RR Lyrae Stars as Probes of Structure in the Milky Way Galaxy: Tools and Applications

por

GABRIEL TORREALBA ARANCIBIA.

Tesis presentada a la Facultad de Física de la Pontificia Universidad Católica de Chile, como requisito para optar al grado académico de Magíster en Astrofísica.

PROFESOR GUÍA : Dr. Márcio Catelan
COMISIÓN EXAMINADORA : Dr. Andrés Jordán
Dr. Julio Chanamé

Abril de 2013
SANTIAGO – CHILE
Agradecimientos

Me gustaría agradecer a todas aquellas personas que me han apoyado en dar este importante paso en mi vida profesional. En particular me gustaría agradecer a mi profesor guía, Márício, que me ha entregado infinidad de oportunidades para poder desarrollarme a lo largo de este magíster, además de toda la ayuda y confianza que me ayudaron a conseguir los objetivos que plantean este desafío. A mis padres, Pablo y Verónica, que han hecho de mí lo que soy y me han guiado a lo largo de mi vida. Y muy especialmente a Giovanna, que sin duda es un pilar fundamental.

This work was supported by the Chilean Ministry for the Economy, Development, and Tourism’s Programa Iniciativa Científica Milenio through grant P07-021-F, awarded to The Milky Way Millennium Nucleus; by the BASAL Center for Astrophysics and Associated Technologies (PFB-06); by Proyecto Fondecyt Regular #1110326; by Proyecto Anillo ACT-86; by proyecto Fondecyt (11100240); and by Conicyt Master’s Scholarship.

A todos ustedes: ¡Muchas Gracias!

Gabriel Torrealba
Abstract

We present a compendium of two works regarding computational approach to RR Lyrae in surveys. We believe that improving programming skills and statistical knowledge is fundamental to tackle the challenges that future surveys like GAIA and LSST will bring. In this sense, we present a set of tools that will result useful to get the better science from RR Lyrae and variable stars in general from the huge amounts of data that will be available when the next survey generation starts. In particular, we first present a theoretical calibration of the metallicity and physical parameters (temperature, luminosity, gravity, mass, radius) for RR Lyrae stars using the $ugriz$ SDSS photometric system in the form of tight analytical fits that makes use of photometric data only. These fits are based on calculations of synthetic horizontal branches and are able to provide all the quantities mentioned with very high (internal) precision. Second, we present algorithms and techniques to extract RRab stars from the southern part of the Catalina Surveys. We developed two algorithms, Automatic Period Selection (APS) and Automatic Fourier Decomposition (AFD), to find better period and lightcurves from V band photometric data. With the APS output we classified 10541 RRab stars in an area that covers more than 14800 square degrees in the declination range of $-75^\circ \leq \delta \leq -15^\circ$. The selected sample is provided with photometric metallicities, unreddened distances and galactocentric positions. We performed an automatic study of the overdensities on the southern part of the halo finding 49 overdensity candidates with significances as high as 10$\sigma$ compared to halo models.
# Contents

Abstract .................................................................................................................. 1
Table of contents ...................................................................................................... 2

1 Introduction ......................................................................................................... 3

2 RR Lyrae stars .................................................................................................... 5
   2.1 Properties ...................................................................................................... 5
   2.2 Variability mechanisms ............................................................................... 6
   2.3 Metallicities and Distances ........................................................................... 9
   2.4 RRL Applications ........................................................................................ 10

3 Physical parameters and metallicities on the SDSS system ................................. 12
   3.1 Models ........................................................................................................ 14
   3.2 Metallicity ................................................................................................... 15
   3.3 Physical Parameters .................................................................................... 18
      3.3.1 Temperature ....................................................................................... 18
      3.3.2 Luminosity ......................................................................................... 19
      3.3.3 Gravity ............................................................................................... 20
      3.3.4 Radius ............................................................................................... 24
Chapter 1

Introduction

The Sloan Digital Sky Survey (SDSS; York et al. 2000; Aihara et al. 2011) represents the dawn of a new era in Astronomy: the era of wide-field sky surveys. Such surveys are of key importance for our understanding of structure in the Universe, both locally and at cosmological distances. This new era in Astronomy comes with great benefits from large-scale studies that were impossible to carry out in the past, but also carries great challenges both from the technology needed and the development of fast and efficient algorithms to deal with huge amounts of data. In the future, increasingly powerful surveys are planned (e.g., GAIA, LSST), creating increasingly complex difficulties. In this sense, it is important that, besides the particular scientific motivation, statistical and programming knowledge and skills are developed to deal with the upcoming challenges. In this thesis we developed two different works, both linked by wide-field astronomy, surveys and RR Lyrae.

Of particular interest in this new era of wide-field sky surveys, given the transient nature of most of the upcoming surveys, are variable stars. We focus mainly on RR Lyrae because they excel as distance estimators, a key element to understand the local nature of our universe. In this sense, RR Lyrae can be used to trace the history of galaxy formation (e.g., Catelan 2009 and references therein). We give a further summary of RR Lyrae and their properties in §2
In the wake of the SDSS survey itself, the SDSS photometric system (Fukugita et al. 1996, 2011) has naturally gained much visibility. Indeed, SDSS filters are now commonly available in all major observatories, and many of the current and future wide-field dedicated telescopes and surveys, including (to name just a few) the Large-Scale Synoptic Telescope (LSST; Ivezic et al. 2008b), PanSTARRS (Kaiser et al. 2002; Stubbs et al. 2007), VLT Survey Telescope (VST; Kuijken et al. 2002), the Dark Energy Survey (DES; Tucker et al. 2007), and SkyMapper (Keller et al. 2007; Bessell et al. 2011), will be (or already are being) carried out using filter systems that generally bear close resemblance to the original SDSS system. In this sense, in §3 we carried an intensive study of the properties of physical parameters and metallicities on the SDSS system.

On the other side, the study of galaxy formation and evolution is another of the greatest challenges of modern astrophysics, in this sense, the Milky Way provides us with a unique laboratory to unveil the details of these processes, in particular, our galactic halo is fundamental to the understanding of the evolution of our galaxy due to its remnants of old formation history that can be studied with great detail. The formation of galaxy halos is believed to be due to hierarchical processes driven by the gravitational forces of the large-scale distribution of cold dark matter (e.g., Freeman & Bland-Hawthorn 2002) where monolithic collapse (Eggen et al. 1962) and the accretion of galactic fragments (Searle & Zinn 1978) combine to form galaxies. As shown by simulations (e.g., Johnston et al. 1996; Harding et al. 2001), during these accretion processes some of the galaxy satellites can be tidally disrupted leaving a trail of stars that can be recognized as coherent overdensities in the halo. In §4 we carry an extensive study of the RRab stars in the Catalina surveys, both to characterize and classify them, and to use them as distance estimators to trace our galactic halo in the search of fossil substructure.
Chapter 2

RR Lyrae stars

RR Lyrae (RRL) are periodic variable stars found on the horizontal branch (HB) whose characteristic lightcurves, high brightness, and uniform absolute magnitude make them excellent distance estimators and unique tools to understand the properties of the Milky Way and our Local Group. In the following sections we will discuss some of the basic properties of the RRL.

2.1 Properties

RRL are old (Population II) low-mass and low-metallicity core helium-burning stars that radially pulsate due to opacity-driven mechanisms. Its pulsations are periodic with periods between 0.2 and 1.2 days and amplitudes that range between 0.2 and 2.0 magnitudes in Johnson’s V band. A summary of their physical properties can be seen on Table 2.1. They are giant stars that lie on the HB in a color-magnitude diagram. Their position on the HR diagram is located at the intersection between the HB and the Cepheids instability strip and it was originally known as the “RR Lyrae Gap” (Catelan 2009). This gap, however, is artificial and the lack of stars at this location was due to the variable nature of the RRL that made their positioning on the diagram difficult, since reliable average colors and magnitudes needed a complete
Table 2.1: Properties of RR Lyrae (Smith 1995)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Fe/H]</td>
<td>0.0 to -2.5</td>
</tr>
<tr>
<td>$\log g$</td>
<td>2.5 to 3.0</td>
</tr>
<tr>
<td>$M_V$</td>
<td>0.6 ± 0.2 mags</td>
</tr>
<tr>
<td>$T_{eff}$</td>
<td>6100 K to 7400 K</td>
</tr>
<tr>
<td>Amplitude</td>
<td>0.2 to 2.0 mags in V</td>
</tr>
<tr>
<td>Mass</td>
<td>$\sim 0.7 M_\odot$</td>
</tr>
<tr>
<td>Period</td>
<td>0.2 to 1.2 days</td>
</tr>
<tr>
<td>Radius</td>
<td>4 to 6 $R_\odot$</td>
</tr>
</tbody>
</table>

mapping of their pulsation cycle.

The RRL were first identified as radially pulsating stars by Shapley (1914); however, their classification based on their lightcurve shape (a, b, and c) was originally proposed by Bailey & Pickering (1902). Today, we identify types a and b with fundamental-mode pulsators, and type c as first-overtone pulsators (Catelan 2009); accordingly, types a and b were joined in a single type ab, introduced first by Schwarzschild (1940). Other types of variability among the RRL are known; however, they are much more difficult to recognize and less frequent compared to RRab and RRc. There are double-mode pulsators, or RRd, first identified by Jerzykiewicz & Wenzel (1977), that pulsate at the fundamental mode and first overtone simultaneously. Also, a large period modulation that range between 5.309 days and 530 days of the lightcurve has been observed (Catelan 2009 and references therein), an effect known as the Blazhko effect, first identified by Blazhko (1907), which has been related to non-radial pulsation modes (e.g., Kovács 1995; Kolenberg 2002; Gruberbauer et al. 2007 and references therein). In Figure 2.1 we show example lightcurves of the RRab and RRc types.

2.2 Variability mechanisms

As stated before, RRL are intrinsically variable stars. Their variability comes from a thermodynamic heat valve mechanism first proposed by Eddington (1926) and known as the “Eddington valve”. Basically,
this mechanism consists on something that puts heat into the star at some location when it is at full compression and releases it when at full expansion. Eddington proposed that this mechanism may be an opacity-driven valve (Smith 1995) whose location has been related to the outer envelope of the star where helium is doubly ionized (Zhevakin 1953). The location of the valve is closely related to the Rosseland mean opacity, defined as:

$$\kappa = \kappa_0 \rho^s T^{-\delta},$$  \hspace{1cm} (2.2.1)

where $\rho$ is the density and $T$ the temperature. As explained by De Lee (2008), for an ionized region the value of “$s$” is close to 0, making the opacity primarily a function of the density providing maximum opacity at maximum compression, exactly what is needed to drive an Eddington valve. The value of “$s$” in the ionized zone is what locates the position of the valve, since outside the ionization zone its values are positive ($s \sim 3.5$), generating a strong dependence of the opacity with the temperature, which is also high at full compression. As a result, this generates low opacities away from the ionization region, positioning the valve at the ionization region.

The previously explained mechanism is known as the “$\kappa$ mechanism” (King & Cox 1968), since it is driven by the Rosseland mean opacity. This mechanism is what primarily drives the pulsations of the RRL stars. However, there is a second mechanism that significantly drives their pulsations, known as
Figure 2.2: Spectrum change during the pulsation cycle of an RRab star (RW Dra). Maximum to minimum light goes from top to bottom. We can clearly see the change of the strengths of the lines. Note that the change of the Ca II K line is significantly smaller than the change of the other lines (Smith 1995)

the “γ mechanism” (King & Cox 1968). Basically, this mechanism takes into account the absorption of heat by the ionizing matter. This effect will leave this zone cooler than its surrounding by taking some of the energy that normally raises the temperature of the region to ionize the helium. As a result, this layer will tend to absorb more heat during compression leading to a pressure maximum at minimum volume, providing a force to drive the pulsations King & Cox (1968).
2.3 Metallicities and Distances

The fact that RRL stars lie on the HB on a short color range makes them excellent standard candles; there is, however, a dependence of the absolute magnitude of the RRL with its metallicity (Di Criscienzo et al. 2004). Usually, the dependence of the absolute magnitude with the metallicity is approximated in a linear form:

\[ M_V = \alpha [\text{Fe/H}] + \beta, \]  

(2.3.1)

where \( \alpha \sim 0.23, \beta \sim 0.928 \) (Chaboyer 1999), [Fe/H] the metallicity, and \( M_V \) is the absolute magnitude of the RRL in the Johnson’s V band.

This dependence makes the metallicity one of the most important parameters of the RRL, yet, fairly difficult to obtain. The intrinsic variable nature of the RRL brings along a change of temperature which changes the strengths of the absorption lines in the spectra during its pulsation cycle, as can be clearly seen in figure 2.2. Also, to determine spectroscopic metallicities complete phase coverage is needed, in order to know at what stage of the pulsation cycle the spectra are going to be taken. The usual method to obtain spectroscopic metallicities of RRL involves the measure of the strength of the CaII K line instead of iron lines, and use relations between [Ca/H] and [Fe/H] to obtain the metallicity.

Since spectra are not frequently available in surveys, where we can mine for RRL stars, several alternatives have been proposed to obtain the metallicities from photometric information. Of particular interest is the method proposed by Jurcsik & Kovács (1996), where metallicities are related to the shape of the lightcurve of the RRL. This method is based on a Fourier decomposition of the phased lightcurve of the star, and uses its parameters to obtain the metallicity. In particular, they found the following relation:

\[ [\text{Fe/H}] = -5.038 - 5.394p + 1.345f_{31}, \]  

(2.3.2)
where \( P \) is the period in days and \( \phi_{31} = \phi_3 - 3\phi_1 \) is a parameter that relates the phases of the third and first harmonics of the Fourier series. This relation works fairly well with errors that rose \( \sigma \sim 0.1 \text{dex} \) (Jurcsik & Kovács 1996) for RRL with a well-sampled V-band lightcurve; however, this method fails to correctly recover the metallicity for lightcurves in other bands, and significantly increases the error for RRab with low phase coverage unless template fitting procedures are used to complete the lightcurve (Kovács & Kupi 2007). One of the aims of this work is to develop a method to obtain the metallicity in these cases, in particular in the SDSS filter system, that will be the only photometric bandpasses available on several of the next-generation surveys. Basically, this method relates mean colors and magnitudes with the metallicity of the RRL. This will be explained in detail in Chapter 3.

### 2.4 RRL Applications

One of the main motivations of this work is the usefulness of the RRL for astrophysical purposes. In this section we will outline some examples where RRL have been successfully used as astrophysical probes.

As argued before, RRL are old stars and excellent distance estimators. This fact has been used in many works involving the study of our Galactic halo. A good example of this is the work by Layden (1995), that made a comprehensive study of the Galactic structure using RRL. He used a sample of 302 RRab to work out rotational velocities for both disk and halo components of the Milky Way finding a slight prograde motion of the halo (De Lee 2008). This is just an example of how RRL can be used to map the kinematics of the halo, but several other works have taken advantage of RRL to study it, letting us have a high level of understanding of our halo (e.g.; Lee & Carney 1999; Kinman et al. 2007; Sesar et al. 2012). The properties of RRL have even made possible the study of the kinematics of nearby galaxies; for example, in a recent work by Wagner-Kaiser & Sarajedini (2013), the properties of the Large Magellanic Cloud were traced based on OGLE-III RRL.

Another approach that takes advantage of the properties of the RRL is the study of structure and
streams on our halo, in the form of overdensities. A recent work by Drake et al. (2013b) is a good example of this, where RRL were used to trace a stream of stars that reaches over a hundred kpc from the center of the Galaxy. Several works have been using RRL to trace the substructure of the halo (e.g., Ivezić et al. 2000; Vivas et al. 2001; Vivas & Zinn 2006; De Lee et al. 2013), most of these reported structures are associated with tidal debris of tidally disrupted satellites of the Milky Way. A notable example is the Sagittarius dwarf spheroidal (dSph) galaxy (Ibata et al. 1994), and its extense tidal stream that can be traced around most of the halo. Other streams have been observed, and extensively studied using RRL, namely, the Virgo stellar stream, the Piscis overdensity, the Monoceros stream and the Cetus stream (Drake et al. 2013a and references therein).

In the next years, RRL will become increasingly important to carry these studies due to the transient nature and the depth of future surveys. This will let us find a huge amount of RRL, which will let us make an increasingly precise 3D map of our halo. Part of this work aims to do this as we report more than a thousand RRab stars in the southern hemisphere using data from the Catalina surveys and carry substructure studies. These are explained in detail in Chapter 4.
Chapter 3

Physical parameters and metallicities on the SDSS system

It is known that the SDSS system possesses great scientific potential for a variety of science applications, including stellar populations (e.g., Lenz et al. 1998; Helmi et al. 2003; Ivezić et al. 2008a; Lardo et al. 2011; Vickers, Grebel, & Huxor 2012), the behavior of variable stars in general in such filter systems has not yet been as extensively studied as in the case of more traditional filter systems, particularly the Johnson-Cousins system. For instance, no systematic studies of RR Lyrae variability in globular clusters has been carried out in the SDSS system yet, and similarly theoretical analyses of RR Lyrae variable stars in the SDSS system are almost entirely lacking – the studies by Marconi et al. (2006) and Cáceres & Catelan (2008) seemingly being the sole exceptions. In order to extract the maximum amount of information from extensive RR Lyrae databases that are increasingly becoming available in the SDSS (or similar) systems (e.g., Sesar 2011; Sesar et al. 2007, 2010, 2011), more extensive theoretical analyses

Based on Catelan et al. (2013), submitted
are clearly needed.

In this sense, we have started a systematic study, based on theoretical models and synthetic calculations for horizontal branch (HB) stars, to define precise relations that should allow one to calculate distances, reddenings, metallicity, and physical parameters of RR Lyrae stars from SDSS photometric observations. In this sense, in Cáceres & Catelan (2008) it is presented the first detailed calibration of the RR Lyrae period-luminosity (PL) and period-color (PC) relations in the SDSS system.

One shortcoming of the Cáceres & Catelan (2008) PL and PC calibrations is that they require a priori knowledge of the metallicity, which is frequently not available, especially for field stars. Currently this requires either spectroscopic information or Fourier decomposition of V-band light curves, whose parameters have been calibrated in terms of metallicity (Jurcsik & Kovács 1996; Morgan, Wahl, & Wieckhorst 2007). However, as pointed out by Jurcsik & Kovács, such calibrations of Fourier decomposition parameters are not applicable to all RR Lyrae stars, requiring very well-behaved light curves. In addition, exceedingly complete phase coverage is required for the computation of reliable Fourier parameters. In this sense, calibrations that are based solely on the average photometric properties of the RR Lyrae stars are certainly desired.

In this section we extend the Cáceres & Catelan (2008) study, by providing an additional set of analytical expressions that allow one to compute metallicities, luminosities, temperatures, masses, gravities, and radii of RR Lyrae stars, solely on the basis of their average photometric properties. Our numerical experiments have shown that it is possible to derive fairly precise relations in the multi-band SDSS system.

We present in §3.1 the theoretical framework upon which our study is based. We then provide, in §3.2, an analytical expression that allows one to compute logZ values for RR Lyrae stars based solely on their photometrically measured average magnitudes. In §3.3 we provide similar such relations for the derivation of the physical parameters of RR Lyrae variables, including temperature (§3.3.1), luminosity (§3.3.2), surface gravity (§3.3.3), radius (§3.3.4), and mass (§3.3.5). We analyze the impact of an en-
hancement in the helium abundance upon the derived relations in §3.4.1. In §3.4.3 we test the metallicity relation and compared our estimations with spectroscopic observations.

3.1 Models

The HB simulations used in this work follow the same techniques described in Catelan (2004) and Catelan, Pritzl, & Smith (2004), to which the reader is referred for further details and references about the HB synthesis method. Here, we use the same HB simulations as already utilized in the study of the PL and PC relations in the SDSS filter system (Cáceres & Catelan 2008). For ease of reference, in the following few paragraphs we summarize some key information about our sets of models and HB simulations.

We employed four sets of evolutionary tracks to compute our HB simulations. These tracks were computed by Catelan et al. (1998) for $Z = 0.0005$ and $Z = 0.0010$, and by Sweigart & Catelan (1998) for $Z = 0.0020$ and $Z = 0.0060$. The evolutionary tracks assume a main-sequence helium abundance of $Y_{MS} = 23\%$ by mass and scaled-solar compositions. Helium-enhanced tracks were also computed. The mass distribution along the HB in our simulations is represented by a normal deviate with a mass dispersion $\sigma_M = 0.02 M_\odot$. Bolometric corrections given by Girardi et al. (2004) were incorporated to compute the magnitudes and colors in the ugriz photometric system.

The blue edge of the instability strip is given by equation (1) in Caputo et al. (1987), with a shift by $-200$ K. To obtain the red edge, we adopt for the width of the instability strip a value $\Delta \log T_{\text{eff}} = 0.075$. As discussed by Catelan (2004), these choices lead to an improved agreement with more recent theoretical prescriptions and the observations. The computed periods are based on equation (4) in Caputo et al. (1998); therefore, the periods of first-overtone pulsators (RRc's or RR1's) must be “fundamentalized” before they can be compared with our derived equations. This can be achieved by adding 0.128 to the logarithm of the period (see, e.g., Catelan 2009).

The total number of stars in our simulations for $Y_{MS} = 0.23$, including all metallicities, is 423,765.
This is the star sample that we use, in the next sections, to search for relations involving metallicity and the physical parameters of RR Lyrae stars as a function of quantities that can be directly measured from photometric observations, including magnitudes, colors, and periods, in the SDSS filter system.

In the present work we will make use of two pseudocolors, which are defined as follows:

\[
C_0 = (u - g)_0 - (g - r)_0, \quad (3.1.1)
\]

\[
m_0 = (g - r)_0 - (r - i)_0. \quad (3.1.2)
\]

These indices are patterned after the well-known gravity and metallicity indices of the \textit{Strömgren} (1963) systems. Similarly to what happens in the \textit{Strömgren} system, both of these indices are fairly insensitive to reddening, with \(E(C_1) = -0.32E(B-V)\) and \(E(m_1) = -0.38E(B-V)\), where the “1” subscript indicates reddened quantities. Cáceres & Catelan (2008) already showed that the introduction of \(C_0\) allows one to compute very precise relationships for the derivations of absolute magnitudes and colors in the SDSS system; here we show that \(m_0\) is also very helpful, as far as derivation of metallicity and the remaining physical parameters of RR Lyrae stars goes (see also Catelan et al. 2013).

### 3.2 Metallicity

As we have seen, equation 3.1.2 defines our SDSS “metal-line index.” We have searched for correlations involving the several photometric parameters provided by our simulations, including \(m_0\) and \(C_0\), and the input values of \(\log Z\). We found that the following, fairly simple, fit allows one to obtain metallicity

---

Figure 3.1: Residuals around the fits provided by equation (3.2.1) for a random sample of 5000 stars in our simulations. Synthetic stars with \( Z = 0.005 \) are shown as red squares; those with \( Z = 0.001 \) as orange circles; green triangles are used for the stars with \( Z = 0.002 \); and, finally, blue inverted triangles indicate synthetic stars with \( Z = 0.006 \). The standard deviation (\( \sigma \)) of the residuals is also shown. Clearly, our expression allows the calculation of photometric metallicities to within an internal (1\( \sigma \)) precision that is better than \( \sim 0.1 \) dex.

values to within an internal precision that is better than 0.1 dex:

\[
[\text{Fe/H}] = a[\text{Fe/H}]_0 + b m_0 (\log P)^2 + c m_0 [\text{Fe/H}]_0^2 + d,
\]

(3.2.1)

with

\[
[\text{Fe/H}]_0 = e + f \frac{C_{\alpha} m_0}{(u - g)_0},
\]

(3.2.2)

where the (fundamentalized) periods are given in days. The coefficients of the fits, along with goodness-
of-fit diagnostics (correlation coefficient $r$ and standard deviation of the residuals ($\sigma$), are provided in Table 3.1.

In Figure 3.1 we show the residuals between the metallicity values implied by this fit and the input values in the HB simulations. As can be seen, the analytical fit provides $[\text{Fe/H}]$ values that are correct at the $\pm 0.08$ dex level ($2\sigma$).

In order to go from $\log Z$ values to the more directly measurable $[\text{Fe/H}]$, we used the following equation, based on the work by Salaris, Chieffi, & Straniero (1993) (see Cortés & Catelan 2008 for a discussion):

$$\log Z = [\text{Fe/H}] + \log(0.638f + 0.362) - 1.765,$$

where $f = 10^{[\alpha/\text{Fe}]}$. In this particular case, the transformation is given as follows:

$$\log Z = [\text{Fe/H}] - 1.5515.$$  

These expressions should be used with due care for metallicities $Z > 0.003$ (VandenBerg et al. 2000).
3.3 Physical Parameters

3.3.1 Temperature

Our derived best fits for the RR Lyrae equilibrium temperatures take on the following form:

Figure 3.2: As in Figure 3.1, but for the effective temperature (eq. [3.3.1]) using \((u-g)_0\) (top), \((g-r)_0\) (middle top), \((r-i)_0\) (middle bottom) and \((i-z)_0\) (bottom).
\[
\log T_{\text{eff}} = a (\text{Col})_0 + b m_0^2 + c m_0 + d (\text{Col})_0 (\log P)^2 + e
\]  

(3.3.1)

where the (fundamentalized) periods are given in days, and \((\text{Col})_0\) represents any of \((u-g)_0\), \((g-r)_0\), \((r-i)_0\), \((i-z)_0\). The coefficients of the fits are provided in Table 3.2. Residuals around these fits are provided in Figure 3.2. From this point in our discussion onwards, we will find that some of the fits are so remarkably good that the correlation coefficients, when rounded up to four decimal places, become in fact equal to unity.

For reference, note that the standard errors of the fits imply temperatures, close to the center of the instability strip, that are internally precise to within 6 to 12 K – which is indeed quite remarkable, considering the relatively simple form of the equation adopted in these fits. More in general, these relationships allow one to estimate temperatures that have an internal precision at the level of \( \pm 0.0003 \) (1σ) in \( \log T_{\text{eff}} \).

In practice, equation 3.3.1 requires a priori knowledge of the unreddened colors and pseudo-colors, and thus of the foreground reddening. While the latter can be independently estimated using reddening maps and other techniques, reddening values can also be computed by minimizing the difference between reddened and unreddened RR Lyrae colors, with the latter being provided by the Cáceres & Catelan (2008) expressions. These expressions do require prior knowledge of the metallicity, but this can be obtained by means of equation 3.2.3 (see also eq. [3.5.1] below).

### 3.3.2 Luminosity

Our derived best fits for the RR Lyrae equilibrium luminosities take on the following form:

\[
\log \left( \frac{L}{L_\odot} \right) = a \log P(Mag)^2 + b(Mag) + c(\text{Col})_0 + d \log P \\
+ e(\text{Col})_0 \log P + f m_0^2 + g,
\]

(3.3.2)
with the (fundamentalized) periods given in days, \((\text{Col})_0\) representing any of \((u - g)_0\), \((g - r)_0\), \((r - i)_0\), \((i - z)_0\), and \(\text{Mag}\) the absolute magnitude associated with any of \(u\), \(g\), \(r\), and \(i\). The coefficients of the fits are provided in Table 3.3. Residuals around these fits are provided in Figure 3.3. As in the case of temperature, we again find that luminosities can be derived, on the basis of such a simple fit, with very high internal precision, i.e., typically to within around ±0.0012 dex (1σ).

Note that, in practice, equation 3.3.2 requires a priori knowledge of the absolute magnitude. Again, the latter can be obtained by means of the relations provided in Cáceres & Catelan (2008), using the metallicities provided by equation 3.2.3 (see also eq. [3.5.1] below).

### 3.3.3 Gravity

As far as \(\log g\) is concerned, our derived best fits have the following form:

\[
\log g = a \log P + b m_0 + c (\text{Col})_0 + d, \tag{3.3.3}
\]

where \((\text{Col})_0\) represents any of the SDSS colors as used in §3.3.1, and the (fundamentalized) periods are again given in days. The coefficients of the fits are provided in Table 3.4. Residuals around the fits are
Figure 3.3: As in Figure 3.1, but for the luminosity (eq. [3.3.2]). From top to bottom, the fits are obtained using $(u-g)_0$ and $u$, $(g-r)_0$ and $g$, $(r-i)_0$ and $r$, and $(i-z)_0$ and $i$, respectively.

provided in Figure 3.4, showing that these relations typically allow one to derive log $g$ values that are (internally) correct to within $\pm 0.001$ dex (1$\sigma$).
<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u - g)</td>
<td>(r = 1.0000, σ = 0.00056)</td>
<td></td>
<td>(g - r)</td>
<td>(r = 0.9997, σ = 0.00185)</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>+0.00912</td>
<td>0.00077</td>
<td>a</td>
<td>+0.07933</td>
<td>0.00693</td>
</tr>
<tr>
<td>b</td>
<td>-0.40223</td>
<td>0.00067</td>
<td>b</td>
<td>-0.49001</td>
<td>0.00841</td>
</tr>
<tr>
<td>c</td>
<td>+0.08114</td>
<td>0.00074</td>
<td>c</td>
<td>+1.03816</td>
<td>0.02323</td>
</tr>
<tr>
<td>d</td>
<td>+0.10903</td>
<td>0.00126</td>
<td>d</td>
<td>-0.39673</td>
<td>0.01592</td>
</tr>
<tr>
<td>e</td>
<td>-0.10287</td>
<td>0.00816</td>
<td>e</td>
<td>+0.37909</td>
<td>0.01137</td>
</tr>
<tr>
<td>f</td>
<td>+1.94904</td>
<td>0.00853</td>
<td>f</td>
<td>-1.20616</td>
<td>0.01471</td>
</tr>
<tr>
<td>g</td>
<td>+2.26128</td>
<td>0.00945</td>
<td>g</td>
<td>+1.78504</td>
<td>0.00322</td>
</tr>
</tbody>
</table>

Table 3.3: Luminosity, Coefficients of the Fits

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u - g)</td>
<td>(r = 1.0000, σ = 0.000127)</td>
<td></td>
<td>(g - r)</td>
<td>(r = 1.0000, σ = 0.00108)</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>-1.20658</td>
<td>0.00102</td>
<td>a</td>
<td>-1.22714</td>
<td>0.00110</td>
</tr>
<tr>
<td>b</td>
<td>+0.07381</td>
<td>0.00269</td>
<td>b</td>
<td>-0.47173</td>
<td>0.00135</td>
</tr>
<tr>
<td>c</td>
<td>+0.66240</td>
<td>0.00334</td>
<td>c</td>
<td>+0.34033</td>
<td>0.00673</td>
</tr>
<tr>
<td>d</td>
<td>-0.27097</td>
<td>0.00590</td>
<td>d</td>
<td>-0.22430</td>
<td>0.00385</td>
</tr>
<tr>
<td>e</td>
<td>+0.69907</td>
<td>0.01017</td>
<td>e</td>
<td>+1.34637</td>
<td>0.01724</td>
</tr>
<tr>
<td>f</td>
<td>-0.41283</td>
<td>0.02463</td>
<td>f</td>
<td>-0.41351</td>
<td>0.01327</td>
</tr>
<tr>
<td>g</td>
<td>+1.85288</td>
<td>0.00101</td>
<td>g</td>
<td>+1.86453</td>
<td>0.00664</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r - i)</td>
<td>(r = 0.9999, σ = 0.00116)</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>-1.22423</td>
<td>0.00096</td>
</tr>
<tr>
<td>b</td>
<td>-0.12566</td>
<td>0.00184</td>
</tr>
<tr>
<td>c</td>
<td>+0.24050</td>
<td>0.00393</td>
</tr>
<tr>
<td>d</td>
<td>+2.49361</td>
<td>0.00045</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i - z)</td>
<td>(r = 1.0000, σ = 0.00088)</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>-1.23423</td>
<td>0.00096</td>
</tr>
<tr>
<td>b</td>
<td>-0.12412</td>
<td>0.00184</td>
</tr>
<tr>
<td>c</td>
<td>+0.44603</td>
<td>0.00343</td>
</tr>
<tr>
<td>d</td>
<td>+2.50844</td>
<td>0.00017</td>
</tr>
</tbody>
</table>

Table 3.4: Gravity, Coefficients of the Fits
Figure 3.4: As in Figure 3.1, but for the log of the gravity (eq. [3.3.3]). From top to bottom, the fits are obtained using $(u-g)_0$, $(g-r)_0$, $(r-i)_0$, and $(i-z)_0$, respectively.
Figure 3.5: As in Figure 3.1, but for the radii, obtained by combining equations (3.3.1), (3.3.2), and (3.3.6). The order of the panels is the same as in Figure 3.4.

3.3.4 Radius

One can obtain the equilibrium stellar radius from the derived fit for temperature and luminosity directly for the Stefan-Boltzmann equation:

\[ L = 4\pi R^2 \sigma T_{\text{eff}}^4 \]  

(3.3.4)
where $\sigma = 5.67 \times 10^{-8}$ Wm$^{-2}$K$^{-4}$ is the Stefan-Boltzmann constant. This leads to a simple relation that can be expressed in terms of our previous relations as follows:

$$\log \left( \frac{R}{R_\odot} \right) = \frac{1}{2} \left[ \log \left( \frac{L}{L_\odot} \right) - 4\log T_{\text{eff}} + 15.049 \right].$$

(3.3.5)

We have indeed attempted to obtain independent analytical fits for the radius, but were not able to find analytical expressions that led to significantly improved radii, compared to those obtained by combining equations (3.3.1), (3.3.2), and (3.3.4) – and accordingly recommend that these equations be used, whenever equilibrium RR Lyrae radii need to be computed. This procedure leads to the residuals that are shown in Figure 3.5, revealing that RR Lyrae radii can indeed be computed in this way with very high (internal) precision, i.e., to within ±0.00018 dex in $\log(R/R_\odot)$, or ±0.0027 $R_\odot$.

### 3.3.5 Mass

As was the case with radii (see previous subsection), RR Lyrae mass values can also be obtained by means of our previously derived fits for surface gravity, luminosity, and temperature. We did also attempt to find specific analytical expressions for the mass, but the candidate relations were found to be less precise and robust. We accordingly recommend making use of our derived expressions for the stellar gravity (eq. [3.3.3]) and radius (eq. [3.3.5]), along with the definition of gravity, to estimate the mass; thus

$$M = \frac{gR^2}{G},$$

(3.3.6)

where $G = 6.67 \times 10^{-11}$ m$^3$kg$^{-1}$s$^{-2}$ is the Newtonian gravitation constant. Using equation 3.3.6, we can write a simple form for the mass in terms of the provided equations, in logarithmic form:

$$\log \left( \frac{M}{M_\odot} \right) = \log g + 2\log R - 4.43892.$$  

(3.3.7)
Figure 3.6: As in Figure 3.1, but for the masses (eq. [3.3.7]), obtained by combining equations (3.3.3), (3.3.4), and (3.3.5)

The residuals of the mass values estimated in this way are shown in Figure 3.6, revealing that RR Lyrae masses can also be computed in this way with high (internal) precision, i.e., to within about $\sigma = 0.0017 M_\odot$. 

26
3.4 Robustness of the relations

The relations presented in this work were obtained for a specific set of HB models (Catelan et al. 1998; Sweigart & Catelan 1998) covering a relatively limited metallicity range \(Z = 5 \times 10^{-4} - 2 \times 10^{-3}\), and for a specific value of the helium abundance \(Y_{\text{MS}} = 0.23\). In this section, we analyze how robust the derived relations are, in regard to different \(Y\) values, metallicity values outside the input range, and alternative theoretical models. We also compare the derived metallicities with those obtained from a set of low-resolution spectra from De Lee (2008).

3.4.1 Impact of a Helium Enhancement

There have been many suggestions in the recent literature that globular clusters may host multiple populations with different levels of He enhancement (see, e.g., Valcarce & Catelan 2011; Gratton, Carretta, & Bragaglia 2012, for recent reviews and extensive references). As recently pointed out by, e.g., Kunder et al. (2011, 2013), He-enhanced RR Lyrae stars might thus exist, at least in principle, in some globular clusters (see also Catelan 2009, and references therein). Therefore, it is important to check whether an enhancement in the helium abundance may impact the derived metallicities and physical parameters in an important way.

We thus computed extensive simulations for a helium abundance \(Y_{\text{MS}} = 0.28\) and a metallicity \(Z = 0.002\), and applied the fits provided in the previous sections – obtained for a \(Y_{\text{MS}} = 0.23\) – to the resulting synthetic photometry. The resulting residuals, for a subsample of 5000 stars randomly drawn from the enhanced-\(Y\) simulations, are shown in Figure 3.7.

This figure immediately reveals that all quantities studied in this work – namely, metallicities, temperatures, luminosities, gravities, masses, and radii – can be obtained, even for stars with enhanced

Note, on the other hand, that there is at present no clear evidence that He-enhanced RR Lyrae stars may be present in \(\omega\) Centauri (NGC 5139; Sollima et al. 2006; Marconi et al. 2011).
Figure 3.7: *High-helium test*: Residuals around the fits are shown for 5,000 randomly selected synthetic stars computed with an enhanced helium abundance, $Y_{\text{MS}} = 0.28$, and a metallicity $Z = 0.002$. Shown in the upper left are the log $Z$ values; in the middle top, the luminosities; in the upper right, the temperatures; in the bottom left, the gravities; in the middle bottom, the masses; and, finally, in the bottom right, the radii. For each case, the standard errors ($\sigma$) are provided as insets. Clearly, an increase in $Y$ does not greatly affect the results obtained on the basis of our equations (3.2.3) through (3.3.7) in a particularly strong way.

helium, with a fairly high precision, systematic effects brought about by the high helium abundance been clearly of secondary importance.
3.4.2 Tests Using Alternative HB Models

As already pointed out, the relations derived in §3.2 and §3.3 are based on the HB models of Catelan et al. (1998) and Sweigart & Catelan (1998). To evaluate the robustness of these relations, and give us a better sense of possible systematic effects stemming from the assumptions in the underlying HB models, we have computed additional sets of HB simulations, using completely independent HB models as input. In particular, we computed HB simulations using evolutionary tracks from Pietrinferni et al. (2004). This includes the following \((Y_{MS}, Z)\) combinations: \((0.245, 0.0001); (0.245, 0.0003); (0.248, 0.002); (0.259, 0.01)\). These cases will be referred to in the following discussion as C1 to C4, respectively. These additional simulations cover a range in HB morphologies, from very red to very blue. The results are discussed in the next sub-sections. In addition, we also check the \(z\)-band period-luminosity relation from Cáceres & Catelan (2008), since absolute magnitudes are required for the computation of the luminosity (see discussion in §3.3.2), and, therefore, also for the radius and the mass.

A couple of additional simulations was also computed, to check the behavior of the simulations in an “extreme” case, namely that of solar metallicity. We consider both Castellani, Chieffi, & Pulone (1991) and Sweigart & Catelan (1998) models for \(Z = 0.02\). These will be referred to as the “E1” and “E2” cases, respectively, in what follows.

3.4.2.1 Metallicity

Application of equation 3.2.1 to the synthetic photometric dataset derived from the Pietrinferni et al. (2004) simulations gives discrepancies between predicted and “observed” \(\log Z\) values that are typically lower than 0.1 dex, the situation however becoming noticeably poorer only at the extremely metal-poor end. More specifically, the average deviations that we have found are as follows: \(-0.31 \pm 0.09\) dex (C1); \(-0.07 \pm 0.08\) dex (C2); \(+0.02 \pm 0.05\) dex (C3); \(-0.01 \pm 0.08\) dex (C4). Application of the analytical relation to cases E1 and E2 led to average deviations of the order of \(+0.07 \pm 0.14\) dex and \(+0.05 \pm 0.13\) dex, respectively. These results suggest that equation 3.2.1 may underestimate the true metallicity.
at the more extreme metal-poor (Z = 0.0001) end, by about 3 tenths of a dex – and so we recommend using equation 3.2.1 with due care, at the extreme metal-poor end.

### 3.4.2.2 Temperature

Application of equation 3.3.1 to the synthetic photometric dataset derived from the Pietrinferni et al. (2004) simulations gives discrepancies between predicted and “observed” log\(T_{\text{eff}}\) values that are remarkably small, with (average) differences appearing at the level of the fourth decimal place only. This is true for both cases E1 and E2 as well. The sole exception to this behaviour is the relation involving \((u - g)_0\), where the differences increase to \(\sim 0.001 \pm 0.001\) dex. These differences become higher in cases E1 and E2, where they essentially double in size.

### 3.4.2.3 Luminosity

Application of equation 3.3.2 to the synthetic photometric dataset derived from the Pietrinferni et al. (2004) simulations gives discrepancies between predicted and “observed” log\(L\) values that are very small, typically at the level of \(\delta \log(L/L_0) = +0.015 \pm 0.003\). There is no significant loss in performance at solar metallicity (cases E1 and E2).

### 3.4.2.4 Gravity

Application of equation 3.3.3 to the synthetic photometric dataset derived from the Pietrinferni et al. (2004) simulations gives discrepancies between predicted and “observed” log\(g\) values that are again quite small on average, the differences being at the level of 0.008 \(\pm 0.003\). There is again no significant loss in precision at solar metallicity (cases E1 and E2).
Figure 3.8: Comparison between our photometric metallicity estimates and the spectroscopic measurements by De Lee (2008), in the \((m_g, [Fe/H]_{phot})\) plane. The level of agreement is denoted by the color code that is shown in the upper right, in standard deviations. The box shows the cuts adopted in this comparison for the removal of obvious outliers.

3.4.3 Spectroscopic Metallicities

In the framework of the SDSS survey, the central stripe (#82) in the South Galactic Cap was mapped repeatedly. The RR Lyrae stars detected in this strip were the subject of the Ph.D. thesis of De Lee (2008), and will be described in detail elsewhere (De Lee et al. 2013, in preparation). In this section, we use the light curves for a sample of 74 RR Lyrae stars located in this region for which spectroscopic metallicity estimates were derived in De Lee (2008). These measurements are based on spectra obtained with the medium-resolution R-C Spectrograph (RCSPEC) on the Blanco 4m telescope.

Among these 74 RR Lyrae stars, 61 are fundamental-mode variables, whereas 13 are first-overtone pulsators. These stars were extracted from the Light Motion Curve Catalogue (LMCC), a catalogue
containing variable objects in the SDSS strip 82 (see, e.g., Bramich et al. 2008). The RR Lyrae star selection in the LMCC was made by De Lee (2008), who also provided the unreddened average colors and periods for the stars in the sample. The reddening indices were obtained by De Lee (2008) using the High Level Catalog by Bramich et al.. With this photometric dataset, we estimated each star’s metallicity using equation 3.2.3. The results were then compared with De Lee’s spectroscopic measurements. Note that the precision of these spectroscopic metallicities is quite low, at the level of ±0.4 dex, and so the following conclusions should be taken with some caution.

To avoid obvious outliers values and to define a range of validity of our relations, we first perform two cuts, one in the predicted metallicity and the other in \( m_0 \), in the sense that we only retain RR Lyrae stars with 

\[-3 < [\mathrm{Fe/H}]_{\text{sim}} < 0 \quad \text{and} \quad 0.009 < \langle m_0 \rangle < 0.3\]  

– i.e., covering a similar range in these parameters as predicted by our simulations. In Figure 3.8 we show the cuts and how they help remove the most obvious outliers.

This figure confirms that the majority of the stars (i.e., 21 out of 74, or about 28%) outside the “compatibility region” in the \([\mathrm{Fe/H}]_{\text{sim}}, m_0\) plane have predicted metallicities that fall more than 3\( \sigma \) from the spectroscopically favored metallicity. Of the remaining 53 stars, \( \approx 61\% \) agree to within the 1\( \sigma \) level, \( \approx 87\% \) within 2\( \sigma \), and \( \approx 95\% \) within 3\( \sigma \). There are only 2 stars (4% of the sample) which are > 3\( \sigma \) outliers, and for which an obvious explanation could not be found. Obviously, more accurate metallicity measurements are needed for a more meaningful comparison and tests of our calibrations.

### 3.4.4 Internal Robustness of the Relations

In order to further investigate the degree of reliability of our relations in the face of noisy data, we ran several Monte Carlo simulations. More specifically, we added random, non-correlated errors to the simulated, “error-free” magnitudes, to emulate errors in the empirical process of estimating the mean magnitude from actual RR Lyrae light curves. We then used these modified magnitude values in conjunction with our relations to predict the physical parameters and metallicity. Finally, we studied the
resulting precision, based on these new derived values.

Typically, by adding a typical error of 0.01 mag to the mean magnitude, the precision level dropped down by a factor of 5 to 10. While this may at first appear like a very large downgrade in performance, it does not pose a great risk in practice, given the extremely high precision of the original expressions (see §3.3). Proceeding in this way, we conclude that all physical parameters are safely obtained on the basis of our expressions, except when the errors in the mean magnitudes exceed the ~ 0.05 mag level. The only exceptions to this statement are provided by the relations for the metallicity and mass, which should be used with care whenever the errors in the mean magnitudes become higher than the ~ 0.01 mag level.

In this sense, it is important to remind the reader that all magnitudes and colors provided in this work correspond to the equivalent static star. In general, average magnitudes and colors are not necessarily equal to the equivalent static quantities. It is beyond the scope of this work to provide a recipe for computing such quantities from the observations; indeed, to the best of our knowledge, a systematic analysis of this problem is currently lacking in the literature, particularly in regard to the SDSS system. In this sense, Bono, Caputo, & Stellingwerf (1995) provided a systematic analysis of the problem, but only in the case of BVK photometric observations. According to their results, neither intensity- nor magnitude-averaged magnitudes and colors can be directly used in lieu of the equivalent static quantities, but amplitude-dependent corrections must be used. An extension of this analysis to the SDSS filter system is accordingly strongly encouraged.

3.5 Alternative Metallicity Model

Some of the surveys mentioned in §3 do not possess a u-band filter. In order to increase the applicability of our models, we have added an alternative fit to the metallicity that does not require u-band magnitudes. The relation takes on the following form:
Figure 3.9: As in Figure 3.1, but for the alternative metallicity relation. The standard deviation ($\sigma$) of the residuals is also shown. Our expressions allow the calculation of photometric metallicities to within an internal (1$\sigma$) precision that is around than $\sim$ 0.2 dex.

\[
[\text{Fe/H}] = a(g - r)_0 (\log P)^2 + b \frac{[\text{Fe/H}]_{01}}{(\log g)^2} + c m_0^2 [\text{Fe/H}]_{03} + d m_0 [\text{Fe/H}]_{01}^2 + e, \tag{3.5.1}
\]

where

\[
[\text{Fe/H}]_{01} = f + g (\log g)^3 m_0, \tag{3.5.2}
\]

\[
[\text{Fe/H}]_{03} = h + i m_0^2 \log P, \tag{3.5.3}
\]

with the (fundamentalized) periods given in days. The coefficients of the fits, along with goodness-of-fit diagnostics, are provided in Table 3.5. The values of $\log g$ are obtained from our relations using equation 3.3.3.
While this relation has a lower performance by about a factor of two compared with equation 3.2.3, its internal precision is still better than $\sim 0.2$ dex, at the $1\sigma$ level.

We ran the same robustness tests mentioned in §3.4. As a result, we find that equation 3.5.1 is still able to reproduce the results of the enhanced-$Y$ simulations, to within $\sigma \sim 0.3$. A similar behaviour as previously found for the C1 – C4 and E1 – E2 models is also obtained, with the main difference being that the metallicity is now underestimated, at the extreme metal-poor end, by about 0.5 dex.

We next compared with spectroscopic metallicities from De Lee (2008), obtaining encouraging results. More specifically, 24 stars are discarded at the outset by the “compatibility test” described in §3.4.3 – but, of the remaining stars, $\sim 66\%$ agree to within the $1\sigma$ level, $\sim 94\%$ within $2\sigma$, and $100\%$ within $3\sigma$. Finally, the internal robustness test (§3.4.4) applied to this relation revealed a similar performance as found for equation 3.2.1.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>+50.6595</td>
<td>5.91673</td>
</tr>
<tr>
<td>$b$</td>
<td>-4.55467</td>
<td>0.94617</td>
</tr>
<tr>
<td>$c$</td>
<td>-8.09571</td>
<td>1.31580</td>
</tr>
<tr>
<td>$d$</td>
<td>-4.56199</td>
<td>0.35442</td>
</tr>
<tr>
<td>$e$</td>
<td>-1.56952</td>
<td>0.17520</td>
</tr>
<tr>
<td>$f$</td>
<td>-2.84249</td>
<td>0.04020</td>
</tr>
<tr>
<td>$g$</td>
<td>+0.40045</td>
<td>0.01228</td>
</tr>
<tr>
<td>$h$</td>
<td>-2.11945</td>
<td>0.00937</td>
</tr>
<tr>
<td>$i$</td>
<td>-117.653</td>
<td>2.12260</td>
</tr>
</tbody>
</table>

Table 3.5: Metallicity, Coefficients of the Fits
Chapter 4

Substructure and algorithms on the southern hemisphere

Numerous tidal streams and dwarf galaxies have been discovered within the Galactic halo in the last years (e.g., Vivas et al. 2001; Grillmair 2006; Drake et al. 2013b). The most studied of these overdensities is the Sagittarius dwarf spheroidal galaxy (dSph) and its associated tidal stream (Ibata et al. 1994), which is the most prominent feature of the halo and extends completely around our galaxy. In the past years, the number of streams and structures found in the outer Galactic halo (> 15 kpc) has significantly increased (Drake et al. 2013a); however, the number of satellites cannot account for the amount predicted by $\Lambda$CDM models of hierarchical structure formation (e.g., Freeman & Bland-Hawthorn 2002; Bullock et al. 2001). This “missing satellite” problem continues to be important to our understanding of galaxy formation, and therefore further study of substructures in the halo is desired.

For the sake of finding substructure, a good tracer is needed. In this sense, RR Lyrae are fundamental

Based on Torrealba et al. (2013), in preparation
distance probes that can be used to trace the history of galaxy formation (e.g., Catelan 2009 and references therein). We focused on RRL type ab stars because their high brightness ($M_V \sim 0.6$) and characteristic lightcurves (that makes them relatively easy to differentiate from other types of stars) let us put together a relatively clean sample of stars with distances to perform an extensive 3D mapping of the halo.

To date a few tens of thousands of RRL are known in dense regions near the Galactic bulge, where the Sagittarius dSph galaxy is located. However, the Galactic halo itself has only been probed with confirmed RRL over a few thousand square degrees to heliocentric distances of 30 to 100 kpc (e.g., Vivas et al. 2001; Sesar et al. 2010; Drake et al. 2013a). This work is the southern counterpart of the more than ~10,000 RRab that we reported in Drake et al. (2013a). Together, we trace the halo covering more than 34,000 deg$^2$ of the sky, reaching distances in excess of 50 kpc.

In this work we perform a systematic study of the southern hemisphere to characterize its internal structure. The two main objectives are the description of the techniques developed to automate the search, calibration, and classification of RR Lyrae type ab stars and the methodology used to search for substructure in the halo. In §4.1 we introduce the data used in this work, whereas in §4.2 we describe the algorithms developed to find and classify the RRab. The RRab found are presented in §4.4, and in §4.5 we present the methodology used to extract substructure and overdensities.

4.1 The Catalina Surveys

The Catalina Sky Survey began in 2004 and uses three telescopes to cover the sky between declination $-75^\circ < \delta < +65^\circ$ covering more than 33,000 square degrees in order to discover Near-Earth Objects (NEOs) and Potential Hazardous Asteroids (PHAs) (Drake et al. 2013a). Its counterpart, the Catalina Real-time Transient Survey (CRTS), involves the analysis of the same data in order to detect and classify stationary optical transients (Drake et al. 2009). Both surveys work collaboratively to extract the maximum scientific return from the data of the three telescopes operated by the Catalina Sky Surveys. These
consist of the Catalina Schmidt Survey (CSS) telescope, located in Mt. Bigelow in the Catalina Mountains just north of Tucson, Arizona, that observes the region between declinations +70° and −25° using the modified 0.4-/0.6-m Catalina Schmidt telescope. The Mount Lemmon Survey (MLS) telescope, also located in Tucson, Arizona, at an altitude of 2800 m, observes a 10° plane around the ecliptic with a 1.5-m Cassegrain reflector telescope. Finally, the Siding Spring Survey (SSS) telescope in Siding Spring, Australia, uses an Uppsala 0.5-m Schmidt telescope to observe the southern hemisphere between declinations −80° and 0°. All the three sub-surveys avoid the Galactic plane region between declinations of 10° and 15° due to crowdedness, its images are unfiltered to maximize throughput and taken in sequences of four observations separated by ten minutes. Photometry is carried out using the aperture photometry program SExtractor (Bertin & Arnouts 1996).

In this work, we will be using the southern data, trying to extract RRL from it. The data is previously selected to be variable objects with periods between 0.1 and 4 days, according to the Lomb-Scargle period finding method (Lomb 1976; Scargle 1982), and consists of ~20,000 objects with unfiltered magnitudes between 10.5 and 19.2 and ~100 epochs. In Figure 4.1, the sky coverage of the data that we will be working with is shown.

We calibrate the unfiltered Catalina data to the standard Johnson V filter using equation 1 of Drake et al. (2013a):

\[ V = V_{\text{CSS}} + 0.31(B - V)^2 + 0.04. \]  

(4.1.1)

This relation has a dispersion of \( \sigma = 0.059 \). Since in this study we will focus on RRab stars, we will study the impact of color variations over a pulsation cycle from their perspective. If we consider an average color for RRab stars of \((B - V) \sim 0.3\) and their color range of \(0.1 \lesssim (B - V) \lesssim 0.5\) (Guldenschuh et al. 2005), the maximum difference in color will be \(\sim 0.2\) in \((B - V)\). This adds an extra uncertainty of \(\sim 0.05\) magnitudes to eq. 4.1.1, which combined in quadrature with the dispersion of the relation, results in total dispersion of 0.077 in the \(V\) magnitude when taking an average color for RRab stars. We
studied the impact of taking better color estimations by making a simple model for the color variation of the RRab following Figure 2 of Guldenschuh et al. (2005). We used average minimum and maximum colors as found in M3 (NGC 5272) data, and used them to make a differentiated magnitude calibration over the pulsation cycle. With the new lightcurve we found that the changes in the average magnitude and metallicity (quantities that are directly affected by a better color estimation) do not exceed $\sim 0.01$ and $\sim 0.1$, respectively. This is the reason to keep the simplest approximation and use the mean color for the RRab stars.

### 4.2 Automatic Algorithms

To select and classify variables we need to characterize their lightcurves as well as we can. The period is undoubtedly the most important parameter to understand the behaviour of the variable star, which is
why we developed a method to obtain the best period we can from the data, with reasonable computing times. The objective of this process is to classify stars with a better period and characterization of their lightcurves. To address these two objectives we developed the Automatic Fourier Decomposition and the Automated Period Selection routines to characterize their lightcurves and find their periods, respectively. With these two tools we automatically study the lightcurves of all our variable star candidates to extract the RR Lyrae stars from the sample. In section 4.2.1 and 4.2.2 we explain more in-depth these two routines, respectively.

4.2.1 Automatic Fourier Decomposition

The Automatic Fourier Decomposition (AFD) is a method that lets us compute a Fourier decomposition of a periodic sample in a way that the number of harmonics in the series is given by the statistical significance of the fit itself. A similar approach was made by Petersen (1986) where a discriminator was used to find the optimum order of the Fourier series. The main difference lies in the type of discriminator, while we use a statistical test to check if additional orders are worth in terms of explaining the data with the Fourier decomposition, Petersen uses the "unit lag autocorrelation" (Baart 1982) to check for correlation between the residues of the fit. When their statistic determines that there is no trend in the residues, they stop adding orders to the Fourier series. A further analysis is needed to compare the two methods.

To find the parameters of the harmonics we perform a weighted least-squares fitting procedure to a series of Fourier harmonics described by

\[ f(x) = A_0 + \sum_{n=1}^{n_{\text{max}}} A_n \sin \left( \frac{2n\pi t}{P} + \phi_n \right), \]

where \( A_n \) are the amplitudes, \( \phi_n \) the phases, \( P \) the period and \( n_{\text{max}} \) the number of harmonics in the decomposition. To determine the order of the fit, \( n_{\text{max}} \), we perform several fits with increasing order until

40
it is not statistically significant to add an extra order. The process starts with just one harmonic:

\[ f(x) = A_0 + A_1 \sin \left( \frac{2 \pi t}{P} + \phi_1 \right). \]  

(4.2.2)

We perform a least-squares fit to determine \( A_0, A_1, \) and \( \phi_1, \) after which we quantify the goodness of the fit through the reduced \( \chi^2 \) index, calculated as follows:

\[ \chi^2 = \frac{1}{N-n-1} \sum_{k=1}^{N} \left( \frac{O_k - E_k}{\sigma_k} \right)^2, \]

(4.2.3)

where \( N \) is the number of observations, \( n \) the number of parameters of the Fourier decomposition, \( \sigma \) the error of the observation, and \( O \) and \( E \) the observed and the expected value as given by the Fourier series, respectively. We then compare this fit to the following two series with higher orders, this is, with 2 and 3 extra harmonics. To determine if the higher-order series gives us a significant improvement on the fit we perform a statistical F-test to check if the null hypothesis (i.e., the two fits are statistically equal in terms of their performance) can be discarded. The \( F \) statistic is calculated as follows:

\[ F = \left( \frac{\chi^2_1}{\chi^2_2} - 1 \right) \frac{b}{a}, \]

(4.2.4)

where \( b = N - p2 \) is the degrees of freedom of the second fit and \( a = p2 - p1 \) is the difference in parameters between the second and first fits (note that always \( p2 > p1 \)). The test is now compared to a critical value, \( \alpha; \) if \( F \) is less than this critical value, then the null hypothesis cannot be discarded and the two fits are statistically equal in their ability to describe the observed data. To determine \( \alpha, \) we ask that the rejection probability is 0.99, that is, that the null hypothesis is discarded when we are 99% sure that adding a new order to the series is statistically significant.

The process to select the order of the fit is now straightforward: We begin with one order on the Fourier decomposition and we add extra harmonics to the series until we cannot discard the null hypothesis, i.e. the addition of an extra order to the series is not a statistically significant improvement to
Figure 4.2: Three example lightcurves and theirs AFD. From left to right, we got an AFD with two, four, and six harmonics on their Fourier series, respectively explain the observed data. In addition to this, we set the maximum number of orders to 6, to avoid extreme wriggles and over-fitting to outliers. In Figure 4.2 we can see three examples of AFD that achieve different orders to explain the lightcurve.

4.2.2 Automatic Period Selection

The Automatic Period Selection (APS) is an algorithm that we developed to get more precise periods from the period-searching algorithms. In particular, we compare the period proposed by two algorithms, namely Analysis Of Variance (AOV) (Schwarzenberg-Czerny 1989, 1996; Schwarzenberg-Czerny & Beaulieu 2006) and Lomb-Scargle (LS; Lomb 1976; Scargle 1982). The comparison is made by phasing the lightcurves with the five best periods proposed by each routine and calculating a goodness-of-fit parameter to discriminate the period that better fits the data.
Figure 4.3: The impact of the first sigma clipping. The dashed black line represents the AFD without sigma clipping, and the green line represents the AFD with sigma clipping. The impact of the outliers at \( \phi \sim 0.1 \) is clearly seen on the black dashed line.

### 4.2.3 Selection Methodology

The comparison between the two period-search algorithms is performed mainly using the AFD method described in section 4.2.1. The process begins by taking the first five guesses from both LS and AOV and phasing the lightcurve to these periods. We perform an initial AFD to find the mean and the amplitude of the lightcurve. With these two quantities we first remove the most obvious outliers that lie more than 3\( \sigma \) from the mean. With the outliers removed we then perform a second AFD. In Figure 4.3 we can see the first AFD and the second AFD with the outliers removed as well as the sigma-clipping limits.

There are two other optional processes to further clean the lightcurve. The first process is lightcurve clipping. It is important to mention that this process does not work well on deep eclipsing binaries; therefore, by applying this we are aiming to find RR Lyrae variables only. The process consists of
clipping all the dots of the lightcurve that lie more than $3\sigma$ from the recently computed AFD. In this case, $\sigma$ is the mean error of the observations instead of the standard deviation of the full lightcurve (as was the case for the first sigma-clipping). In Figure 4.4 we can see the lightcurve clipping boundaries and how they remove some of the outliers that were kept after the first cut.

The second process is gap filling. This process is triggered when there is a gap of more than 0.1 in phase on the phased lightcurve. When this happens a “data” point is included in the lightcurve based on the previously calculated AFD. The synthetic observation is positioned by means of a simple linear interpolation between the positions of the edges of the gap. These edges are calculated as the mean of the three closest points to the edge at each side in phase, and uses the previous AFD at these phases to find their magnitudes. As seen in Figure 4.5, this process helps remove extreme wriggles from the estimated lightcurve.
Figure 4.5: Same as F4.3, but for gap closing. As with lightcurve clipping, the effect of this procedure is secondary in comparison to the raw sigma clipping process, but removes considerable wriggles from the AFD.

4.3 Sample Selection

All the processes mentioned in §4.2 are used to have cleaner lightcurves; note that after each one of these processes we perform a new AFD to the data, that is used in the next process as input. The importance of a clean lightcurve is that we then use it to estimate the goodness of fit that will be used to find the best period for the candidates; therefore, the better the lightcurve describes the data, the better we will be able to discriminate between the period candidates.

With the final AFD for each of the phased lightcurves based on the period candidates, we then compute reduced $\chi^2$ to estimate the goodness of the fit and rank them based on it. To select the best period we follow a series of rules. First, if the two best period candidates have a difference of more than 10% in reduced $\chi^2$ then the first period is selected. If this is not met, but the two first candidates of AOV
and LS are equal and match one of the two best periods ranked by the reduced $\chi^2$ test, then that period is selected. If neither of the two criteria are met, but the two first periods are aliases (either double or half the period), then, motivated by Schwarzenberg-Czerny (1996), the one ranked higher by AOV is chosen. Finally, if neither of these conditions are met, we flag the star saying that we cannot differentiate between the two best candidates.

After this process we also flag the stars with reduced $\chi^2 > 3$ and those with periods greater than 4 days, since given the pre-selection condition mentioned in §4.1, we do not expect to find stars with periods that long. It is important to note that, to avoid getting biased by higher-order fits, that will have better $\chi^2$ (even after being corrected by degrees of freedom by using the reduced $\chi^2$), the comparison of the goodness of the fit is done with the highest order found by the AFD over the period sample; nevertheless, this is done only to choose the better period, not to select the order of the final AFD.

With this method, we are able to discriminate good periods from bad periods. To test this, we plotted several ($\sim 1000$) phased lightcurves and inspected them by eye. The inspection shows that this method is able to correctly select the periods when both AOV and LS agree, when they do not agree, and even when they both fail at their first guess.

In order to avoid contamination, we have made this method very restrictive. This leaves out of our sample some of the stars with good period estimations by LS and/or AOV, as shown by the eye inspection; nevertheless, this is just a small percentage ($\sim 1\%$) and aside these stars that got good periods from LS and AOV (but not selected by our procedure), we were able to recover about 20% of the stars that with a single method got a wrong period.

### 4.3.1 RRab Selection

With the period and the lightcurve information, we now select the RRab stars from the sample. Before starting with this process we remove all the stars with bad period estimation as flagged by the previously described method. We then introduce a new discriminator motivated by Kinemuchi et al. (2006) that we
called the M-test. This test is defined by the following statistic:

\[
M_f = \frac{M_{\text{max}} - M_{\text{mean}}}{M_{\text{max}} - M_{\text{min}}},
\]

where M stands for magnitude. This test basically measures if the star spends most of the time above \( M_f > 0.5 \) or below \( M_f < 0.5 \) the mean magnitude of the star. To avoid being biased by outliers, we find all the input information to this test from the AFD lightcurve, in contrast to what is done by Kinemuchi et al., where raw data was used to calculate the M-statistic. To see how this test can differentiate between different types of variable stars, we classified \( \sim 150 \) randomly selected stars by eye between RRab, detached eclipsing binaries and RRc/W UMa-type eclipsing binaries. In Figure 4.6 we can see how this test isolates the different types of stars, in particular, RRab. Based on this, we used this test as the main RRab selection criterion.

In Table 4.1 there is a summary of all the cuts that were imposed to select RRab. As for the RRc type, we spotted a possible region in the M-statistic/period space (see figure 4.6) that is mainly populated by this kind of stars, but the contamination with W UMa stars is strong, and, even looking at the lightcurves, it is difficult to identify the difference between these two classes. Accordingly, in order to have the cleanest selection possible, we did not attempt to disentangle RRc’s and W UMa’s, and focused only on RRab stars.

To further clean our sample, we removed all the lightcurves that have more than 1 peak in phase.
Figure 4.6: M-statistic vs period map. This figure clearly shows how this parameter help us to discriminate the RRab stars from the rest of the variables, letting us have a far cleaner sample using the M-statistic.

space. This is achieved by tagging all the stars whose amplitude in the higher harmonics of the AFD is bigger than the amplitude of the first harmonic.

With the final sample ready, we now calculate several parameters for the RRab stars. First, we calculate their metallicity based on Jurcsik & Kovács (1996):

$$[\text{Fe/H}] = -5.038 - 5.394P + 1.345\phi_{31},$$

(4.3.2)

where $P$ is the period in days, $\phi_{31} = \phi_3 - 3\phi_1$, and the metallicity is given in the scale of Jurcsik (1995). It is important to note that all the amplitudes have to be positive and the phases between 0 and $2\pi$. We transform the metallicity of equation 4.3.2 to the standard Zinn & West (1984) scale using equation 4 of
Jurcsik (1995):

\[
[\text{Fe/H}]_{ZW} = \frac{1}{1.431} \left( [\text{Fe/H}]_j - 0.88 \right).
\] (4.3.3)

To check the reliability of these metallicities we calculate the \( D_{111} \) factor (see Jurcsik & Kovács 1996eq. 6). To fully calculate this parameter we need at least 6 orders on the Fourier decomposition. Since we have RRab stars starting from 3 orders in their AFD, we estimated two \( D_{111} \) factors, one using the available orders only (that we called \( D_{np} \)), and another forcing the AFD to have 6 orders, as in the calculations originally made by Jurcsik & Kovács.

With the metallicity we calculate the absolute magnitude of the star, according to the following expression Catelan & Cortés (2008)

\[
M_V = 0.23[\text{Fe/H}]_{ZW} + 0.948.
\] (4.3.4)

Then, by taking reddening values from the Schlegel et al. (1998) reddening maps, we obtain unreddened magnitudes. Finally, by applying the definition of distance modulus, we find the distance, in kpc, from the Sun:

\[
\log D = \frac{1}{5} (m_V - M_V) - 2.
\] (4.3.5)

4.4 Results

From the original sample of 18,288 variable star candidates with period between 0.3 and 4 days, we selected 10,541 stars as RRab variables. In F4.7 we show a sample of the classified stars. All of these stars have derived information for metallicity, distance, Galactic position, and lightcurve shape (through the AFD parameters).
4.4.1 Completeness

To check the completeness of our sample, we generated synthetic lightcurves based on Layden (1998) RRab templates. To make the lightcurves as similar as possible to the ones observed by SSS, we studied how the error of the observations changed with magnitude. In F4.8, we can see how the error of the observations increases with apparent magnitude. We model this behavior using an exponential fit:

$$\sigma_M(M) = 0.04 + 1.84 \times 10^{-3} \exp^{0.55(M-M_{\text{min}})},$$  \hspace{1cm} (4.4.1)

where $M_{\text{min}} = 9.51$ corresponds to the brightest star in our sample. To generate our synthetic RRab we selected a random RRab from our sample, we used its period to phase its MJD observations, and its amplitude to modify the templates so we can have synthetic lightcurves with different amplitudes. Using the modified template, the phases of the RRab, and the errors from equation 4.4.1, we create the synthetic lightcurve.

We generated 100 synthetic RRab per magnitude bin, from 12 to 18 magnitudes. We then ran the APS software over them and looked how many RRab we were able to recover, in order to obtain an estimate of the completeness of the APS software as a function of the apparent magnitude (see Fig. 4.9). This figure shows that we recover at least 60% of the RRab at magnitudes lower than 18. The period efficiency is even higher, reaching $\sim 90\%$. It is important to mention that the APS was set to have the cleanest sample possible in a fully automatic procedure, which is why at the faintest levels the efficiency drops down. This is rather unavoidable since at these magnitudes the error becomes comparable to the amplitudes of the stars, making it very difficult for an automatic procedure to identify correctly their type of variability. Nevertheless, as part of the output of the program all the information of the discarded stars is also included, so that the faint RRab discarded by the procedure can be recovered by making a non-automatic inspection of their lightcurves.

In T4.4.1 we show an extract of our derived RRab catalog and all the final information given by our
Table 4.2: RRab output table sample. The id corresponds to the Catalina id whereas RA and Dec are in degrees. 

- $n_{obs}$ represents the number of observations and $P$ is the period in days. $m_V$ and $M_V$ are the apparent magnitude and the absolute magnitude, respectively. $D_{appl}$ represents the metallicity reliability factor as described above, $E(B-V)$ the reddening factor and $D_0$ the distance in kpc. calculated using reddening-corrected magnitudes. 

4.4.2 Mean Halo Metallicity

As part of the information that we got from the APS, we can estimate the mean metallicity of the halo. We achieve this by making a histogram of all the metallicities obtained from the RRab's using equation 4.3.2. It is important to mention that we set the metallicity to $-1.5$ to all the stars whose metallicity falls beyond $-2.5 < [\text{Fe/H}] < 0.5$, but these were removed from the calculation of the mean. In Figure 4.10, we show two histograms, showing the mean metallicity distribution of the halo. We report a peak mean density of $[\text{Fe/H}] = -1.4 \pm 0.22$. This result is in fairly good agreement with other spectroscopic analyses of halo stars, as summarized for instance by Kinman et al. (2000). A more detailed analysis, including spatial variations, will be carried out in the near future.
4.5 Halo Substructure

We have found a total of 10,541 RRab in the halo with their respective distances. They cover an area on the sky of about 14,800 square degrees and go as deep as 45 kpc from the Sun. There should exist stars beyond this limit, but as argued in §4.4.1, the APS software was set up to get the cleanest sample possible reducing the efficiency at higher distances. We used this sample to study the internal structure of the halo by using the RRab as tracers in an algorithm designed to search for overdensities. To avoid extreme errors in the distance estimation, we removed 224 RRab with $E(B - V) > 0.4$.

The methodology used is based mainly on the local density. To measure it we used the N-th neighbour method, originally proposed by Dressler (1980). This method consists of selecting a position and use the distance to the N-th nearest star as the radius of a sphere containing N stars. The number density at the selected position is then calculated as

$$\rho_l = \frac{N}{\frac{4}{3}\pi d_N^3},$$

(4.5.1)

where $d_N$ is the distance to the N-th star from the selected position. The first step of the method is to calculate the local number density around every star in the sample. In order to identify overdensities, we then compare the local number density to the local density as predicted by halo models. We use the model proposed by Sesar et al. (2010), defined by

$$\rho_M = \rho_0 \left( \frac{R_0}{\sqrt{x^2 + y^2 + (z/q_H)^2}} \right)^{n_H},$$

(4.5.2)

where $\rho_0 = 4.2\text{kpc}^{-3}$ is the number density of RRab stars at $R_0 = 8.0\text{ kpc}$ (Vivas & Zinn 2006), $q_H = 0.64$ indicates the oblateness of the halo (Sesar et al. 2010), $n_H = 2.77$ is the power-law index (Jurić et al. 2008), and $x$, $y$, and $z$ are coordinates in the galactocentric system.

To make a more precise study of the overdensities and to characterize them individually as overdensities and not as stars, we calculate the local density in a mesh around the stars that has a significant
higher local density than the one predicted by the model of equation 4.5.2. In particular, we choose as overdensity candidates the stars whose local density is 5 times higher than the model-predicted density. We used that cut to select a reasonable (~500) number of candidates. In these positions we find the centroid based on the stars that lie within the volume that defines its local density in order to group candidates that may be part of the same overdensity. There we create a $100 \times 100 \times 100$ mesh on its vicinity (defined by 3 times the typical distance between stars expected from the model density) and calculate the local density in each point of the mesh using the method defined in equation 4.5.1.

To better define the overdensity in terms of its shape, star members, and boundaries, we perform a scan of the mesh in the three spatial axes. In each axis we perform contours on the planes defined by the density mesh that correspond to 0.4 times the peak density of the mesh. If a star is on that plane and lie inside the mentioned contour, we select it as part of that overdensity. At the same time, we put together all the contours to characterize qualitatively the overdensity shape and size. We do this process in the 3 spatial axes, making three 2D scans on the mesh. Finally, we selected the final overdensity candidates grouping all the raw overdensities that have at least one star in common. In F4.11 we show an example of the 3 projections on the 3 spatial axes of an overdensity candidate (FOVD01).

With this method we selected 49 overdensity candidates, one of which corresponds to the Sagittarius dSph galaxy. To check the significance of the overdensities, we compare the number of observed stars versus the expected stars by the model of eq. 4.5.2 in the volume around the overdensity. Due to the complex shapes of the overdensities, we characterized their volume as a rectangular bounding-box around them to estimate their significance, which often results in an underestimation of their significance. This effect is particularly high in stream-shaped overdensities where this setup adds considerable volume that is not part of the overdensity. In conclusion, this method proposes overdensity candidates, but a spectroscopic follow-up is necessary in order to get radial velocities to confirm their coherence in terms of motion. The significance of the candidates is calculated in standard deviations from the expected
number density using a Poisson uncertainty:

\[ \sigma = \frac{N_{\text{obs}} - N_{\text{exp}}}{\sqrt{N_{\text{obs}}}}. \]  

(4.5.3)

In T4.3 we show an extract of the overdensity candidates found in the SSS data. We have removed overdensities 9, 11, 14, 15, 33, 39, 42 and 45 from the table due to their low significance based on equation 4.5.3. It is important to note that the significance was calculated in a bounding box around the overdensity candidate, and the stars selected as part of the overdensity by the algorithm are not exactly the same as the ones contained in the bounding box, which is why we see overdensities containing a single star. Even when this may sound absurd, we must remember that we are using RRL as tracers of the stellar population and that we are using automatic algorithms to map the entire zone covered by the SSS. At these locations we have spotted an overdense zone compared to the model, but the algorithm just formally selected one star as part of it.

A notable overdensity is FOVD10, which correspond to the Sagittarius dSph galaxy. We estimated its significance only at the very center (180 kpc) of the galaxy due the extension and irregular shape of its trailing stream, which extends for \(~17\) kpc in x and z, and have a width of \(~6\) kpc in y. A rectangular bounding box around the complete extension of the overdensity would greatly overestimate its volume, leading to wrong estimation of its significance. In F4.12, we show the three projections of this overdensity.

Visualization of the overdensities was a great challenge. Projections on 2d density maps is a useful, but in our experience an insufficient alternative, given the complexity of some of the overdensities. This method works very well when you know exactly where you are looking. Since our approach was a wide study of a largely unexplored part of the Milky Way, we needed a more powerful visualizing tool. In this spirit, we developed a 3D web application to let us visualize our overdensities. This application will be included in our actual journal paper, with accompanying instructions, for the reader to gain better feeling of the detected overdensities.
<table>
<thead>
<tr>
<th>id</th>
<th>Shape</th>
<th>RA (J2000)</th>
<th>Dec</th>
<th>D (kpc)</th>
<th>Nov</th>
<th>Nbb</th>
<th>Nmo</th>
<th>significance</th>
<th>x (kpc)</th>
<th>y (kpc)</th>
<th>z (kpc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOVD01</td>
<td>Stream-like</td>
<td>190.910</td>
<td>-15.871</td>
<td>11.450</td>
<td>9</td>
<td>14</td>
<td>4.3</td>
<td>2.6</td>
<td>4.397</td>
<td>-6.754</td>
<td>8.367</td>
</tr>
<tr>
<td>FOVD02</td>
<td>Stream-like</td>
<td>167.533</td>
<td>-36.096</td>
<td>15.761</td>
<td>15</td>
<td>47</td>
<td>17.2</td>
<td>4.4</td>
<td>5.554</td>
<td>-14.305</td>
<td>6.005</td>
</tr>
<tr>
<td>FOVD03</td>
<td>Stream-like/Clumpy</td>
<td>242.017</td>
<td>-17.246</td>
<td>3.5300</td>
<td>49</td>
<td>166</td>
<td>24.8</td>
<td>11.0</td>
<td>5.136</td>
<td>-0.238</td>
<td>1.485</td>
</tr>
<tr>
<td>FOVD04</td>
<td>Clumpy</td>
<td>146.906</td>
<td>-17.819</td>
<td>9.4230</td>
<td>20</td>
<td>48</td>
<td>9.6</td>
<td>5.6</td>
<td>10.798</td>
<td>-8.049</td>
<td>4.233</td>
</tr>
<tr>
<td>FOVD05</td>
<td>Spherical</td>
<td>157.220</td>
<td>-20.397</td>
<td>16.815</td>
<td>8</td>
<td>11</td>
<td>1.7</td>
<td>2.8</td>
<td>10.022</td>
<td>-14.273</td>
<td>8.728</td>
</tr>
<tr>
<td>FOVD06</td>
<td>Spherical</td>
<td>170.283</td>
<td>-18.335</td>
<td>9.3310</td>
<td>3</td>
<td>6</td>
<td>0.5</td>
<td>2.3</td>
<td>7.738</td>
<td>-7.174</td>
<td>5.936</td>
</tr>
<tr>
<td>FOVD07</td>
<td>Spherical</td>
<td>238.236</td>
<td>-18.335</td>
<td>4.7470</td>
<td>1</td>
<td>10</td>
<td>1.1</td>
<td>2.8</td>
<td>4.131</td>
<td>-0.582</td>
<td>2.137</td>
</tr>
<tr>
<td>FOVD08</td>
<td>Stream-like/Clumpy</td>
<td>96.7730</td>
<td>-41.711</td>
<td>7.3830</td>
<td>19</td>
<td>6</td>
<td>0.3</td>
<td>5.1</td>
<td>10.7</td>
<td>-6.425</td>
<td>-2.758</td>
</tr>
</tbody>
</table>

Table 4.3: Table showing a sample of the detected overdensities. RA and Dec are in degrees, D in kpc and x y z represent the Galactocentric positions in kpc. Nov stands for the number of stars selected as part of the overdensity, Nbb is the number of stars inside the bounding box used to calculate significance, and Nmo is the number of stars predicted by the model of eq. 4.5.2. The significance number is the standard deviation calculated with equation 4.5.3.
Figure 4.7: From top to bottom, sample RRab selected with order 3, 4, and 6, respectively.
Figure 4.8: Error of the observations vs. apparent magnitude for the SSS telescope. We fitted an exponential to model the error and used the resulting fit to obtain the synthetic lightcurves described in section 4.4.1
Figure 4.9: The completeness of both the period selection (red crosses) and the RRab selection (blue dots) in terms of the apparent magnitude. We mainly focused on finding false negatives, since our procedure aims to have the cleanest sample possible. Eye inspection test of our sample shows that the false positives are less than 1% of the sample.
Figure 4.10: Metallicity distribution of the halo. The left panel shows the metallicities with $D_m < 5$. The right panel contains all the selected RRab. The peak that we see at $[\text{Fe/H}] = -1.5$ is spurious, corresponding to a hand-imposed metallicity of all the RRab with metallicities outside the range $-2.5 < [\text{Fe/H}] < 0.5$. In both histograms we see a Gaussian distribution with a peak at $[\text{Fe/H}] = -1.4$. 
Figure 4.11: Three 2D projections of the overdensity FOVD01. From left to right the projection are on the xz plane, the xy plane and the yz plane, respectively. The color code shows the logarithm of the ratio between the observed and the expected local density. This overdensity in particular has a significance of 2.6σ and has an elongated shape that suggests a stream-like overdensity.

Figure 4.12: Same as F4.11 but for the overdensity of the Sagittarius dSph. The huge (>10 kpc) stream of stars can be seen right next to it. However, to confirm its members and real extension, further spectroscopic studies are required.
Chapter 5

Conclusions

We have presented two different approaches to RR Lyrae in surveys in this work, each with different specific motivations, but within the same background context, namely, the challenges that wide-field astronomy is bringing now, and will continue to bring in the future surveys generation. On the first side, we presented an extensive theoretical calibration of the metallicity and physical parameters (temperature, luminosity, gravity, mass, radius) of RR Lyrae stars in the $ugriz$ SDSS photometric filter system. Extensive, but rather simple, analytical expressions are provided, which are shown to provide the quoted physical quantities with remarkable (internal) precision. The effect of an increase in the helium abundance is investigated, and found to have but a minor impact upon the derived results. Similarly, the extrapolation of the relations to the very metal-poor and metal-rich regimes is also discussed, on the basis of independent sets of HB evolutionary tracks, generally showing that most of the relations perform very well, even at $Z \lesssim 0.0003$ or $Z \gtrsim 0.01$. Photometric metallicities obtained on the basis of our metallicity relation were also compared with spectroscopic measurements from De Lee (2008), showing that, statistically, the relations work as expected. We anticipate that these relations can be very useful in current and future large-scale variability surveys, including the LSST.

On the other hand, we also presented a wide systematic study of the southern hemisphere ($-75^\circ \leq$
δ ≤ −22°) to find and classify RRab stars, discovering more than 10,500 of them in the SSS data. The full sample covers a range in brightness from 10.5 ≤ V ≤ 19.3 which means a range in heliocentric distances from 1 to 48 kpc. The stars were found using completely automatic procedures that are available for general-purpose use, and could be applied to any other survey. These algorithms have, depending on the quality of the photometry, a completeness of RRab classification greater than 60% at V = 18 magnitudes, and gets the correct period in more than 90% of the cases. This software leaves the possibility of manual classification of the faintest stars, including other types of periodic variables, using the output given by the software in terms of lightcurve information and period.

A search for substructure in the (largely unexplored) southern celestial hemisphere was performed by means of an intensive, automatic, study to spot the most obvious overdensities. We classified 41 overdensity candidates with significances typically greater than 2σ, and some as high as 10σ (compared to halo models). We provide these overdensities with galactocentric positions and their tentative RRab members. Most of the overdensities show a spherical geometry, but some show an elongated distribution suggesting a possible stream of stars left behind by tidal disruption of a satellite orbiting the galaxy.

In the future, we expect to perform spectroscopic follow-up to confirm the overdensity candidates by means of their coherence in velocity space. Also, a deeper search for RRL is desirable, since the APS aim is to have the cleanest automatic sample of RRab, thus leaving outside the faintest RRL. In this sense, visual inspection of the lightcurves of the faintest stars of the sample could reveal more distant RRab stars, and hence, will let us trace the halo even further. Another idea is to classify other types of variables, like eclipsing binaries, that are clearly visible on the output lightcurves of the APS. Finally, a more focused search for substructure in the southern Milky Way halo is also desirable, looking at places where models predict the locations of the known streams in the southern hemisphere that have been poorly studied with RR Lyrae (and other tracers) so far.
Bibliography


Baart, M. L. 1982, IMA JNA, 2, 241


Blazhko, S. 1907, Astron. Nachr., 175, 325


Chaboyer, B. 1999, in Post-Hipparcos Cosmic Candles, 237, 111
De Lee, N. 2008, Ph.D. Thesis (Michigan State University)
De Lee, N. M., Smith, H. A., Beers, T. C., et al. 2013, AAS Meeting Abstracts, 221, #316.03


Freeman, K., & Bland-Hawthorn, J. 2002, ARAA, 40, 487


Jerzykiewicz, M., & Wenzel, W. 1977, AcA, 27, 35


Jurcsik, J. 1995, Acta Astronomica, 45, 653


Keller, S. C., Schmidt, B. P., Bessell, M. S., et al. 2007, PASA, 24, 1


Petersen, J. O. 1986, 170, 59


Schwarzschild, M. 1940, Harvard College Observatory Circular, 437, 1


Sesar, B. 2011, RR Lyrae Stars, Metal-Poor Stars, and the Galaxy, 135


Strömgren, B. 1963, QJRAS, 4, 8


68


Zinn, R., & West, M. J., 1984, APJS, 55, 45

Zhevakin, S. A. 1953, Russ. A.J. 30, 161