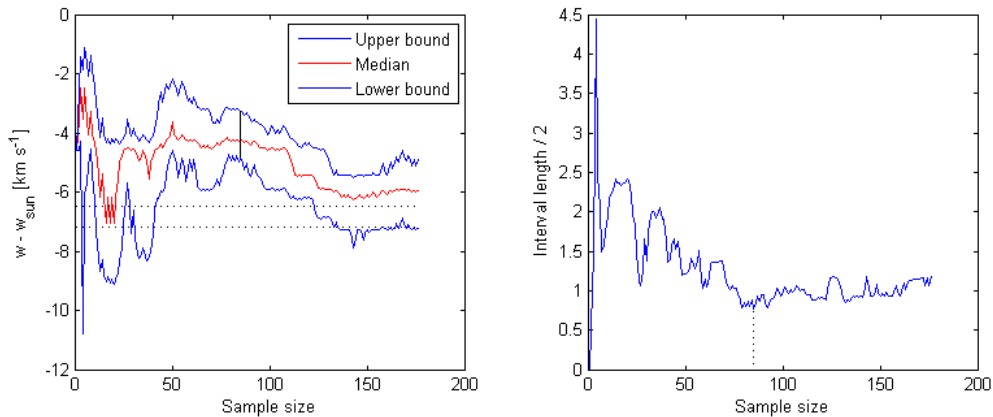
Figure E.8: Confidence intervals in terms of velocity for 1998-Lu.
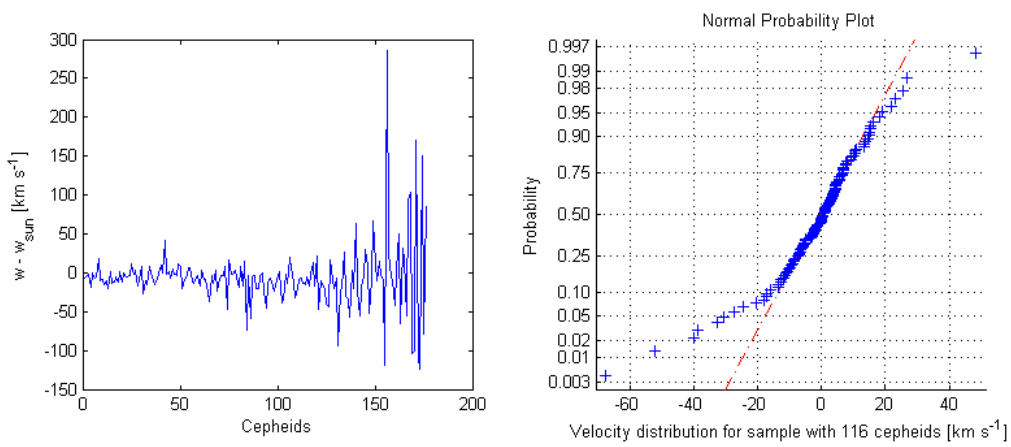
Figure E.9: Velocity distributions for 1999-La.

Figure E.10: Confidence intervals in terms of velocity for 1999-La.

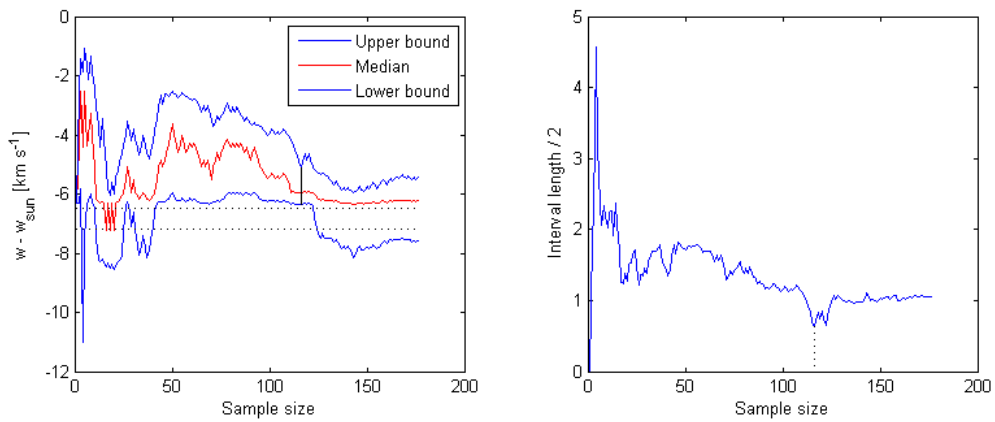Figure E.11: Velocity distributions for 2001-Fr.

Figure E.12: Confidence intervals in terms of velocity for 2001-Fr.

Figure E.13: Velocity distributions for 2003-Ta.

Figure E.14: Confidence intervals in terms of velocity for 2003-Ta.

Figure E.15: Velocity distributions for 2007-vL.

Figure E.16: Confidence intervals in terms of velocity for 2007-vL.

Figure E.17: Velocity distributions for 2013-An.

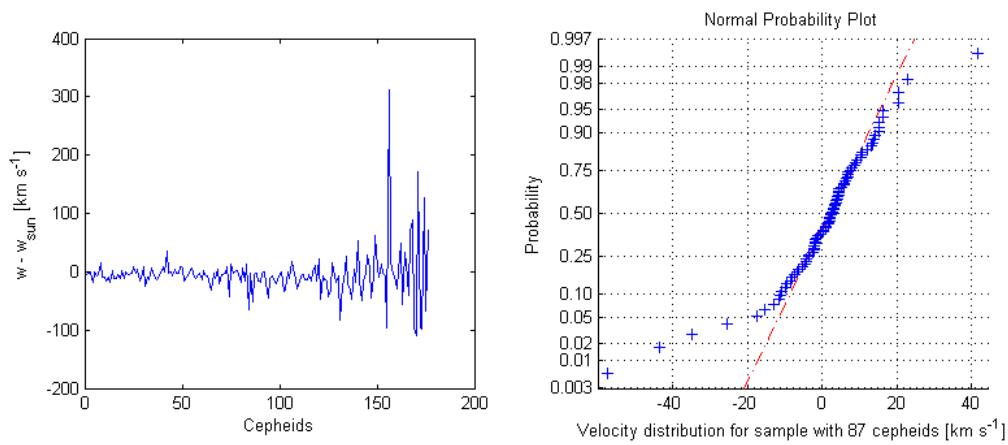Figure E.18: Confidence intervals in terms of velocity for 2013-An.
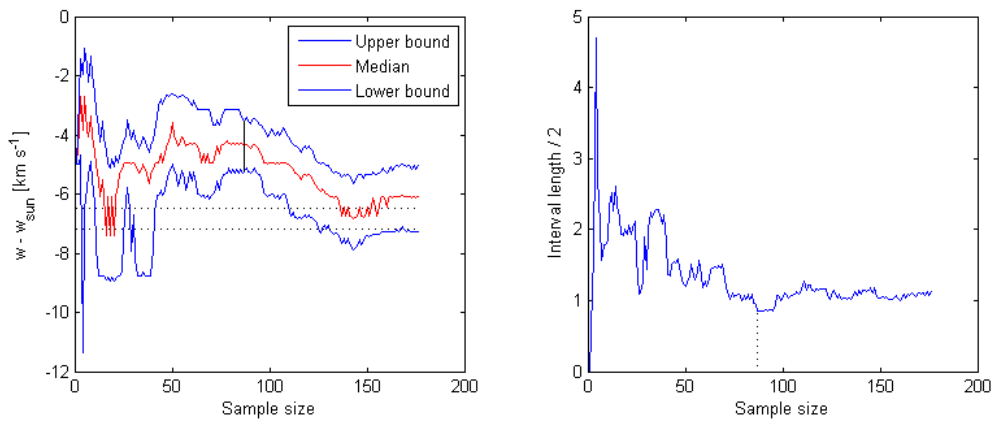

Figure E.19: Velocity distributions for 2013-Ma.

Figure E.20: Confidence intervals in terms of velocity for 2013-Ma.