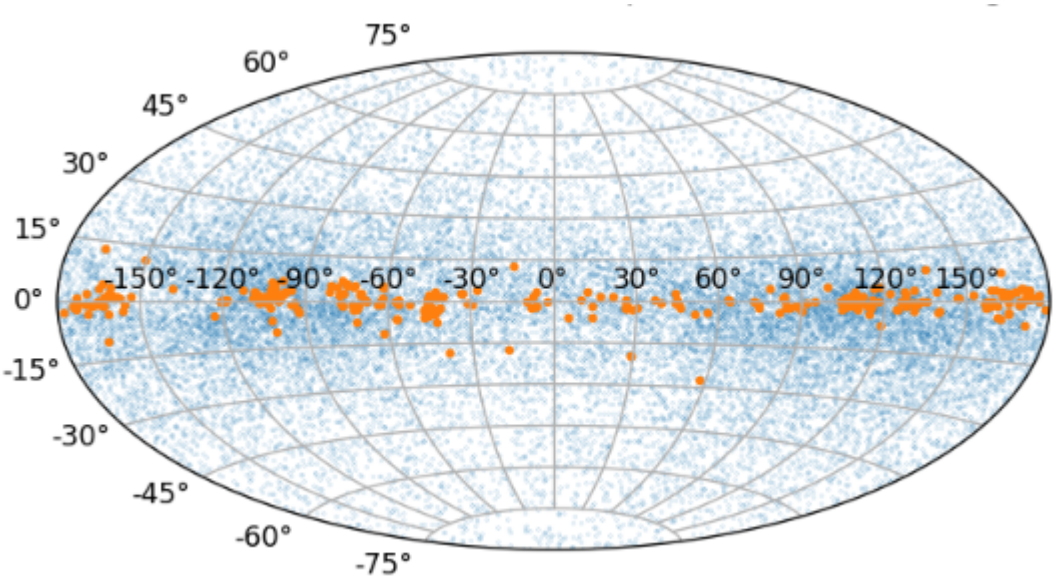
An exploration of O-star data contained in early data release 3 of the GAIA satellite

Ruben van Erp

July 3, 2022

Bachelor thesis Mathematics and Physics & Astronomy

Supervisors: prof. dr. Lex Kaper & prof. dr. Frank Pijpers



Institute of Physics
Korteweg-de Vries Institute for Mathematics
Faculty of Sciences
University of Amsterdam

Abstract

In 2013 ESA launched the GAIA satellite into space to take a survey of all visible sources and determine their position, proper motion and some colour information. This data has higher precision than anything that has come before, the satellite's unparalleled ability to accurately measure parallax allows us to make more accurate visualisations of our interstellar neighbourhood. Main sequence stars are classified according to their spectral types. O-type stars are the most massive and most luminous main sequence stars. The life cycle of an O-stars is extremely short compared to other types of stars due to their rapid fuel burning. This short life makes O-stars interesting to study; if we observe an O-star then we know that it must have formed near where we see it, thus allowing us to study active star formation regions in space. Describing star position and motion is done in reference to the ICRF (International Celestial Reference Frame), this reference frame can be transformed to align with the galactic plane to get the Galactic Reference Frame, which helps us study star distribution in our local milky way. We find that O-stars tend to be concentrated near the galactic plane, where other stars are more spread out. An important aspect of astrometry (the science of measuring objects in space) is keeping track of possible biasses in the data. An important bias is the Malmquist bias, which tells us that dim stars are only visible up to a certain point, if you are measuring with a sensor which has a detection limit for low luminosity. Understanding this bias helps us correct for it in the data that we have. Another bias is the Lutz-Kelker bias which describes how, when measuring parallaxes, we will underestimate the distance to stars on average. It is important to correct for the Lutz-Kelker bias when using data to calibrate astrometric equipment.

Title: An exploration of O-star data contained in early data release 3 of the GAIA satellite

Authors: Ruben van Erp, revanerp@gmail.com, 12844454
Supervisors: prof. dr. Lex Kaper & prof. dr. Frank Pijpers
Second graders: Prof. dr. J.M. Mooij & Prof. dr. A. de Koter
End date: July 3, 2022

Institute of Physics
University of Amsterdam
Science Park 904, 1098 XH Amsterdam
http://www.iop.uva.nl

Korteweg-de Vries Institute for Mathematics
University of Amsterdam
Science Park 904, 1098 XH Amsterdam
http://www.kdvi.uva.nl

Contents

1. Introduction

In this paper we will explore some aspects of the incredible data gathered by the GAIA mission. We will explore known phenomena in astronomy and show how these can be represented with the measurements made by GAIA. The data from the GAIA satellite is publicly available and can be retrieved from [1] with the use of ADQL (Astronomical data query language) which is a programming language based on SQL. In this paper we will make use of a specific sample of O-type stars as gathered by [2], and some random samples which were retrieved with the random_index parameter in the GAIA data. This random sample will be used to make comparisons with the O-star data. The most important result in this paper is figure 3.9, which shows a three dimensional map of 472 O-stars and the motion of a subset of these. It is the first time in history that such a map can be made with this level of precision. The data used is from early data release 3 which became available on December 3rd 2020. Early data release 3 is the first installment of the third data release of the GAIA satellite, it's second installment (data release 3) became available the 13th of june 2022, which was after all analysis was completed for this paper. The first half of this paper will focus on the results of the data analysis and is intended to provide general knowledge of some important aspects of astronomy and will use plots from the GAIA data to illustrate this general knowledge. The second half of this paper explores two important systematic errors in astronomy, the Malmquist bias and the Lutz-Kelker bias. These biasses will be explained and some common corrections will be derived. The section on the Malmquist bias will follow the results from Malmquist's original paper [3]. In this section results from mathematician Charlier will be used which can be found in [4], we will not go in depth into these results since they are beyond the scope of this paper. In the section on the Lutz-Kelker bias we will follow the work done by Lutz and Kelker in [5]. We have updated some of the figures used by Lutz and Kelker with the help of modern software.

2. Astronomy

2.1. Astrometry

Astrometry is the branch of astronomy involved with making precise position and velocity measurements for various extraterrestrial objects. For a long time all we had to observe the sky above were our own sadly inadequate eyes. The advent of telescopes and precision measuring devices brought new life to the field of astrometry. A big question is the one of scale, we know that everything outside of our solar system is very far away, at least on human scales, but how far away is everything exactly? The best way we have to measure how far away a (not too distant) star is is parallax. Parallax measurements rely on the movement of our earth around the Sun. We can measure the position of a nearby star compared to some (far away) background stars at one point in the year, if we do the same half a year later we see that the star has seemingly shifted compared to the background stars. This shift is measured in degrees (or arcseconds (") or milliarcseconds (mas)), and allows us to calculate how far away the star is. The usual distance unit associated with parallax measurements is a parsec (pc): 1 parsec is defined as the distance to a star whose parallax is equal to 1". A parsec is $3.086 \cdot 10^{13}$ kilometres. Measuring parallax with a telescope from earth is not very precise, the first person to succeed in measuring the distance to a star is Friedrich Bessel in 1838, his measurement of 61 Cygni provided a parallax of $p = 0.314$ [6], more recent measurements have concluded that the actual value for the parallax is equal to $p = 0.286$. Measuring parallaxes by hand is a time consuming and difficult process. A massive leap was made in 1989 when ESA built and launched the Hipparcos satellite, this satellite was able to measure parallaxes of many stars at the same time to an accuracy of about a milliarcsecond [7]. But this is not where the history of parallax measurements ends, a more recent telescope has once again made a massive leap in the field of astrometry.

2.2. The GAIA Satellite

The GAIA satellite was launched in 2013 by ESA and started operating in 2014. Its main goal is to create the largest and most precise three dimensional map of the milky way [8]. GAIA's main strength lies in its extremely precise ability to pinpoint the distance to nearby stars. It is the successor of the Hipparcos mission, with the similar task of mapping our interstellar neighbourhood. The GAIA satellite is an order of magnitude more precise and offers not only a quantitative leap forward compared to Hipparcos but also a qualitative leap. The accuracy of its parallax measurements is about 0.4 microarcseconds (μas) for bright ($G < 14$ mag) sources and about 0.1 mas

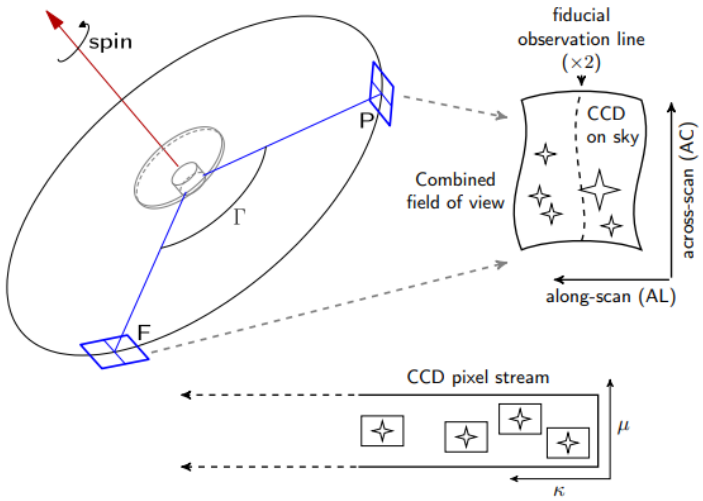


Figure 2.1.: Schematic of the GAIA satellite. Taken from [10].

for sources around $G = 17$ mag [9]. It is able to achieve this high precision distance mapping because of the sophisticated equipment on the spacecraft. On board of the GAIA satellite are two telescopes which are separated by a large angle (Γ =106.5°) which is well known and very stable. The spacecraft rotates at a rate of one rotation every 6 hours. The light from both telescopes is redirected into a single focal plane, as shown in figure 2.1. The satellite is also equipped with two low resolution prisms which disperse the light to have measurements in two different color bands, one operates at wavelengths of 330–680nm (BP) and the other 640–1050nm (RP). GAIA also has a spectrometer on board to accurately measure spectra in the range of 847–874nm. With these instruments we are able to determine, for a single star, its distance, its proper motion, its apparent magnitude, and some color information. This paper is based on data gathered mostly from Early data release 3 (GAIA EDR3), in this data release radial velocity is only included for some stars, so this information has to supplemented with data from different sources. The detection limit for GAIA's main telescopes in the G band is approximately 21 mag. In section 4.1 on the Malmquist bias we will discuss the implications of this limit.

2.3. Star classification

All stars are divided into different so called "spectral types", the different spectral types are: O, B, A, F, G, K, M. O-type stars are the hottest and most luminous stars and M-type stars are the coolest and faintest. This classification is applicable only to main

sequence stars, these are stars that fuse hydrogen in their cores to produce helium. Our Sun is a main sequence G-type star. A useful tool to study star classification is the Hertzsprung-Russel diagram (HR diagram), this diagram plots stars based on their absolute magnitude and temperature. Since temperature is closely related to color for stars, instead of plotting temperature we can use the difference in apparent magnitude of two different color-bands. Figure 2.2 shows a Hertzsprung-Russell diagram compiled from a random sample of stars (in blue) contained in the GAIA early data release 3 (GAIA EDR3). In orange a sample of O-type stars is shown. In this diagram the main sequence is clearly visible. The O-star sample appears to diverge from the main sequence, this is most likely due to an astrometric effect called reddening and extinction. Due to dust contained in our milky way and in our line of sight to stars certain stars will appear "redder" and less bright than they actually are. This is due to radiation being scattered by this interstellar dust. Shorter wavelengths are scattered more than longer wavelengths, leaving more red light. So without this dust, the O-star sample would be moved up and to the left in the HR diagram and align with the main sequence [11]. Since the HR diagram in figure 2.2 is composed of a random sample of GAIA data, it is susceptible to other statistical biases and may not be an actual representation of reality, we will focus on some of these biases in detail later in this paper. Another helpful tool for star classification are color-color diagrams, where two different colors are plotted against each other. Color here is defined as the difference in apparent magnitude of two color bands. In figure 2.3 we have a color-color diagram of a random star sample from the GAIA EDR3 (in blue) and O-type stars (in orange). We can clearly see that, while there is some overlap, the O-type stars differentiate themselves from a random sample. Figure 2.4 shows the combination of our HR diagram and the color-color diagram, here we clearly see our HR-diagram and also we get an even better distinction between the O-type stars and the random sample. Historically we did not have access to accurate distance measurements so all we had were color-color diagrams to make classifications with, this muddled the data set and affected analyses negatively.

2.3.1. O-stars

In this paper we are most interested in O-type stars, these are the hottest and most luminous main sequence stars. What makes these O-stars so interesting is the fact that they burn through their fuel quickly. Due to this quick fuel burning their lives are very short compared to other types of stars, this means that when we observe an O-type star we can be certain that this star was formed relatively recently. Studying these objects can give us insight in where star formation occurs and what happens to stars at the beginning of their lives. The thing that makes observing O-type stars difficult is the fact that they are the rarest type of main sequence stars. Throughout this paper we will use a sample of O-stars collected by Apellániz et al [2]. This sample was cross-referenced with the GAIA catalogue to find the corresponding GAIA EDR3 ID's of these stars, and from this data was collected about these stars, the full list of the used O-type stars and their GAIA EDR3 ID's can be found in the appendix A, as well as the ADQL code to retrieve the data from the GAIA catalogue. The sample contains 472 (usable) stars.

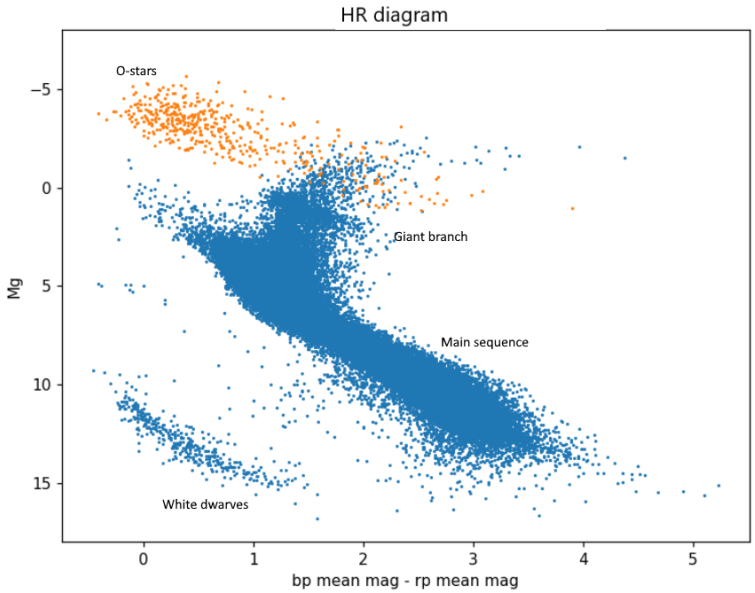


Figure 2.2.: Hertzsprung-Russell diagram from GAIA early data release 3. A random sample of stars is used (blue) and a sample of exclusively O-type stars (orange). The main features are highlighted.

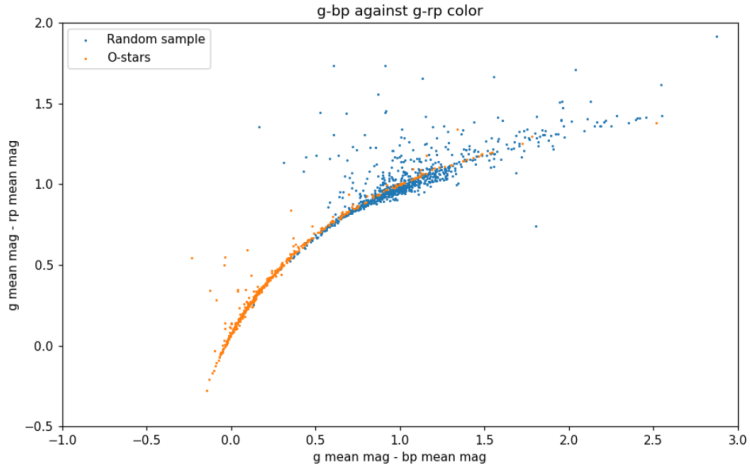


Figure 2.3.: Color-Color diagram of a random GAIA sample (blue) and O-stars (orange).

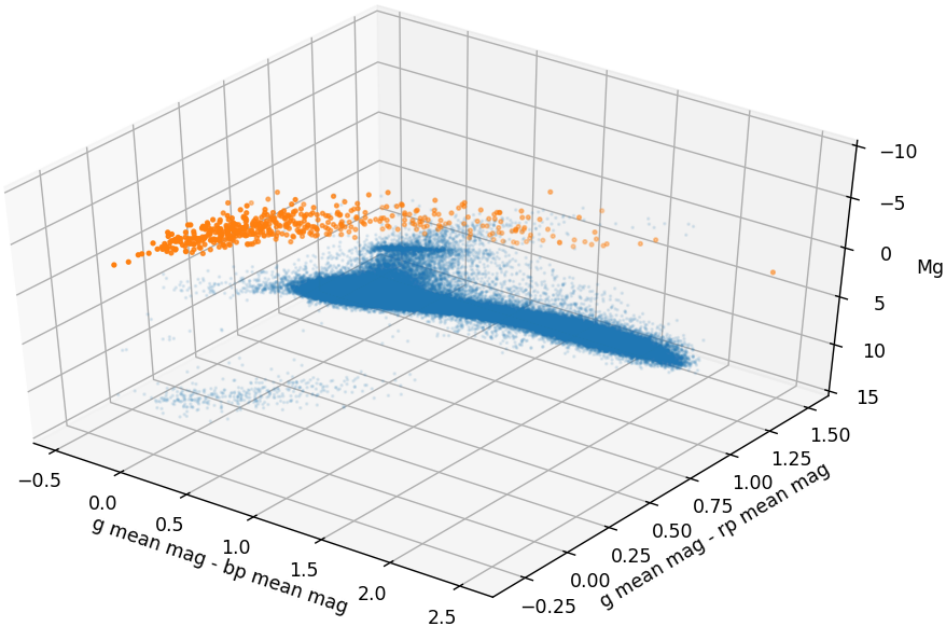


Figure 2.4.: Color-Color diagram of a random GAIA sample (blue) and O-stars (orange) with an added third axis for absolute magnitude.

An attempt was made to train a neural network to look for more O-stars within GAIA EDR3, but the attempt was unsuccessful, as the data used (distance, color, magnitude) was not discriminatory enough to distinguish between O- and B-stars, and due to the relative abundance of B-stars compared to O-stars, no new O-stars were identified.

2.4. Galactic coordinates

The position of all stars are defined by the International Celestial Reference System (ICRS), this reference system has at its origin the solar system barycenter. Each point is defined by its right ascension (ra), declination (dec) and distance (r). The right ascension is its azimuthal coordinate and the declination its polar coordinate. The $dec = 0$ plane is aligned with the ecliptic and the $(ra, dec) = (0, 0)$ point is aligned with the vernal equinox or the point where the ecliptic and the celestial equator intersect. Another coordinate system is the Galactic coordinate system. It is also a spherical coordinate system, which is aligned not with the solar system but with the galactic plane. The origin of the galactic coordinate system is still the barycenter of the solar system but its azimuthal coordinate (l) is now aligned with the galactic plane, its polar coordinate is b and distance (r). The $(l, b) = (0, 0)$ point is aligned with the galactic center. The transformation from ICRS to galactic coordinates is given by:

$$r_{Gal} = \begin{pmatrix} -0.05487556041621545 & -0.8734370902348850 & -0.4838350155487132 \\ +0.4941094278755837 & -0.4448296299600112 & +0.7469822444972189 \\ -0.8676661490190047 & -0.1980763734312015 & +0.4559837761750669 \end{pmatrix} r_{ICRS}, \tag{2.1}$$

(As taken from [12])
where r_{ICRS} is the position vector

$$r_{ICRS} = \begin{pmatrix} x_{ICRS} \\ y_{ICRS} \\ z_{ICRS} \end{pmatrix} = \begin{pmatrix} \cos(\alpha)\cos(\delta) \\ \sin(\alpha)\cos(\delta) \\ \sin(\delta) \end{pmatrix}$$

where α is the right ascension and δ the declination.
These reference frames are defined by so called reference sources. Since everything moves in the universe it is very difficult to find some anchor to base an inertial reference system on, so to compromise a selection has been made of very distant sources, since these appear to be stationary from our perspective. These sources must be extremely bright to still be visible at a vast distance, so mostly quasars are used for this purpose [13].
In GAIA EDR3 the proper motion is included as the change in right ascension (pmra) and declination (pmdec) in miliarcseconds (mas) per year, these are μ_α and μ_δ respectively. Since right ascension is an azimuthal coordinate, the actual right ascension velocity component of the proper motion is given by $\mu^*_\alpha = \mu_\alpha \cos(\delta)$. In this paper we will also follow a cartesian coordinate system based on the galactic reference frame. The origin of this cartesian coordinate system is the solar system barycenter, it's x-axis is aligned with the galactic center, and the z-axis is perpendicular to the galactic disc, it's coordinates

are denoted $(x_{Gal}, y_{Gal}, z_{Gal})$. The transformation from the galactic coordinate system to the cartesian one is given by:

$$\begin{pmatrix} x_{Gal} \\ y_{Gal} \\ z_{Gal} \end{pmatrix} = \begin{pmatrix} \cos(l)\cos(b) * r \\ \sin(l)\cos(b) * r \\ \sin(b) * r \end{pmatrix}$$

3. Structure of the milky way

3.1. Spatial structure

Our milky way is a spiral galaxy. Most of the stars in our galaxy are concentrated in a disc with sprial arms. At the moment 5 spiral arms have been identified, in order of furthest to the galactic center to closest we have the Outer, Perseus, Local, Sagittarius and Scutum arms [14]. Due to dust and extinction effects we can currently not look deeper into our own galaxy past these 5 arms. The distance to the black hole at the center of the milky way is roughly 8 kiloparsec [15]. The star density roughly decreases exponentially with the distance perpendicular to the disc. We can indeed see in figure 3.3 that most stars are indeed concentrated around $z = 0$ where z is the distance to the galactic plane in kiloparsec. We expect that the distributions of x_{Gal} and y_{Gal} should be roughly homogeneous at this scale, since the disc is roughly homogeneous in our close neighbourhood. We see for these coordinates however that they are also slightly concentrated around 0, this is due to the Malmquist bias which is discussed in section 4.1, in short, we see fewer stars the further out we look. This bias results in the fact that we are sampling a roughly spherical section of space. If we were sampling a homogeneous sphere we would expect the x, y and z distributions to be roughly dome shaped. For x_{Gal} and y_{Gal} we recognise a dome-like structure in their distribution. The distribution for the z coordinate also suffers from the Malmquist bias to equal extent as x_{Gal} and y_{Gal}, but we can still clearly see that the peak around 0 is more significant for z_{Gal} than for x_{Gal} and y_{Gal} as expected for a disc-shaped galaxy.
Figure 3.2 shows the position of O-stars (orange) and a random sample (blue) projected on the celestial sphere in galactic coordinates. We see that the O-stars are clumped together and do not show a homogeneous distribution around the equator. We thus expect to see clustering when we look at a three dimensional representation of this data. In figure 3.5 we see the distance distribution for a random sample of GAIA EDR3 data up to 0.5 kiloparsec. We would expect, if the sample was complete, a roughly quadratic distribution, since we are sampling from a sphere. We would expect some flattening at further distances, due to the exponential decrease in density in the z_{Gal} direction, but what we see instead is a more drastic flattening than expected. This flattening is again due to the Malmquist bias (section 4.1), we can correct for this bias by randomly dropping sources from our sample based on a the desired final distribution and the current distribution. For a source at distance r we allocate a random number between 0 and 1, we drop the source from the sample if the random number is lower than $c * \frac{p_{desired}(r)}{p_{current}(r)}$, where $p_{desired}(r)$ (green line in figure 3.5) is the desired distribution at distance r and $p_{current}(r)$ is the current distribution at distance r (orange line in figure 3.5), and c is a constant to normalise the expression. Now we have a distance

distribution which is more accurate to reality, but the sources contained in this sample are still subjected to the Malmquist bias, which entails that at further distances we remain with only bright sources (see 4.1 for the reason for this). If we now represent this data as a density plot, as we have done in figure 3.6, we see that before the sampling it looks like most stars are concentrated around the Sun. After the sampling we see that this concentration disappears and a more homogeneous plot is produced.

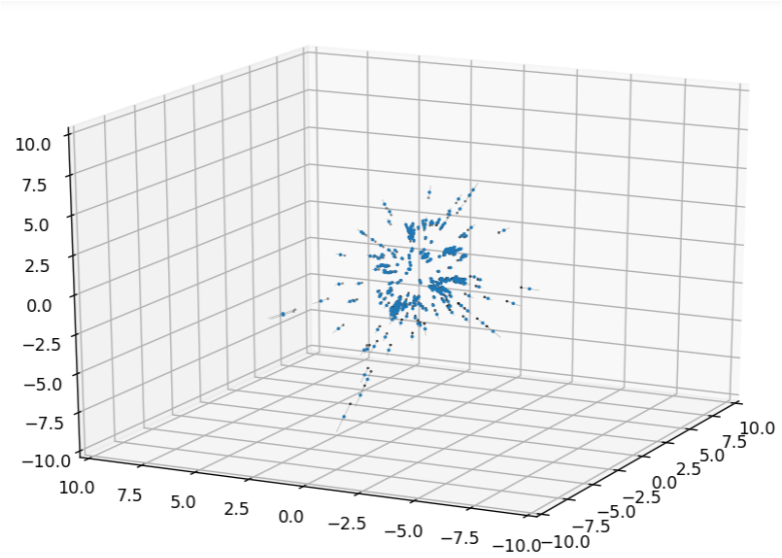


Figure 3.1.: Spatial distribution of our O-star sample with error bars on the distance.

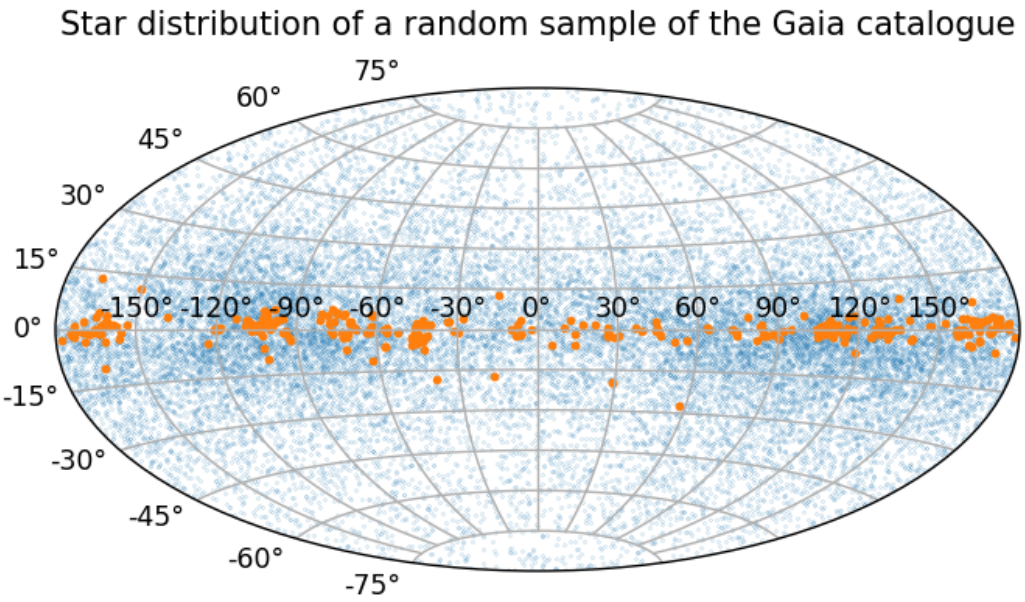


Figure 3.2.: Distribution of a random sample of stars in the GAIA catalogue (blue), and O-stars (Orange)

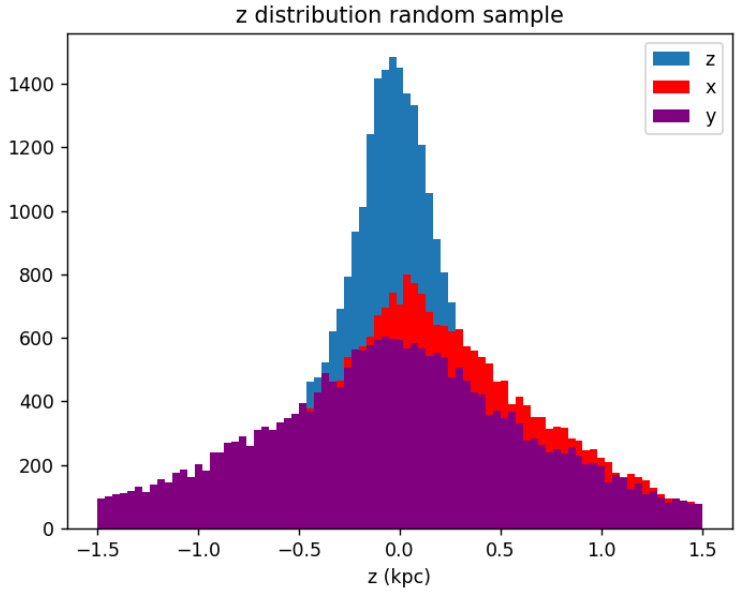


Figure 3.3.: Distributions of $x_{Gal}, y_{Gal}, z_{Gal}$ of a random sample.

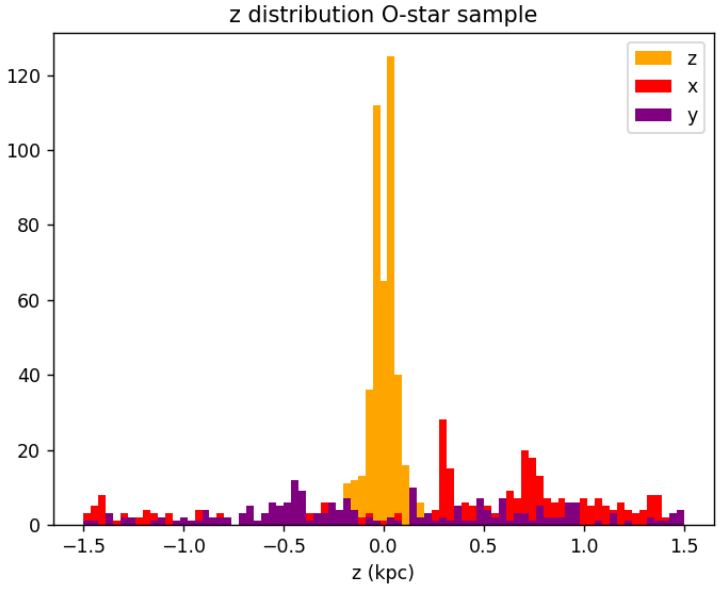


Figure 3.4.: Distributions of $x_{Gal}, y_{Gal}, z_{Gal}$ of our O-star sample.

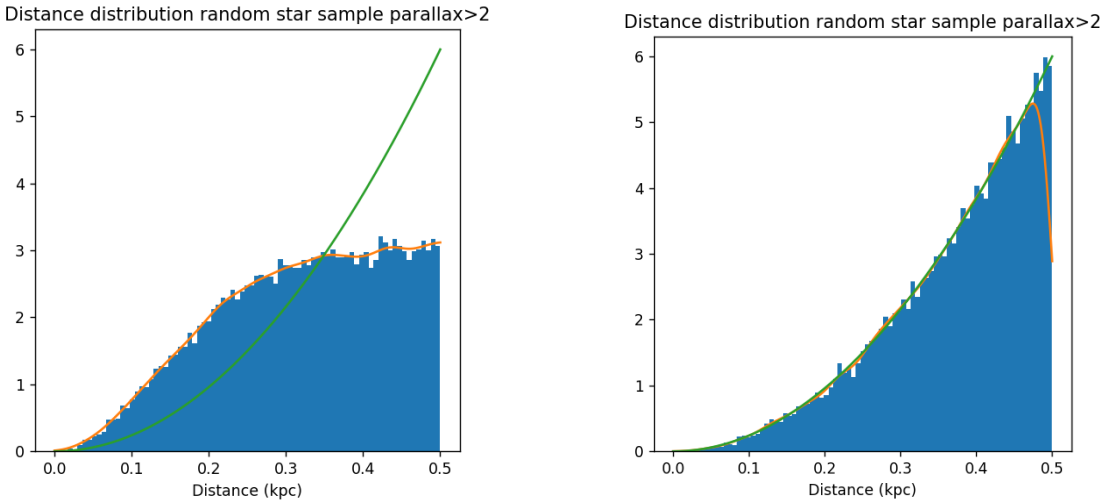


Figure 3.5.: Distribution of the distance to stars in a random sample of the GAIA data up to a distance of 0.5 kpc (left) and the distance distribution to a subsample of this random sample to correct the distribution to a desired distribution (right).

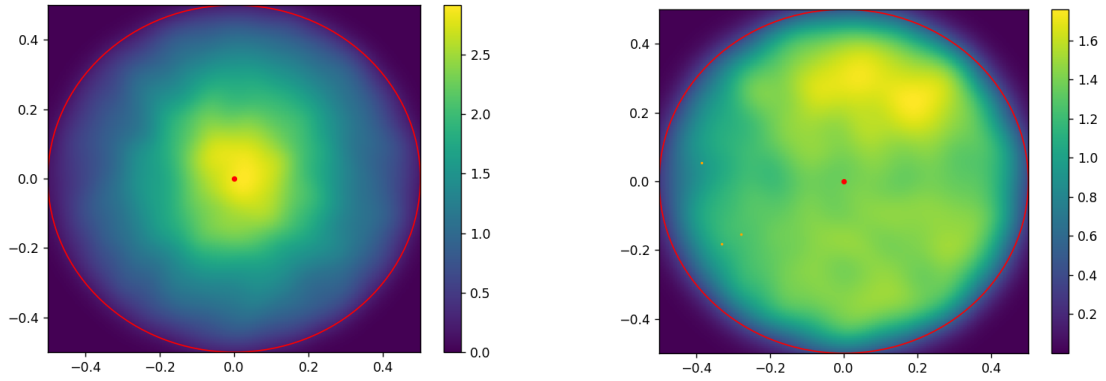


Figure 3.6.: Spatial star density in a 0.5kpc neighbourhood around the Sun (red dot) of a random sample of the GAIA data (left) and the Spatial star density of a subsample of this random sample to correct the distribution to a desired distribution (right).

3.2. Differential galactic rotation

All stars in the milky way move, most move in a circular motion around the galactic center. Describing this motion helps understand the milky way and its kinematics. In 1927 J.H.Oort [16] published a now famous paper about the differential galactic rotation. In this paper he describes the expected motion of stars, based on their galactic coordinates. He shows that the expected value of μ_{l^*} is only dependent on the galactic coordinates l, b and A, B the so called Oort constants, which may be experimentally determined. Where μ_{l^*} is the component of the proper motion in the longitudinal direction in galactic coordinate reference frame. The relation is given by

$$\overline{\mu_{l^*}} = \frac{A}{4.74}\cos(2l)\cos(b) + \frac{B}{4.74}\cos b. \tag{3.1}$$

He also states that the expected radial velocity ($\overline{\rho}$) is given by

$$\overline{\rho} = V_0\cos(\lambda) + rA\sin(2l)\cos^2(b). \tag{3.2}$$

Where λ is the distance to the solar apex.

3.2.1. deriving the expected tangential and radial velocities

We will derive expression 3.1 and 3.2 for the case $b = 0$, so for stars in the galactic plane. In figure 3.7 (adapted from [17]) we have all the quantities we need to derive the expression. "Center" represents the Galactic center "Sun" our Sun and "Star" the observed star. We assume that the Sun and the observed star follow a circular motion around the galactic center. V_0 is the velocity of the Sun and V the velocity of the observed star. The relative tangential velocity (v_t) is simply given by

$$v_t = V\sin(\alpha) - V_0\sin(90 - l) \tag{3.3}$$

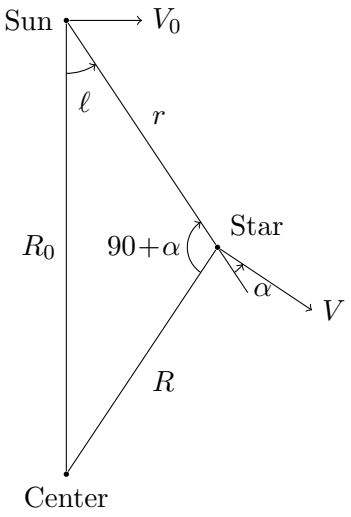


Figure 3.7.: Schematic of the situation for the derivation of the expected radial and tangential velocities of a star in the galactic plane.

By using the fact that $\sin(90 - x) = \cos(x)$, and the fact that $\sin(\alpha) = \frac{R_0 \cos(l) - r}{R}$ we see that 3.3 can be rewritten as

$$v_t = \frac{V}{R}(R_0 \cos(l) - r) - \frac{V_0}{R_0} R_0 \cos(l). \tag{3.4}$$

$\frac{V}{R}$ defines the angular velocity ω so we can rewrite this expression as

$$v_t = R_0(\omega - \omega_0)\cos(l) - \omega r.$$

Where ω_0 is the Suns angular velocity around the galactic center and ω is the observed star's angular velocity.

$$R^2 = R_0^2 + r^2 - 2rR_0 \cos(l) \tag{3.5}$$

which we divide by R_0^2 we see that for relatively close stars $\left(\frac{r}{R_0}\right)^2$ will be small, so we set this to 0. By using the series approximation $(1 + x)^n \approx 1 + nx$ we get that

$$\frac{R}{R_0} = \left(1 - \frac{2r\cos(l)}{R_0}\right)^{\frac{1}{2}} \approx 1 - \frac{r\cos(l)}{R_0} \tag{3.6}$$

or $R \approx R_0 - r\cos(l)$. From the first order taylor expansion $(f(x) \approx f(a) + (x - a)f'(a))$ we get

$$(\omega - \omega_0) \approx \left(\frac{d\omega}{dR}\right)_{R_0} (R - R_0) \tag{3.7}$$

Now we see, since $\omega = \frac{V}{R}$, that

$$\frac{d\omega}{dR} = \frac{d}{dR}\frac{\omega}{R} = \frac{1}{R}\frac{dV}{dR} - \frac{V}{R^2} \tag{3.8}$$

For stars in our neighbourhood $\omega r \approx \omega_0 r$, now substituting 3.7 in 3.5 we get

$$v_t = \left[\frac{V_0}{R_0} - \left(\frac{dV}{dR}\right)_{R_0}\right] r\cos^2(l) - \frac{V_0}{R_0} r. \tag{3.9}$$

Now with the relation $\cos(2l) = 2\cos^2(l) - 1$ we can rewrite this as

$$v_t = \frac{A\cos(2l) + B}{4.74} r \tag{3.10}$$

where

$$A = \frac{1}{2}\left[\frac{V_0}{R_0} - \left(\frac{dV}{dR}\right)_{R_0}\right] \quad B = \frac{1}{2}\left[\frac{V_0}{R_0} + \left(\frac{dV}{dR}\right)_{R_0}\right]. \tag{3.11}$$

The factor $\frac{1}{4.74}$ comes from the conversion from kilometers per second to arcseconds per year. Now we have simply that $\mu_l = \frac{v_t}{r}$, to get the desired result. The differential galactic rotation is visible in GAIA data, in figure 3.8, courtesy of [18] we clearly see the double cosine relation given by Oort.

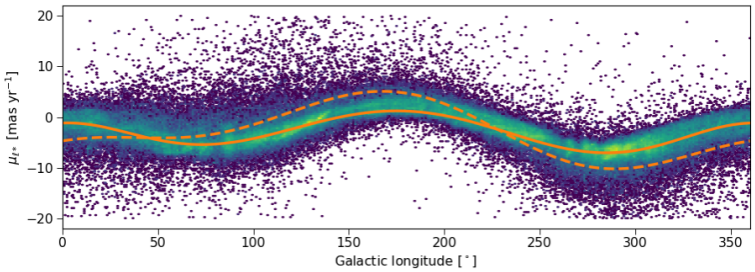


Figure 3.8.: A scatter plot of the galactic coordinate l against the velocity component μ_{l^*}, the double cosine relation is clearly visible. Courtesy of [18]

For the relative radial velocity component (v_r) we see that

$$v_r = V\cos(\alpha) - V_0\cos(90 - l) \tag{3.12}$$

Now using the sine rule we can simply express the expected radial velocity in terms of the difference in angular velocity as:

$$v_r = R_0(\omega - \omega_0)\sin(l). \tag{3.13}$$

Now we substitute equation 3.7 into the expression above to get

$$v_r = R_0\left(\frac{d\omega}{dR}\right)_{R_0}(R - R_0)\sin(l). \tag{3.14}$$

Which becomes

$$v_r = \frac{1}{2}\left[\frac{V_0}{R_0} - \left(\frac{dV}{dR}\right)_{R_0}\right] r\sin(2l) \tag{3.15}$$

using equation 3.8, and noting that V evaluated at R_0 is V_0 and using the substitution $R - R_0 \approx -r\cos(l)$ and finally using the relation $\sin(x)\cos(x) = \sin(2x)$.

3.3. O-Stars

Since O-stars are the stars with the shortest lives, we can learn where star formation happens in the galaxy by studying where we find O-stars. Figure 3.4 shows that O-stars are highly concentrated near the galactic plane (z=0) compared to a random sample (figure 3.3). This shows that star formation is likely to occur more near the galactic plane, where star density is highest.
Figure 3.9 shows the position of 472 O-stars, these O-stars were selected from the O-star catalogue by Apellániz et al. [2]. The shown O-stars were the O-stars which have data in GAIA EDR3 and with a parallax error smaller than 10%. Since no radial velocity data is included for these O-stars in GAIA EDR3 these had to be sourced from literature. A full list of sources can be found in appendix A as well as all the names and ID's of the used O-stars.
For the O-stars for which a radial velocity was found in literature, their motion is represented in figure 3.9 by the red line, this shows their direction of travel relative to our Sun. The size of the red line indicates relative speed.
We can see from figure 3.9 that O-stars do not necessarily follow the pattern expected from differential galactic rotation. This is also noted by Oort in [19]. The O-stars do show obvious signs of clustering. One of these clusters is highlighted in figure 3.10, we can clearly see an "arm" of O-stars which move in similar directions. These clusters are likely very active star formation regions.
Another interesting phenomenon we see in figure 3.9, which is highlighted in figure 3.11, we clearly see some O-stars which are spatially separated from the galactic plane with peculiar velocities corresponding to motion away from the plane. These are examples of so called runaway O-stars, more information about these runaway O-stars and what mechanisms may cause this motion can be found in [20].

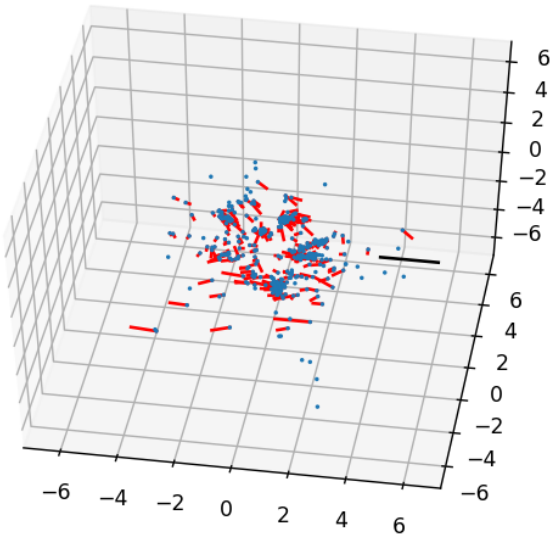


Figure 3.9.: Spatial distribution of the O-star sample in the galactic coordinate system. The red lines represent the relative peculiar motions of the O-stars. The black line points towards, and ends at the galactic center.

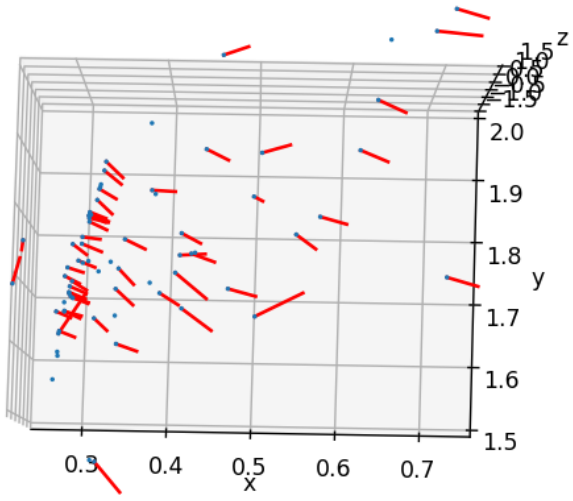


Figure 3.10.: A cluster of O-stars, $x_{Gal} \in [0.25, 0.75], y_{Gal} \in [1.5, 2]$.

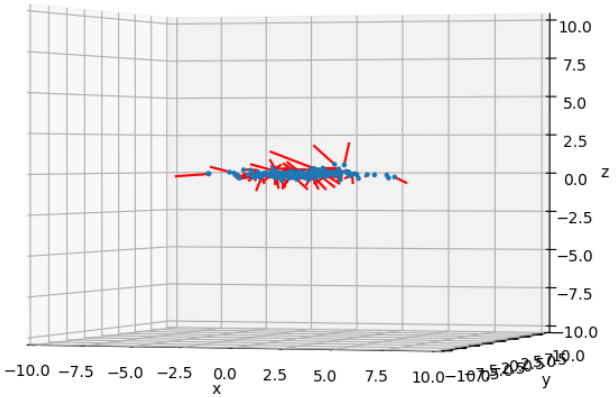


Figure 3.11.: Side view of the spatial O-star distribution, notice the two O-stars which are clearly separated from the bulk in the galactic plane, and their peculiar motion away from the plane.

4. Statistics

4.1. Malmquist Bias

In observational astronomy, collection of data is usually limited by brightness of the observed sources, dim sources are harder to detect than bright sources, this leads to a selection bias known as the Malmquist bias. Since observed brightness of stars is inversely proportional to its distance squared, it is easy to see that dim sources which are far away are harder to detect than dim sources nearby. This leads to the fact when far sources are observed the dim sources will not be measured due to some hard limit of the detector. Due to the Malmquist bias, one may erroneously derive that the average magnitude of a star survey is higher than it actually is.

The GAIA satellite also suffers from the Malmquist bias, this is clearly demonstrated in a plot of some random sample where calculated magnitude is plotted against distance, as done in figure 4.1. We can clearly see some hard detection limit. We would have expected to see no correlation between distance and absolute magnitude, but clearly one is present.

A traditional correction for the Malmquist bias, as first mentioned by Malmquist himself in [3], is to correct the calculated average ($\overline{M}$) by adding $1.382\sigma^2$. While simple, this correction is based on some very idealized assumptions, namely:

1. Light follows only the inverse square law and is not affected by an interstellar medium.
2. The universe is completely homogeneous.
3. The universe is completely isotropic.
4. The sample is complete, no sources are missing above some apparent magnitude (m_{lim}) .
5. The distribution of luminosities can be approximated by a gaussian centered on an intrinsic absolute magnitude (M_0)
6. All stars are of the same spectral type with intrinsic mean absolute magnitude M_0 and dispersion σ.

Clearly these assumptions don't hold in a dust filled flat spiral galaxy, centered around a point roughly 8 kpc away from the Sun, containing many different types of stars.

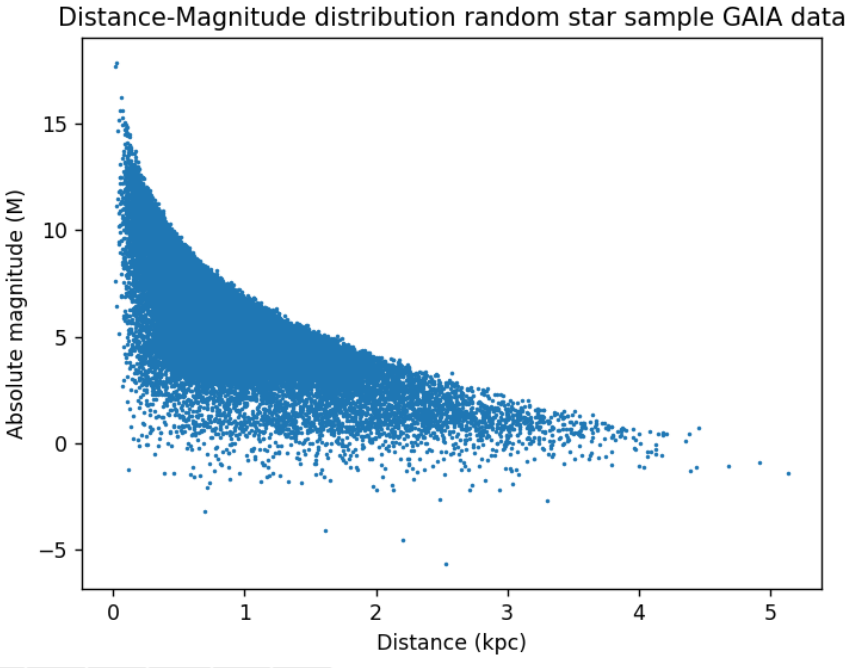


Figure 4.1.: Distribution of absolute magnitude of stars and their distances. We can clearly see a correlation between the maximum observed magnitude and the distance. A lower magnitude corresponds to a brighter star.

4.1.1. Malmquist in GAIA

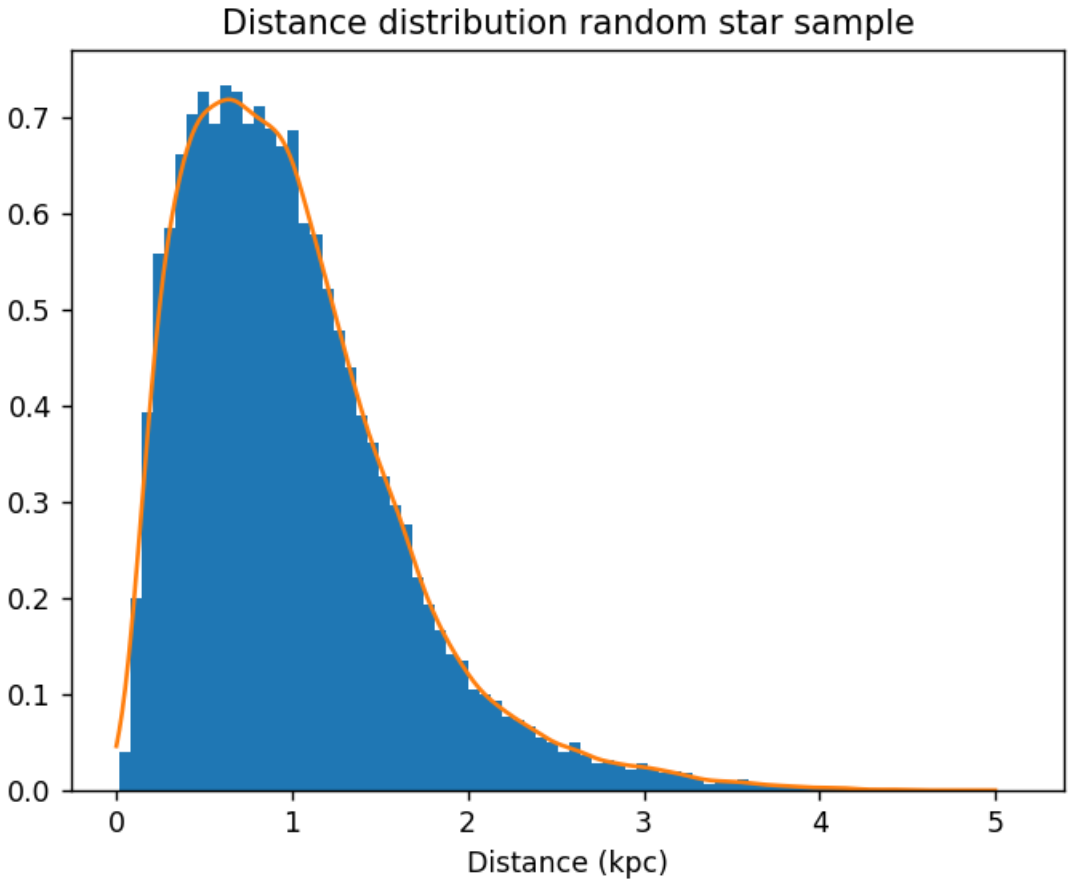


Figure 4.2.: A distribution of the distance to stars in a random sample of GAIA data. We clearly see a strong decline in the amount of sources measured at greater distances.

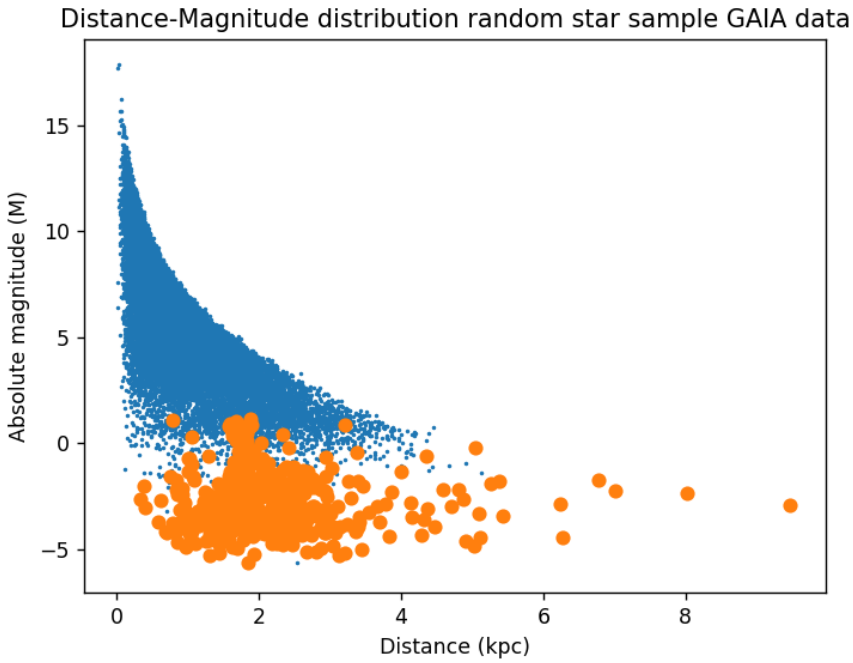


Figure 4.3.: Distribution of absolute magnitude of stars and their distances. We see that the O-stars sample available does not suffer directly from the malmquist bias. A lower magnitude corresponds to a brighter star.

In figure 4.3 we have added the O-star sample to our distance-absolute magnitude plot. As we can clearly see, the O-stars are much brighter than non-O-stars. Since most O-stars in our sample are closer than 6 kpc, we can clearly say that we do not expect any O-stars within this distance to be missed due to the Malmquist bias. Of course not all the O-stars in our galaxy are within 6 kpc, so this sample is in no way complete beyond this distance. We also see in figure 4.4 that most O-stars in our sample are concentrated around 2 kpc, and less so around 5 kpc, this shows that our sample is most likely also not complete within the 6 kpc bound, even though this is not due to the Malmquist bias.
Now we understand the Malmquist bias we can take another closer look at the star distributions shown in figure 3.3. If the universe was homogeneous and our sample up to distance r was complete, we would expect the distribution as shown in figure 3.3 to show perfect domes. But taking the Malmquist bias into account we would expect the amount of sources in our sample to decrease with distance, this is exactly what we see happen in figure 3.3 for the x_{Gal} and y_{Gal} coordinates. We do not see this same pattern in figure 3.4, as expected since we see that our O-star sample does not suffer the effects of the Malmquist bias.

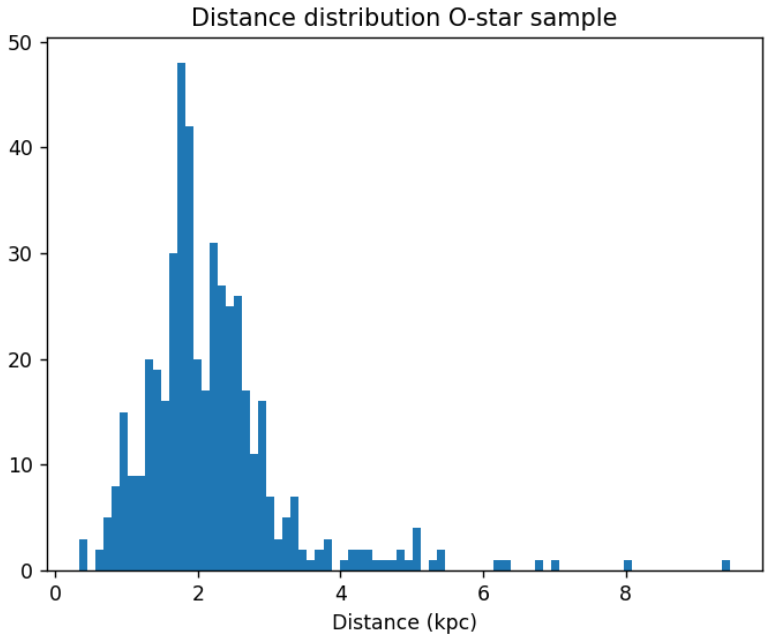


Figure 4.4.: Histogram of the distance distribution of our O-star sample.

4.1.2. Deriving the correction

This derivation will follow the derivation given by Malmquist in [3].
Let $a(m)$ be the distribution function for stars of apparent magnitude m. It is clear then, that

$$\int_{-\infty}^{\infty} a(m)\, dm = N$$

is the total number of stars. We now define the relative distribution function $\varphi(M)$ of absolute magnitudes such that

$$\int_{-\infty}^{\infty} \varphi(M) dM = 1.$$

Next take $D(r)$ as the star density at distance r from us. We now get that

$$a(m) = \omega \int_{0}^{\infty} r^2 D(r) \varphi(m - \frac{1}{b}\log(r))\, dr. \tag{4.1}$$

We use here the established relationship between apparent and absolute magnitude

$$M = m - 5\log_{10}(r) = m - \frac{1}{b}\log(r) \tag{4.2}$$

where $\frac{1}{b} = \frac{5}{\log(10)} \approx= 2.1714$ or $b \approx 0.4605$, and with $\omega = 4\pi$ which comes from the fact that we integrate over the whole sky. We see that in 4.1 we have three unknown

functions, which means that if we know two of these we can derive the third. Before we continue, we will recapitulate some assumptions, as Malmquist first posed them in his work:

1. *There exists no appreciable absorption of light in space.*
2. *The distribution function of the absolute magnitudes is a normal curve of type A.*
3. *This function is independent of the distances of the stars.*

The units used by Malmquist we those introduced by Charlier in [4] where the unit for distance is a siriometer (sir) which is equal to 10^6 times the mean distance of the earth to the Sun. And the unit for time is a stellar year (st) wich is equal to 10^6 times the length of a tropical year. Charlier defines absolute magnitude as the apparent magnitude at a distance of one siriometer. These units are not relevant for our derivation. By assumption (2.) we may write the distribution function of the absolute magnitudes as:

$$\varphi(M) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(M-M_0)^2}{2\sigma^2}}. \tag{4.3}$$

With M_0 the mean absolute magnitude and σ the corresponding standard deviation. From this we will now look at this expression for a given apparent magnitude (m), it is not necessary that this expression $F_m(M)$ should be gaussian, but it will also not differ greatly according to assumption (3.), thus we may write this expression as a so called Charlier Series described by Charlier in [4]. With the tools of Charlier series in hand we may continue by rewriting our expression for the distribution of Absolute magnitudes for a given apparent magnitude as a Charlier series

$$F_m(M) = \varphi_m(M) + A_3\varphi_m^{(3)} + A_4\varphi_m^{(4)} + A_5\varphi_m^{(5)} + A_6\varphi_m^{(6)} + ... \tag{4.4}$$

with $\varphi_m(M)$ the normal curve

$$\varphi_m(M) = \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{(M-\overline{M_m})^2}{2\sigma_m^2}}. \tag{4.5}$$

With $\overline{M_m}$ the mean absolute magnitude for constant apparent magnitude m and σ_m its corresponding standard deviation. These quantities need not be equal to M_0 and σ. For practical reasons no higher order terms are included than $A_4\varphi_m^{(4)}$. Charlier introduced two abstract quantities, the skewness (S) and the excecss (E) defined as:

$$S = \frac{3A_3}{\sigma_m^3} \tag{4.6}$$

$$E = \frac{3A_4}{\sigma_m^4} \tag{4.7}$$

A positive skewness expresses that the median and mode are greater than the mean, and a negative skewness the opposite. By neglecting higher order terms than $A_4\varphi_m^{(4)}$ we get indeed that if the excess is small that

$$\text{median} = \text{mean} + \frac{1}{3}\sigma_m \cdot S$$

and

$$\text{mode} = \text{mean} + \sigma_m \cdot S$$

For our derivation, we wish to learn about the behavior of 4.4, or, we wish to express $\overline{M_m}, \sigma_m, A_n$ in terms of M_0,σ and $a(m)$. To do this we will use the method of moments developed by Charlier [4].
The relative moment about the mean of the n'th order is given by

$$\nu_n(m) = \overline{(M - M_m)^n} = \omega \int_0^\infty r^2 D(r)(M - \overline{M_m})^n \varphi(m - \frac{1}{b}\log(r))\, dr \tag{4.8}$$

Where to use the notation that a line above a quantity denotes the mean value of that quantity.$\overline{M - m}$ is defined by the relation

$$\nu_1(m) = 0 \tag{4.9}$$

and σ_m from

$$\sigma_m^2 = \nu_2(m). \tag{4.10}$$

To get the value of $\overline{M_m}$ we start with 4.1

$$a(m) = \omega \int_0^\infty r^2 D(r)\varphi(m - \frac{1}{b}\log(r))\, dr.$$

We have that

$$a(m) \cdot \overline{M_m} = \omega \int_0^\infty M r^2 D(r)\varphi(m - \frac{1}{b}\log(r))\, dr. \tag{4.11}$$

By differentiating 4.1 with respect to m we get

$$\frac{d\, a(m)}{dm} = -\omega \int_0^\infty \frac{M - M_0}{\sigma^2} r^2 D(r)\varphi(m - \frac{1}{b}\log(r))\, dr, \tag{4.12}$$

from which we get

$$\omega \int_0^\infty M r^2 D(r)\varphi(m - \frac{1}{b}\log(r))\, dr = M_0 a(m) - \sigma^2 \frac{d\, a(m)}{dm} \tag{4.13}$$

Combining this with 4.11 we see that

$$\overline{M_m} a(m) = M_0 a(m) - \sigma^2 \frac{d\, a(m)}{dm}. \tag{4.14}$$

By setting

$$\log(a(m)) = f_1(m) \tag{4.15}$$

we find

$$\frac{1}{a(m)} \frac{d\, a(m)}{dm} = \frac{d\, f_1(m)}{dm}, \tag{4.16}$$

now by 4.14 we have

$$\overline{M_m} = M_0 - \sigma^2 \frac{d\, f_1(m)}{dm}. \tag{4.17}$$

Now if we once again differentiate 4.12 with respect to m we find:

$$\begin{aligned}\frac{d^2\,a(m)}{dm^2} =&\omega\int_0^\infty r^2 D(r)\frac{(M-M_0)^2}{\sigma^4}\varphi(m-\frac{1}{b}\log(r))\,dr\\ &-\omega\int_0^\infty r^2 D(r)\frac{1}{\sigma^2}\varphi(m-\frac{1}{b}\log(r))\,dr,\end{aligned} \tag{4.18}$$

from which we get

$$\begin{aligned}&\sigma^4\frac{d^2\,a(m)}{dm^2}+\sigma^2 a(m) =\\ =&\omega\int_0^\infty r^2 D(r)(M-\overline{M_m}+\overline{M_m}-M_0)^2\varphi(m-\frac{1}{b}\log(r))\,dr\\ =&\omega\int_0^\infty r^2 D(r)(M-\overline{M_m})^2\varphi(m-\frac{1}{b}\log(r))\,dr\\ &+2\omega(\overline{M_m}-M_0)\int_0^\infty r^2 D(r)(M-\overline{M_m})\varphi(m-\frac{1}{b}\log(r))\,dr\\ &+\omega(\overline{M_m}-M_0)^2\int_0^\infty r^2 D(r)\varphi(m-\frac{1}{b}\log(r))\,dr\end{aligned} \tag{4.19}$$

By dividing this result by $a(m)$ and taking into consideration the definition of 4.8 of the moments we get, according to 4.9

$$\int_0^\infty r^2 D(r)(M-\overline{M_m})\varphi(m-\frac{1}{b}\log(r))\,dr=0, \tag{4.20}$$

the expression

$$\sigma^4\frac{1}{a(m)}\frac{d^2\,a(m)}{dm^2}+\sigma^2=\nu_2(m)+(\overline{M_m}-M_0)^2. \tag{4.21}$$

From 4.16 we get

$$\frac{d^2\,f_1(m)}{dm^2}=\frac{1}{a(m)}\frac{d^2\,a(m)}{dm^2}-\left(\frac{1}{a(m)}\frac{d\,a(m)}{dm}\right)^2, \tag{4.22}$$

which gives, by 4.16

$$\frac{d^2\,f_1(m)}{dm^2}+\left(\frac{d\,f_1(m)}{dm}\right)^2=\frac{1}{a(m)}\frac{d^2\,a(m)}{dm^2}. \tag{4.23}$$

Inserting this into 4.21, we get, by 4.17

$$\nu_2(m)=\sigma^2+\sigma^4\frac{d^2\,f_1(m)}{dm^2} \tag{4.24}$$

or, rewritten by 4.10

$$\sigma_m^2=\sigma^2(1+\sigma^2\frac{d^2\,f_1(m)}{dm^2}). \tag{4.25}$$

From 4.18 we derive

$$\begin{aligned}&\frac{d^3\,a(m)}{dm^3} = \\ &= -\omega \int_0^\infty r^2 D(r) \frac{(M-M_0)^3}{\sigma^6} \varphi(m - \frac{1}{b}\log(r))\,dr \\ &+ 3\omega \int_0^\infty r^2 D(r) \frac{M-M_0)}{\sigma^4} \varphi(m - \frac{1}{b}\log(r))\,dr\end{aligned} \tag{4.26}$$

now 4.23 gives us

$$\frac{d^3\,f_1(m)}{dm^3} + 3\frac{d\,f_1(m)}{dm}\frac{d^2\,f_1(m)}{dm^2} = \frac{1}{a(m)}\frac{d^3\,a(m)}{dm^3} - \frac{1}{a(m)^2}\frac{d\,a(m)}{dm}\frac{d^2\,a(m)}{dm^2}, \tag{4.27}$$

which we can rewrite with the help of 4.16 and 4.23 as

$$\frac{d^3\,f_1(m)}{dm^3} + 3\frac{d\,f_1(m)}{dm}\frac{d^2\,f_1(m)}{dm^2} + \left(\frac{d\,f_1(m)}{dm}\right)^3 = \frac{1}{a(m)}\frac{d^3\,a(m)}{dm^3}. \tag{4.28}$$

Now by once again dividing by $a(m)$ and rewriting $M - M_0$ as $M - \overline{M_m} + \overline{M_m} - M_0$, we obtain, with 4.8,4.20 and 4.28

$$\begin{aligned}&\sigma^6\frac{d^3\,f_1(m)}{dm^3} + 3\sigma^6\frac{d\,f_1(m)}{dm}\frac{d^2\,f_1(m)}{dm^2} + \sigma^6\left(\frac{d\,f_1(m)}{dm}\right)^3 = \\ &= -\nu_3(m) - 3(\overline{M_m} - M_0)(\nu_2(m) - \sigma^2) - (\overline{M_m} - M_0)^3.\end{aligned} \tag{4.29}$$

From this, with the help op 4.17 and 4.24 we find

$$\nu_3(m) = -\sigma^6\frac{d^3\,f_1(m)}{dm^3}. \tag{4.30}$$

Repeating the steps taken, we find in the same manner

$$\nu_4(m) - 3\nu_2^2(m) = \sigma^8\frac{d^4\,f_1(m)}{dm^4} \tag{4.31}$$

$$\nu_5(m) - 10\nu_2(m)\nu_3(m) = -\sigma^{10}\frac{d^5\,f_1(m)}{dm^5} \tag{4.32}$$

$$\nu_6(m) - 10\nu_3^2(m) - 15\nu_2(m)\nu_4(m) + 30\nu_2^3(m) = \sigma^{12}\frac{d^6\,f_1(m)}{dm^6} \tag{4.33}$$

We will now make use of the relations between moments and the characteristics (A_3, A_4...) given by Charlier in [4]. The relations take the form

$$\begin{aligned}A_3 &= -\nu_3(m), \\ A_4 &= \nu_4(m) - 3\nu_2^2(m), \\ A_5 &= -\nu_5(m) + 10\nu_2(m)\nu_3(m), \\ A_6 &= \nu_6(m) - 15\nu_2(m)\nu_4(m) + 30\nu_2^3(m)\end{aligned} \tag{4.34}$$

Now relating this to the expressions for the moments found in 4.30,4.314.32 and 4.33 we get

$$\begin{aligned} A_3 &= \sigma^6 \frac{d^3 f_1(m)}{dm^3}, \\ A_4 &= \sigma^8 \frac{d^4 f_1(m)}{dm^4}, \\ A_5 &= \sigma^{10} \frac{d^5 f_1(m)}{dm^5}, \\ A_6 &= \sigma^{12} \frac{d^6 f_1(m)}{dm^6} + 10\sigma^1 2 \left(\frac{d^3 f_1(m)}{dm^3} \right)^3 \end{aligned} \tag{4.35}$$

Now adding this to 4.17 and 4.25 we get

$$\overline{M_m} = M_0 - \sigma^2 \frac{d f_1(m)}{dm} \tag{4.36}$$

$$\sigma_m^2 = \sigma^2 \left(1 + \sigma^2 \frac{d^2 f_1(m)}{dm^2} \right) \tag{4.37}$$

To derive the commonly used correction of $\overline{M_m} = M_0 - 1.382\sigma^2$, we make the simple assumption that $a(m)$ is of the form

$$a(m) = C \cdot e^{\beta m}$$

we get that $\log(a(m)) = f_1(m) = \log(C) + \beta m$ and $\frac{d f_1(m)}{dm} = \beta$, and higher derivatives vanish, thus we have:

$$\overline{M_m} = M_0 - \beta\sigma^2 \tag{4.38}$$

$$\sigma_m^2 = \sigma^2 \tag{4.39}$$

and

$$A_3 = A_4 = ... = 0 \tag{4.40}$$

Which gives us a distribution function of absolute magnitudes for a given m of

$$F_m(M) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(M - M_0 + \beta\sigma^2)^2}{2\sigma^2}} \tag{4.41}$$

which is a gaussian. In the simplest case, one of constant star density $D(r) = c$ where c is a constant, we get from

$$a(m) = 2\pi \int_0^\infty r^2 D(r) \varphi(m - \frac{1}{b}\log(r))\, dr \tag{4.42}$$

with

$$\begin{aligned} a(m) &= C \cdot e^{\beta m} \\ \varphi(m - \frac{1}{b}\log(r)) = F_m(m - \frac{1}{b}\log(r)) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{((m - \frac{1}{b}\log(r)) - M_0 + \beta\sigma^2)^2}{2\sigma^2}} \\ D(r) &= c \end{aligned}$$

which gives the solution:

$$\frac{2\pi e^{\frac{3}{2}b\left(\sigma^2(3b+2\beta)+2m-2M_0\right)}}{\sigma\sqrt{\frac{1}{b^2\sigma^2}}},$$

and if we assume $\sigma = 1$

$$C \cdot e^{\beta m} = 2\pi c \frac{e^{\frac{3}{2}b(3b+2(\beta+m-M_0))}}{\sqrt{\frac{1}{b^2}}}$$

from which we find, by rearranging:

$$C \cdot e^{\beta m} = 2\pi c \frac{e^{\frac{3}{2}b(3b+2(\beta-M_0))}}{\sqrt{\frac{1}{b^2}}} \cdot e^{3bm}.$$

And thus we get that

$$\beta = 3b = 3\frac{\log(10)}{5} \approx 1.382$$

which gives

$$\overline{M_m} = M_0 - 1.382\sigma^2.$$

4.2. Lutz-Kelker bias

The Lutz-Kelker bias arises when measuring parallaxes. On average, due to measurement errors, the measured parallax will be larger than the true parallax. This can be intuitively seen when we want to measure how many stars are contained within a distance r to us. Due to measuring errors we will measure that some sources which in reality are outside radius r are within this radius, but inversely we will also measure that some sources are outside of the radius r which in reality are within this radius. You may expect that these errors will correct each other on average, but this is not the case, since there are more stars just outside radius r than within. Since this is true for arbitrary r we gather that on average we underestimate the distance to stars. It is important to correct for this bias, since distance measurements are directly correlated with absolute magnitude calculations. Some astrometric measuring devices need to be calibrated to correctly measure magnitude, and if we use wrong magnitude values for this calibration the device's measurements will be off by some systematic error. We will derive a common correction for the average absolute magnitude in the next section.

4.2.1. Deriving the correction

This derivation roughly follows the one posed by Lutz and Kelker in [5].
We make the assumption that the error of parallax measurements is normally distributed, such that for some fixed true parallax of p the measured parallax (p_0) will be distributed as:

$$g(p_0|p) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(p_0-p)^2}{2\sigma^2}} \tag{4.43}$$

with σ the standard deviation. In a universe where stars are homogeneously distributed (which is an approximation for our own for stars with large parallaxes), we expect the number of stars between r and $r + dr$ to be equal to:

$$N(r)dr = 4\pi r^2 dr, \tag{4.44}$$

or, the expected number of stars between p and $p + dp$ is equal to:

$$N(p)dp = \frac{4\pi}{p^4}dp. \tag{4.45}$$

Now we can use this to find the distribution of true parallaxes based on a fixed measured parallax using Bayes' theorem. We get

$$g(p|p_0) \propto \frac{1}{p^4} e^{-\frac{(p_0-p)^2}{2\sigma^2}}. \tag{4.46}$$

This form is currently not normalised, but by substituting $\frac{1}{p}$ for $\frac{p_0}{p}$ we "normalise" it in the sense that $g(p|p_0) = 1$ when $p = p_0$. This form: $\frac{p}{p_0}$ is in a sense a dimensionless parallax, we define:

$$Z = \frac{p}{p_0}$$

We get

$$g(p|p_0) \propto G(Z) \equiv \frac{1}{Z^4} e^{-\frac{(Z-1)^2}{2(\sigma/p_0)^2}}. \tag{4.47}$$

We clearly see that the shape of this relation is only dependent on $\frac{\sigma}{p_0}$, we can visualize this with a plot of $G(Z)$ for some different values of $\frac{\sigma}{p_0}$, this is done in figure 4.5 As we can see in figure 4.5 for $\frac{\sigma}{p_0} = 0.15$ we see that near 0 $G(Z)$ diverges to infinity, this is also the case for $\frac{\sigma}{p_0} = 0.05$ and $\frac{\sigma}{p_0} = 0.25$, in fact this happens for all $\frac{\sigma}{p_0}$, near 0 the $\frac{1}{Z^4}$ dependence takes over, clearly this is nonphysical and such errors would stand out as erroneous and not be included in any application of the data. We see that for small $\frac{\sigma}{p_0}$ the distribution is almost normal, but shifted slight towards smaller Z. This shift becomes more pronounced as $\frac{\sigma}{p_0}$ increases until there is no more distinction possible between the $\frac{1}{Z^4}$ dependence and the physical distribution, at which point no useful correction can be derived.

If parallax measurements are used for absolute magnitude calibration, one needs to consider the Lutz-Kelker bias. We wish to determine some true absolute magnitude M_{true}, from some calculated absolute magnitude M_{observed}, we define

$$\Delta M \equiv M_{\text{true}} - M_{\text{observed}} = 5\log_{10}(\frac{p}{p_0}) = 5\log_{10}(Z).$$

The expected value of this correction is given by:

$$\overline{\Delta M} = \frac{5\int_0^\infty \log_{10}(Z)G(Z)dZ}{\int_0^\infty G(Z)dZ}$$

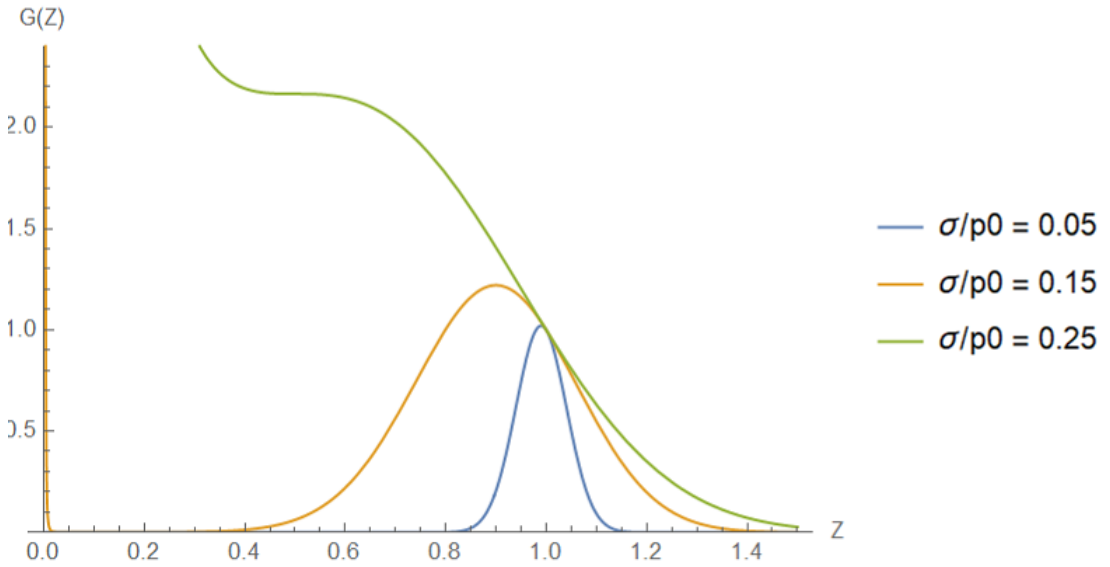


Figure 4.5.: Plot of $G(Z)$ for $\frac{\sigma}{p_0} = 0.05$,$\frac{\sigma}{p_0} = 0.15$ and $\frac{\sigma}{p_0} = 0.25$.

But as we observed earlier, this integral will diverge due to the divergence of $G(Z)$ near 0. To approximate the correction however, we may instead numerically calculate

$$\overline{\Delta M} \approx \frac{5\int_{\epsilon}^{\infty} \log_{10}(Z)G(Z)dZ}{\int_{\epsilon}^{\infty} G(Z)dZ}.$$

To get useful information from this expression, we wish to find some ϵ_{σ/p_0} for some fixed $\frac{\sigma}{p_0}$, such that we ignore most of the $\frac{1}{Z^4}$ dominated region and include most of the actual distribution. Clearly the higher $\frac{\sigma}{p_0}$ the more we need to exclude, but at some point we will no longer be able to find a useful ϵ. For instance, if we look at the graph of $G(Z)$ for $\frac{\sigma}{p_0} = 0.25$ in figure 4.5, there is no clear location for ϵ, while for the $\frac{\sigma}{p_0} = 0.15$ case, any ϵ between 0.1 and 0.3 would yield reasonably accurate estimations. We can plot

$$\overline{\Delta M} \approx \frac{5\int_{\epsilon}^{\infty} \log_{10}(Z)G(Z)dZ}{\int_{\epsilon}^{\infty} G(Z)dZ}.$$

In figure 4.6 $\overline{\Delta M}$ is plotted for different values of ϵ and $\frac{\sigma}{p_0}$, here we see that those lines which appear reasonably horizontal at some point yield accurate estimations. The best estimation for $\overline{\Delta M}$ is equal to the value at these horizontal regions. We see that the maximum $\frac{\sigma}{p_0}$ for which we can reasonably estimate the $\overline{\Delta M}$ is around $\frac{\sigma}{p_0} = 0.175$.
Figure 4.7 shows the best estimations for $\overline{\Delta M}$ for different values of $\frac{\sigma}{p_0}$.

Now correcting for the Lutz-Kelker bias is done by simply adding the value of $\overline{\Delta M}$ to the calculated average absolute magnitude of the sample based on the right $\frac{\sigma}{p_0}$.

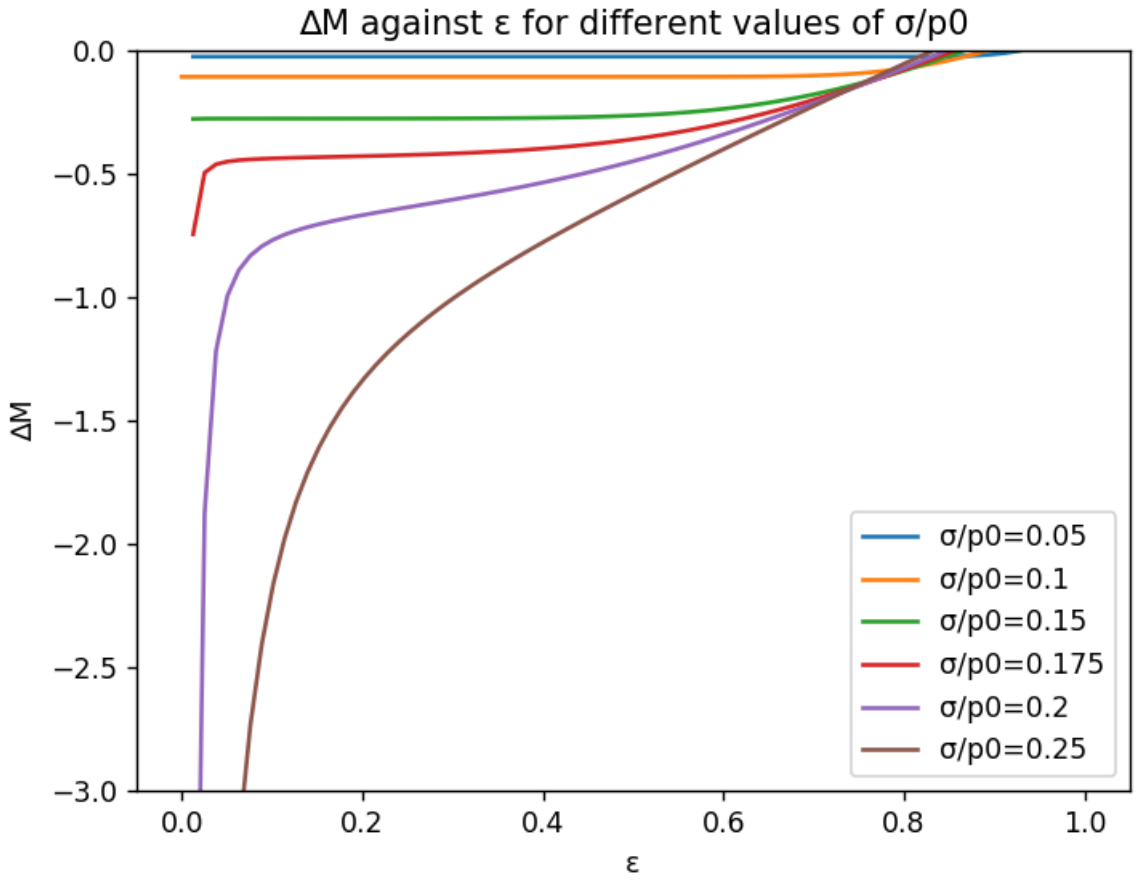


Figure 4.6.: Estimations for $\overline{\Delta M}$ for different values of ϵ, plotted for various values of $\frac{\sigma}{p_0}$.

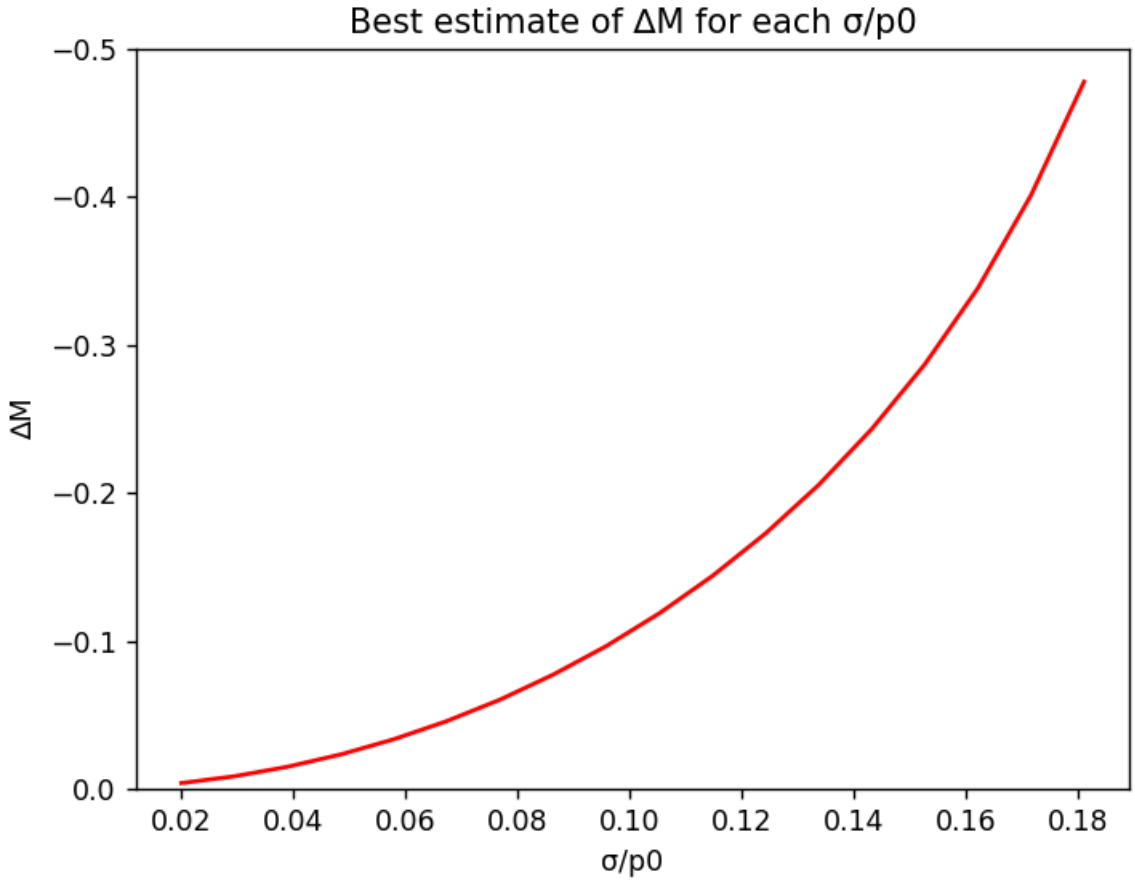


Figure 4.7.: The best estimations of $\overline{\Delta M}$ for different values of $\frac{\sigma}{p_0}$.

5. Conclusion

We have explored some concepts of astronomy, and shown how these concepts are represented in GAIA data. We learned about different reference frames to describe the position and motion of objects in our galaxy. We looked at how stars are distributed around us found that O-stars are more concentrated in the galactic plane. We found that the Malmquist bias has great effects on star distributions at greater distances and we learned how to correct for this by sampling to a known distribution. We have also derived a correction for the average luminosity of a sample affected by the Malmquist bias, even though we had to make some big assumptions. We learned about the differential galactic rotation and derived expressions for expected tangential and radial velocities. We saw that O-stars do not follow these expected velocities and instead show that they are spatially clustered with similar velocities. Finally we derived that when measuring parallaxes we underestimate the distance to measured sources on average and how to correct for this.

5.1. Further research

5.1.1. Modifications Malmquist correction

In our calculation for the common correction for the Malmquist bias, we made some pretty rough assumptions about how stars are distributed. We have shown in this paper that the homogeneity assumption is incorrect. For our interstellar neighbourhood, we expect a rough homogeneity in the l coordinate but not in the z_{Gal}. We expect the stellar density to decrease exponentially away from the disk. The integral (4.42) we used to derive the correction can be modified to more accurately represent reality. The original integral assumes that the density $D(r)$ is equal everywhere ($D(r) = 1$), we can modify this to represent our expectations. For instance if we expect rotational symmetry in the galactic plane and exponential decrease perpendicular to it, we may modify this density function to $D(\rho, z) = be^{-a|z|}$. The integral now becomes, with the transformation to cylindrical coordinates:

$$\int_0^{2\pi}\int_0^{\infty}\int_{-\infty}^{\infty}\rho D(\rho,z)\varphi(m-\frac{1}{b}\log(\sqrt{\rho^2+z^2}))dzdrd\theta = 2\pi\int_0^{\infty}\int_{-\infty}^{\infty}\rho D(\rho,z)\varphi(m-\frac{1}{b}\log(\sqrt{\rho^2+z^2}))dzd\rho.$$

$$= 2\pi\int_0^{\infty}\int_{-\infty}^{\infty}\rho be^{-a|z|}\varphi(m-\frac{1}{b}\log(\sqrt{\rho^2+z^2}))dzd\rho$$

Where $z = z_{Gal}$ and ρ is the distance to the z_{Gal} axis.
It is interesting to try to solve this integral (numerically) to see how the correction is

modified. Other density functions, which take into account measurements of the galaxy we are in, will lead to even more accurate results, but are harder to calculate.

5.1.2. Increase O-star sample size

As a brief exploration, a neural network was trained on our O-star sample and an equal sized random sample to try and distinguish between these to possibly find more O-stars within the GAIA data. As we have shown with figure 4.3 the O-star sample is not complete up to 6kpc, but since at this distance the sample should not be affected by the Malmquist bias we do expect more O-stars to be present in the data but not identified as such. In our brief attempt, the absolute magnitude and color data was used. In training, the neural network showed promise in distinguishing between O-stars and the random sample. But once the neural network was ran on a big random data set the network identified a bunch of known B-stars as O-stars. The data used was not able to distinguish between O-stars and some B-stars, and since B-stars are much more abundant, no conclusions could be drawn. It would be interesting to see if a better neural network could be made by including more aspects of the GAIA data in training, for example the l and b coordinates, since most O-stars are closely located near the galactic plane. Also the network could be trained specifically against B-stars to increase it's discriminatory capabilities between these closely related classifications.

Bibliography

[1] "gaia archive." https://gea.esac.esa.int/archive/, accessed on 01-07-2022.

[2] J. Apellániz, A. Sota, N. Morrell, R. Barbá, N. Walborn, E. Alfaro, R. Gamen, J. Arias, and A. Calvente, "The galactic o-star spectroscopic catalog (gosc) and survey (gosss): first whole-sky results and further updates," *Massive Stars: From α to Ω*, 06 2013.

[3] K. G. Malmquist, "On some relations in stellar statistics," *Meddelanden fran Lunds Astronomiska Observatorium Serie I*, vol. 100, pp. 1–52, Mar. 1922.

[4] C. Charlier, *Researches Into the Theory of Probability, by C.V.L. Charlier*. H. Ohlsson, 1906.

[5] T. E. Lutz and D. H. Kelker, "On the use of trigonometric parallaxes for the calibration of luminosity systems: Theory," *Publications of the Astronomical Society of the Pacific*, vol. 85, p. 573, oct 1973.

[6] F. W. Bessel, "Bestimmung der Entfernung des 61sten Sterns des Schwans.," *Astronomische Nachrichten*, vol. 16, p. 65, Dec. 1838.

[7] *The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission*, vol. 1200 of *ESA Special Publication*, Jan. 1997.

[8] "Gaia overview." http://www.esa.int/Our_Activities/Space_Science/Gaia_overview, accessed on 28-06-2022.

[9] Lindegren, L., Hernández, J., Bombrun, A., Klioner, S., Bastian, U., Ramos-Lerate, M., de Torres, A., Steidelmüller, H., Stephenson, C., Hobbs, D., Lammers, U., Biermann, M., Geyer, R., Hilger, T., Michalik, D., Stampa, U., McMillan, P.J., Castañeda, J., Clotet, M., Comoretto, G., Davidson, M., Fabricius, C., Gracia, G., Hambly, N.C., Hutton, A., Mora, A., Portell, J., van Leeuwen, F., Abbas, U., Abreu, A., Altmann, M., Andrei, A., Anglada, E., Balaguer-Núñez, L., Barache, C., Becciani, U., Bertone, S., Bianchi, L., Bouquillon, S., Bourda, G., Brüsemeister, T., Bucciarelli, B., Busonero, D., Buzzi, R., Cancelliere, R., Carlucci, T., Charlot, P., Cheek, N., Crosta, M., Crowley, C., de Bruijne, J., de Felice, F., Drimmel, R., Esquej, P., Fienga, A., Fraile, E., Gai, M., Garralda, N., González-Vidal, J.J., Guerra, R., Hauser, M., Hofmann, W., Holl, B., Jordan, S., Lattanzi, M.G., Lenhardt, H., Liao, S., Licata, E., Lister, T., Löffler, W., Marchant, J., Martin-Fleitas, J.-M., Messineo, R., Mignard, F., Morbidelli, R., Poggio, E., Riva, A., Rowell, N.,

Salguero, E., Sarasso, M., Sciacca, E., Siddiqui, H., Smart, R.L., Spagna, A., Steele, I., Taris, F., Torra, J., van Elteren, A., van Reeven, W., and Vecchiato, A., "Gaia data release 2 - the astrometric solution," *A&A*, vol. 616, p. A2, 2018.

[10] A. G. Brown, "Microarcsecond astrometry: Science highlights from gaia," *Annual Review of Astronomy and Astrophysics*, vol. 59, pp. 59–115, sep 2021.

[11] B. T. Draine, "Interstellar Dust Grains," *araa*, vol. 41, pp. 241–289, Jan. 2003.

[12] Alexey Butkevich and Lennart Lindegren, "Transformations of astrometric data and error propagation." https://gea.esac.esa.int/archive/documentation/GDR3/Data_processing/chap_cu3ast/sec_cu3ast_intro/ssec_cu3ast_intro_tansforms.html#SSS1, accessed on 4-06-2022.

[13] "International celestial reference system (icrs)." https://aa.usno.navy.mil/faq/ICRS_doc, accessed on 15-06-2022.

[14] M. J. Reid, K. M. Menten, A. Brunthaler, X. W. Zheng, T. M. Dame, Y. Xu, Y. Wu, B. Zhang, A. Sanna, M. Sato, K. Hachisuka, Y. K. Choi, K. Immer, L. Moscadelli, K. L. J. Rygl, and A. Bartkiewicz, "Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way," *apj*, vol. 783, p. 130, Mar. 2014.

[15] The GRAVITY Collaboration, Abuter, R., Amorim, A., Bauböck, M., Berger, J. P., Bonnet, H., Brandner, W., Clénet, Y., Coudé du Foresto, V., de Zeeuw, P. T., Dexter, J., Duvert, G., Eckart, A., Eisenhauer, F., Förster Schreiber, N. M., Garcia, P., Gao, F., Gendron, E., Genzel, R., Gerhard, O., Gillessen, S., Habibi, M., Haubois, X., Henning, T., Hippler, S., Horrobin, M., Jiménez-Rosales, A., Jocou, L., Kervella, P., Lacour, S., Lapeyrère, V., Le Bouquin, J.-B., Léna, P., Ott, T., Paumard, T., Perraut, K., Perrin, G., Pfuhl, O., Rabien, S., Rodriguez Coira, G., Rousset, G., Scheithauer, S., Sternberg, A., Straub, O., Straubmeier, C., Sturm, E., Tacconi, L. J., Vincent, F., von Fellenberg, S., Waisberg, I., Widmann, F., Wieprecht, E., Wiezorrek, E., Woillez, J., and Yazici, S., "A geometric distance measurement to the galactic center black hole with 0.3% uncertainty," *A&A*, vol. 625, p. L10, 2019.

[16] J. H. Oort, "Investigations concerning the rotational motion of the galactic system together with new determinations of secular parallaxes, precession and motion of the equinox (Errata: 4 94)," *bain*, vol. 4, p. 79, Sept. 1927.

[17] Lex Kaper, Jure Japelj, and Uddipta Bhardwaj, "Working lectures stellar populations 2020 – 2021."

[18] Gaia Collaboration and O. L. Creevey, et al., "Gaia data release 3: A golden sample of astrophysical parameters," 2022.

[19] J. H. Oort, "Observational evidence confirming Lindblad's hypothesis of a rotation of the galactic system," *bain*, vol. 3, p. 275, Apr. 1927.

[20] J. van Opijnen, “Locations and peculiar velocities of O-type runaway and field stars based on Gaia Data Release 2.” Bachelor thesis, June 2018.

Popular summary

The universe is a vast and magnificent place with many wonders and mysteries. Everyone has looked up at the sky during a clear night and wondered if we are special and why we are here. These great and complex questions have captivated humanity for thousands of years. To try to get a grip on these conundrums, the science of astronomy was developed. Many great minds have busied themselves with questions about our universe, it all started with just describing what we see when we look up, trying to predict motion of astral bodies and tying meaning to it all. Astronomy has made great leaps since then, but the fundamental questions are still the same, we are still trying to understand our spot in the universe and figure out if we are special. We have made marvellous machines to give us a better understanding of everything around us, but accomplishing this is no small feat. An example of how we are trying to understand the universe around us is a satellite launched into space in 2013, the GAIA satellite. The GAIA satellite hovers in space in the shadow of the earth and is constantly scanning the sky. With the satellite we are measuring the distance to many nearby stars. This way we can get a good picture of where we are and what is around us.

We still have some very fundamental questions about how the stars around us got to be where they are. Studying the youngest stars in our galaxy helps us to understand why and where stars are born and what they go through in the earliest stages of their lives. These young so called O-stars are rare and we know not many of them, but with help of GAIA satellite we can accurately study them and subsequently learn something about them. We see for instance that many O-stars are clustered together, possibly indicating locations of active star formation. And by looking at where they are going we get to understand what happens to them at an early age.

Studying and measuring these stars is no easy feat, there are many thing you have to account for. An example of this is something we call the Malmquist bias, named after the man that first described it. The Malmquist bias states that stars that are far away seem less bright than they actually are. If we are measuring these stars with a telescope that is not sensitive enough to pick out their dim light, then we will miss them, while brighter stars at the same distance will be found. This leads to a discrepancy with reality where it seems that the further away we look the less dim stars are present, which is of course not the case. Other such biasses exist, for instance difficulty arises when measuring how far away a star is. We are more likely to underestimate the distance to that star than to overestimate it. Overcoming these biasses can be done by mathematical analysis of the problem, even though we often have to make pretty bold assumptions to even get an answer as how to deal with them, leaving some inaccuracy in the final results.

A. O-stars

In this table we have included all 472 O-stars used for the data analysis, included are their GAIA EDR3 ID, name and the source of their radial velocity data if any.

Table A.1.: Table of the GAIA EDR3 ID's, names, and bibcodes of the sources of the radial velocity if any, for our O-star sample

Index	GAIA EDR3 ID	Name	Bibcode rv source
0	Gaia EDR3 168450545792009600	V* X Per	2007AN....328..889K
1	Gaia EDR3 181173402856344704	LS V +33 15	1972PASP...84..459J
2	Gaia EDR3 181174025630477056	HD 242935A	
3	Gaia EDR3 181174781544714880	BD+33 1025	1972PASP...84..459J
4	Gaia EDR3 181178389317231616	HD 242908	2006AstL...32..759G
5	Gaia EDR3 182071570715713024	V* AE Aur	2007AN....328..889K
6	Gaia EDR3 183431116844350976	HD 36483	2007AN....328..889K
7	Gaia EDR3 187469760492408960	HD 34656	2018A&A...613A..65H
8	Gaia EDR3 250748269580520320	BD+50 886	1974PDAO...14..283C
9	Gaia EDR3 251961924258399616	HD 24431	2018A&A...613A..65H
10	Gaia EDR3 273895291446188288	* 1 Cam A	
11	Gaia EDR3 339061352757048576	HD 14633	2004A&A...424..727P
12	Gaia EDR3 394976501194687488	V* AO Cas	2004A&A...424..727P
13	Gaia EDR3 423818404880147456	HD 5005A	
14	Gaia EDR3 423818404882207232	HD 5005C	
15	Gaia EDR3 423818409181768704	HD 5005D	
16	Gaia EDR3 427457895747434880	BD+60 134	1953GCRV..C.....0W
17	Gaia EDR3 429470895385555456	HD 225146	2018A&A...613A..65H
18	Gaia EDR3 429927879906030336	HD 225160	2006AstL...32..759G
19	Gaia EDR3 431569519486156032	HD 108	2006AstL...32..759G
20	Gaia EDR3 452357745305268224	HD 15137	2004A&A...424..727P
21	Gaia EDR3 454781584331773312	HD 16832	2018A&A...613A..65H
22	Gaia EDR3 454846489876889216	HD 16691	2018A&A...613A..65H
23	Gaia EDR3 456780633911740288	HD 13745	2006AstL...32..759G
24	Gaia EDR3 456993526850282240	HD 13268	2006AstL...32..759G
25	Gaia EDR3 457262769759711616	HD 15642	2006AstL...32..759G
26	Gaia EDR3 458310054588240384	HD 14434	1970ApJS...19..387A
27	Gaia EDR3 459034873267691264	HD 14947	2006AstL...32..759G

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
28	Gaia EDR3 459303222822598272	HD 14442	
29	Gaia EDR3 460625282472690048	HD 17603	2018A&A...613A..65H
30	Gaia EDR3 462130376453494272	V* CC Cas	1953GCRV..C.....0W
31	Gaia EDR3 464697873547937664	BD+60 586	1996BICDS..48...11F
32	Gaia EDR3 465485776704859520	BD+60 513	1977ApJ...214..759C
33	Gaia EDR3 465527523789596160	HD 15570	2006AstL...32..759G
34	Gaia EDR3 465531750036366080	BD+60 501	1977ApJ...214..759C
35	Gaia EDR3 465535048571192704	HD 15629	2018A&A...613A..65H
36	Gaia EDR3 465538793783729920	BD+60 498	2006AstL...32..759G
37	Gaia EDR3 465538828142411392	BD+60 499	1967IAUS...30..167U
38	Gaia EDR3 465551025849495296	BD+60 497	1932LicOB..16...53H
39	Gaia EDR3 466127062559750528	HD 18326	2007AN....328..889K
40	Gaia EDR3 466603670792030080	HD 18409	2006AstL...32..759G
41	Gaia EDR3 468876464407828096	HD 237211	2011AJ....142..146W
42	Gaia EDR3 469718003120510464	Hilt 412	1979IAUS...30...57E
43	Gaia EDR3 504561286085395712	HD 12323	2004A&A...424..727P
44	Gaia EDR3 506681213224331648	HD 12993	2018A&A...613A..65H
45	Gaia EDR3 506790923870949248	HD 13022	2006AstL...32..759G
46	Gaia EDR3 509891477946331264	BD+60 261	2011AJ....142..146W
47	Gaia EDR3 513431183821175936	V* DN Cas	
48	Gaia EDR3 514124013591624064	BD+62 424	1996BICDS..48...11F
49	Gaia EDR3 523835243884327680	Hilt 63	
50	Gaia EDR3 524109090999225344	HD 5689	
51	Gaia EDR3 528570015826682496	BD+66 1675	1974PDAO...14..283C
52	Gaia EDR3 528592796334452352	NGC 7822 29	
53	Gaia EDR3 528593517887714560	BD+66 1674	1974PDAO...14..283C
54	Gaia EDR3 528594342521399168	V* V747 Cep	
55	Gaia EDR3 969823376583544576	HD 41161	2006AstL...32..759G
56	Gaia EDR3 1821596755283891840	* 9 Sge	2006AstL...32..759G
57	Gaia EDR3 1865519977036620544	HD 201345	2006AstL...32..759G
58	Gaia EDR3 1869256701670871168	V* Y Cyg	2004A&A...424..727P
59	Gaia EDR3 1970299999201029888	HD 202124	2006AstL...32..759G
60	Gaia EDR3 1971467916750747264	* 68 Cyg	1953GCRV..C.....0W
61	Gaia EDR3 1995526747297893504	HD 218915	2006AstL...32..759G
62	Gaia EDR3 2003813319766437120	BD+55 2840	
63	Gaia EDR3 2003815312630659584	LS III +56 109	
64	Gaia EDR3 2005971180064658560	Hilt 1118	
65	Gaia EDR3 2006034745567935360	LS III +55 45	
66	Gaia EDR3 2006219738408337280	BD+55 2722	1974PDAO...14..283C
67	Gaia EDR3 2007416556811088768	HD 215835	1953GCRV..C.....0W

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
68	Gaia EDR3 2007418961992863744	LS III +57 90	1969A&A.....1..356U
69	Gaia EDR3 2009852456101233536	LS III +56 119	
70	Gaia EDR3 2010239999586617216	HD 218195A	
71	Gaia EDR3 2011867311156752256	BD+60 2635	
72	Gaia EDR3 2014149897293278848	BD+60 2522	1977ApJ...214..759C
73	Gaia EDR3 2014960718396258816	2MASS J23133024+6130103	
74	Gaia EDR3 2014960718396556160	TYC 4279-1463-1	
75	Gaia EDR3 2020089356034971904	HD 344782	
76	Gaia EDR3 2020106604616650496	HD 344777	
77	Gaia EDR3 2020407973910357376	HD 344758	
78	Gaia EDR3 2020856746444684032	HD 338931	
79	Gaia EDR3 2020947043840246144	HD 338916	1953GCRV..C.....0W
80	Gaia EDR3 2034587035353232768	HD 186980	2006AstL...32..759G
81	Gaia EDR3 2055609732001730304	HD 228779	2018yCat.1345....0G
82	Gaia EDR3 2057530097777247744	V* V382 Cyg	2004A&A...424..727P
83	Gaia EDR3 2058010275115738880	BD+36 4063	
84	Gaia EDR3 2058259383251666816	HD 227465	
85	Gaia EDR3 2059011376158159488	V* V448 Cyg	2006AstL...32..759G
86	Gaia EDR3 2059046010782859392	HD 191201A	
87	Gaia EDR3 2059070135632404992	HD 190864	2006AstL...32..759G
88	Gaia EDR3 2059130368252069888	HD 191612	2006AstL...32..759G
89	Gaia EDR3 2059383668236814720	HD 226868	2006AstL...32..759G
90	Gaia EDR3 2059434898612310656	HD 227018	2006AstL...32..759G
91	Gaia EDR3 2059455617540980736	HD 227245	1953GCRV..C.....0W
92	Gaia EDR3 2060647728648558464	HD 228766	1951ApJ...113..317H
93	Gaia EDR3 2060824750020658432	HD 192639	2006AstL...32..759G
94	Gaia EDR3 2061039704533786496	HD 193682	1979IAUS...30...57E
95	Gaia EDR3 2061089113838868480	V* V478 Cyg	1953GCRV..C.....0W
96	Gaia EDR3 2061130105002594304	HD 229234	1932LicOB..16...53H
97	Gaia EDR3 2061220333677386496	HD 229232	
98	Gaia EDR3 2061288365972432640	HD 228989	2006ApJ...648..580H
99	Gaia EDR3 2061349732467797504	HD 228841	2011AJ....142..146W
100	Gaia EDR3 2061405257789262080	HD 193595	1979IAUS...30...57E
101	Gaia EDR3 2061457553312825600	HD 193514	2018A&A...613A..65H
102	Gaia EDR3 2062426085622944256	V* V2011 Cyg	2004A&A...424..727P
103	Gaia EDR3 2062998828092126208	BD+36 4145	
104	Gaia EDR3 2064757698797394688	[CPR2002] A24	
105	Gaia EDR3 2067267304028045568	LS II +39 53	
106	Gaia EDR3 2067398416488472704	BD+40 4179	
107	Gaia EDR3 2067468407275542144	HD 229202	

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
108	Gaia EDR3 2067571937468635264	HD 229196	2018A&A...613A..65H
109	Gaia EDR3 2067780054401820544	GSC 03157-00327	2007ApJ...664.1102K
110	Gaia EDR3 2067781016474500864	Schulte 70	2007ApJ...664.1102K
111	Gaia EDR3 2067781905528395264	Schulte 22A	
112	Gaia EDR3 2067781909827710720	Schulte 51	
113	Gaia EDR3 2067781939888133248	RLP 853	2007ApJ...664.1102K
114	Gaia EDR3 2067782936320586240	Schulte 41	2007ApJ...664.1102K
115	Gaia EDR3 2067783623515353728	Schulte 9	2018A&A...613A..65H
116	Gaia EDR3 2067783799613328128	Schulte 24	2007ApJ...664.1102K
117	Gaia EDR3 2067784246289931776	Schulte 8C	2007ApJ...664.1102K
118	Gaia EDR3 2067784516868550016	RLP 666	2007ApJ...664.1102K
119	Gaia EDR3 2067784619950644480	BD+40 4227B	2007ApJ...664.1102K
120	Gaia EDR3 2067784624247057920	BD+40 4227A	2018yCat.1345....0G
121	Gaia EDR3 2067785002204178688	Schulte 8D	2007ApJ...664.1102K
122	Gaia EDR3 2067785070923663104	Schulte 7	2018A&A...613A..65H
123	Gaia EDR3 2067785208362826112	RLP 680	2007ApJ...664.1102K
124	Gaia EDR3 2067818533010251776	[CPR2002] B17	
125	Gaia EDR3 2067826062091598976	BD+40 4212	
126	Gaia EDR3 2067827466541470080	GOS G080.03+00.94 01	
127	Gaia EDR3 2067829596846026880	[CPR2002] A11	
128	Gaia EDR3 2067830941174418048	BD+40 4220	1979IAUS...30...57E
129	Gaia EDR3 2067832624801783040	GSC 03161-01218	2007ApJ...664.1102K
130	Gaia EDR3 2067832968398974208	Schulte 17	2007ApJ...664.1102K
131	Gaia EDR3 2067833204620693632	Schulte 16	2007ApJ...664.1102K
132	Gaia EDR3 2067833243277076864	BD+40 4221	2007ApJ...664.1102K
133	Gaia EDR3 2067834926904094848	Schulte 15	2007ApJ...664.1102K
134	Gaia EDR3 2067835682818357376	BD+40 4219	2018A&A...613A..65H
135	Gaia EDR3 2067840149584105344	RLP 1592	2007ApJ...664.1102K
136	Gaia EDR3 2067847502568383744	Schulte 20	2007ApJ...664.1102K
137	Gaia EDR3 2067862414694092544	Schulte 73	
138	Gaia EDR3 2067874028285659520	Schulte 27	2007ApJ...664.1102K
139	Gaia EDR3 2067883820810670336	RLP 335	2007ApJ...664.1102K
140	Gaia EDR3 2067887802243913216	Schulte 75	2007ApJ...664.1102K
141	Gaia EDR3 2067887840900094848	Schulte 29	2007ApJ...664.1102K
142	Gaia EDR3 2067888218857234304	BD+41 3807	2018A&A...613A..65H
143	Gaia EDR3 2068498409154126720	LS III +41 20	
144	Gaia EDR3 2068647191122709632	LS III +41 14	
145	Gaia EDR3 2068647191122712960	HD 228759	1974PDAO...14..283C
146	Gaia EDR3 2069819545390584192	BD+43 3654	2018yCat.1345....0G
147	Gaia EDR3 2070636448174993024	HD 195592	2006AstL...32..759G

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
148	Gaia EDR3 2071519321655107200	LS III +46 12	
149	Gaia EDR3 2071522345312074240	LS III +46 11	
150	Gaia EDR3 2074539817555034112	HD 191978	2018A&A...613A..65H
151	Gaia EDR3 2074735977289390080	HD 192001	2018A&A...613A..65H
152	Gaia EDR3 2074975395945857920	HD 191423	2007AN....328..889K
153	Gaia EDR3 2075283293565448704	HD 189957	2018A&A...613A..65H
154	Gaia EDR3 2082047179854088064	HD 191781	2018A&A...613A..65H
155	Gaia EDR3 2162889493831375488	[SL2008c] 8	
156	Gaia EDR3 2163075895429683072	HD 199579	2004A&A...424..727P
157	Gaia EDR3 2178158583618706560	HD 206183	2007PASP..119..742B
158	Gaia EDR3 2179492944050949760	HD 204827	1961PDAO...12....1P
159	Gaia EDR3 2199161386015469312	* 14 Cep	2004A&A...424..727P
160	Gaia EDR3 2199904449710846720	* lam Cep	2006AstL...32..759G
161	Gaia EDR3 2203434019472097664	HD 207538	2018A&A...613A..65H
162	Gaia EDR3 2204409870400823936	HD 209339	1953GCRV..C.....0W
163	Gaia EDR3 2205414029452444800	BD+62 2078	2018yCat.1345....0G
164	Gaia EDR3 2205471994328733568	HD 213023A	
165	Gaia EDR3 2205660290000075392	* 19 Cep	2006AstL...32..759G
166	Gaia EDR3 2207167239400168192	HD 216532	2006AstL...32..759G
167	Gaia EDR3 2207199739413172736	HD 216898	2018A&A...613A..65H
168	Gaia EDR3 2207291656008154112	HD 217086	2006AstL...32..759G
169	Gaia EDR3 2210120763852980736	BD+66 1661	2007AN....328..889K
170	Gaia EDR3 2216034525806770560	HD 207198	2006AstL...32..759G
171	Gaia EDR3 2901155648586891648	* mu. Col	2006AstL...32..759G
172	Gaia EDR3 2934257408221360384	TYC 5968-4011-1	
173	Gaia EDR3 3017360936587007616	* tet02 Ori A	1920PDO.....5....1H
174	Gaia EDR3 3017364063330718080	* tet01 Ori C	1972ApJ...174L..79C
175	Gaia EDR3 3044980222001242112	HD 53975	2006AstL...32..759G
176	Gaia EDR3 3045391130111207808	HD 54879	2007PASP..119..742B
177	Gaia EDR3 3046582725837564800	HD 54662	2006AstL...32..759G
178	Gaia EDR3 3047763601330845312	HD 57682	2006AstL...32..759G
179	Gaia EDR3 3047853142806752512	HD 55879	1953GCRV..C.....0W
180	Gaia EDR3 3100935811848497024	HD 52266	2007AN....328..889K
181	Gaia EDR3 3108642147107724544	HD 52533	2004A&A...424..727P
182	Gaia EDR3 3126650296443096192	HD 47432	2006AstL...32..759G
183	Gaia EDR3 3130004665899281024	HD 46573	2006AstL...32..759G
184	Gaia EDR3 3130556139698032512	HD 46485	2006AstL...32..759G
185	Gaia EDR3 3131327653265432704	HD 46056	1924PDAO....2..287P
186	Gaia EDR3 3131331054879415680	HD 46223	1977ApJ...214..759C
187	Gaia EDR3 3131335693444047232	HD 46150	2007PASP..119..742B

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
188	Gaia EDR3 3131384484272514688	HD 46106	1924PDAO....2..287P
189	Gaia EDR3 3131385205826979200	HD 46149	1977ApJ...214..759C
190	Gaia EDR3 3131898712120792064	HD 47129	1953GCRV..C.....0W
191	Gaia EDR3 3132668164099141120	HD 48099	2004A&A...424..727P
192	Gaia EDR3 3327820070594639104	TYC 737-1170-1	
193	Gaia EDR3 3345950879898371712	HD 41997	2007AN....328..889K
194	Gaia EDR3 3346795167389827712	HD 39680	1979IAUS...30...57E
195	Gaia EDR3 3356755643227139584	HD 45314	2007PASP..119..742B
196	Gaia EDR3 3372475738926018560	HD 44811	2006AstL...32..759G
197	Gaia EDR3 3375202836997950592	HD 42088	2018A&A...613A..65H
198	Gaia EDR3 3402983608883878016	HD 36879	2006AstL...32..759G
199	Gaia EDR3 3447570321516606720	HD 37366	2007AN....328..889K
200	Gaia EDR3 3455835109904777472	HD 37737	2004A&A...424..727P
201	Gaia EDR3 4054618559611164288	HD 159176	2004A&A...424..727P
202	Gaia EDR3 4055916430035297792	HD 161853	1972MNRAS.158...85C
203	Gaia EDR3 4057914208312575360	V* V1081 Sco	1979IAUS...30...57E
204	Gaia EDR3 4062481614002853120	HD 164019	2006AstL...32..759G
205	Gaia EDR3 4066021251120282496	NAME Her 36SE	1953GCRV..C.....0W
206	Gaia EDR3 4066022591147527552	* 9 Sgr	1977ApJ...214..759C
207	Gaia EDR3 4066023415781257984	HD 164816	
208	Gaia EDR3 4066063754109686016	HD 165246	
209	Gaia EDR3 4066064956700837248	HD 165052	2019MNRAS.486.2477W
210	Gaia EDR3 4066278846098710016	* 11 Sgr	2006AstL...32..759G
211	Gaia EDR3 4069268658691324672	HD 164492	1977ApJ...214..759C
212	Gaia EDR3 4069523951550440832	HD 163800	2018A&A...613A..65H
213	Gaia EDR3 4069547797210386816	HD 163892	2007AN....328..889K
214	Gaia EDR3 4069585661644839808	HD 313846	1997A&A...317..532C
215	Gaia EDR3 4086465432879867264	HD 175754	1953GCRV..C.....0W
216	Gaia EDR3 4091082415320078848	TYC 6277-1821-1	
217	Gaia EDR3 4094208193745478144	HD 166546	2006AstL...32..759G
218	Gaia EDR3 4094790350826867584	HD 167659	1972AJ.....77..138A
219	Gaia EDR3 4094879303848709376	HD 167771	2004A&A...424..727P
220	Gaia EDR3 4095308873632903168	HD 164438	2007AN....328..889K
221	Gaia EDR3 4095650271976366592	HD 167411	1953GCRV..C.....0W
222	Gaia EDR3 4097600496363078912	HD 167633	2018A&A...613A..65H
223	Gaia EDR3 4097815382164899840	NGC 6618 337	
224	Gaia EDR3 4098005975631742720	2MASS J18210223-1601010	
225	Gaia EDR3 4098127712190306688	LS IV -15 42	
226	Gaia EDR3 4103751095772700032	HD 171589	2018A&A...613A..65H
227	Gaia EDR3 4104196427943626624	V* V479 Sct	

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
228	Gaia EDR3 4104201586232296960	BD-14 5040	
229	Gaia EDR3 4146592707257808384	V* QR Ser	2007AN....328..889K
230	Gaia EDR3 4146599166888812032	HD 168075	1977ApJ...214..759C
231	Gaia EDR3 4146599682284818176	HD 168137	1932LicOB..16...53H
232	Gaia EDR3 4146612494169472256	BD-13 4927	2005A&A...437..467E
233	Gaia EDR3 4146612872133107584	2MASS J18183749-1343393	2005A&A...437..467E
234	Gaia EDR3 4147125180140397440	HD 165319	2006AstL...32..759G
235	Gaia EDR3 4150272944490967680	LS IV -13 3	
236	Gaia EDR3 4152407474508172160	HD 168504	2005A&A...437..467E
237	Gaia EDR3 4153614871414364288	HD 168461	
238	Gaia EDR3 4153657344390915200	BD-12 4979	
239	Gaia EDR3 4153666106124493696	HD 168112	2006AstL...32..759G
240	Gaia EDR3 4153675108375783040	LS IV -12 12	
241	Gaia EDR3 4153933111306297600	BD-11 4586	2018A&A...613A..65H
242	Gaia EDR3 4154232835623350656	BD-10 4682	
243	Gaia EDR3 4155344545012059520	HD 173010	1979IAUS...30...57E
244	Gaia EDR3 4156007447403090304	HD 169582	2018A&A...613A..65H
245	Gaia EDR3 4156616194555562624	BD-08 4617	
246	Gaia EDR3 4157061363634092160	HD 166734	2006AstL...32..759G
247	Gaia EDR3 4165963456443515520	HD 157857	2006AstL...32..759G
248	Gaia EDR3 4252558380933360000	HD 172175	1979IAUS...30...57E
249	Gaia EDR3 4256830086697893760	BD-04 4503	
250	Gaia EDR3 4310497932106715520	V* V1182 Aql	1979IAUS...30...57E
251	Gaia EDR3 4468233187335913472	V* V986 Oph	2006AstL...32..759G
252	Gaia EDR3 5237207155697930496	V* TU Mus	2004A&A...424..727P
253	Gaia EDR3 5241775145481099904	HD 95589	
254	Gaia EDR3 5241879697825137280	HD 94963	2018A&A...613A..65H
255	Gaia EDR3 5254222888417319424	HD 93222	2018A&A...613A..65H
256	Gaia EDR3 5254262161604257408	HD 93028	2018A&A...613A..65H
257	Gaia EDR3 5254268518156437888	HD 93027	2018A&A...613A..65H
258	Gaia EDR3 5254269514589043200	CPD-59 2554	
259	Gaia EDR3 5254269613320423040	CD-59 3274	2018A&A...613A..65H
260	Gaia EDR3 5254269617668284416	HD 93146B	
261	Gaia EDR3 5254269961265754368	HD 305536	
262	Gaia EDR3 5254275630622585344	HD 305438	
263	Gaia EDR3 5254407262783529472	HD 91651	2007AN....328..889K
264	Gaia EDR3 5254408151786430336	HD 91837	
265	Gaia EDR3 5254478593582508288	HD 92607	
266	Gaia EDR3 5255031579201240064	HD 89625	
267	Gaia EDR3 5255042402520416128	HD 90087	1979IAUS...30...57E

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
268	Gaia EDR3 5255424345369668608	GSC 08612-00589	
269	Gaia EDR3 5255667681036173568	SS 215	
270	Gaia EDR3 5312859946571764352	HD 298429	2018A&A...613A..65H
271	Gaia EDR3 5324644031165449600	HD 76968	2018A&A...613A..65H
272	Gaia EDR3 5325671971449181696	CD-49 4263	
273	Gaia EDR3 5328579836109543040	CD-47 4550	
274	Gaia EDR3 5328579939188867200	CD-47 4551	2018A&A...613A..65H
275	Gaia EDR3 5328780634428211456	HD 76556	2011AJ....142..146W
276	Gaia EDR3 5329858843022355072	CPD-45 2920	2018A&A...616A..40C
277	Gaia EDR3 5329862966191325312	HD 74920	
278	Gaia EDR3 5333575329765910400	HD 101436	1977ApJ...214..759C
279	Gaia EDR3 5333575398489755520	HD 101413	1977ApJ...214..759C
280	Gaia EDR3 5333576326198296064	HD 101298	1977ApJ...214..759C
281	Gaia EDR3 5333579556013463936	HD 101191	2011AJ....142..146W
282	Gaia EDR3 5333580930403062016	HD 308813	1979IAUS...30...57E
283	Gaia EDR3 5333860240705973888	HD 99897	2006AstL...32..759G
284	Gaia EDR3 5333957788009701760	HD 101223	2018A&A...613A..65H
285	Gaia EDR3 5333958578283024000	HD 101190	1973A&AS....9...85H
286	Gaia EDR3 5334960920601501184	HD 102415	
287	Gaia EDR3 5336162721186272512	HD 99546	2006AstL...32..759G
288	Gaia EDR3 5337239658437912960	V* EM Car	1964POHP....7...33M
289	Gaia EDR3 5337249828921901312	HD 97319	1979IAUS...30...57E
290	Gaia EDR3 5337284326102045568	HD 96264	2018A&A...613A..65H
291	Gaia EDR3 5337306316337270144	HD 96946	2018A&A...613A..65H
292	Gaia EDR3 5337418015513337472	NGC 3603 18	
293	Gaia EDR3 5337418015513340160	NGC 3603 47	
294	Gaia EDR3 5337637303697246976	HD 97434	1979IAUS...30...57E
295	Gaia EDR3 5337669189536576000	HD 97253	1979IAUS...30...57E
296	Gaia EDR3 5337719045521376640	HD 97166	1977ApJ...214..759C
297	Gaia EDR3 5337740589078804736	HD 96715	2018A&A...613A..65H
298	Gaia EDR3 5337747731560719616	HD 96670	2004A&A...424..727P
299	Gaia EDR3 5337974815066189184	THA 35-II-153	
300	Gaia EDR3 5337975983300032512	UCAC4 145-065595	
301	Gaia EDR3 5338275673288599936	HD 93843	2018A&A...613A..65H
302	Gaia EDR3 5338453278772130688	HD 303492	2018A&A...613A..65H
303	Gaia EDR3 5338501554205059072	HD 96622	2018A&A...613A..65H
304	Gaia EDR3 5338837553741386752	HD 94370A	
305	Gaia EDR3 5339352438763921536	HD 97848	2018A&A...613A..65H
306	Gaia EDR3 5340725660047993600	HD 96917	2007AN....328..889K
307	Gaia EDR3 5350279213497594496	HD 305619	

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
308	Gaia EDR3 5350289865017121024	HD 305612	
309	Gaia EDR3 5350289933736400000	HD 93632	
310	Gaia EDR3 5350290724010326784	HD 93576	
311	Gaia EDR3 5350296840043687552	HD 305539	
312	Gaia EDR3 5350301306809073280	HD 305532	
313	Gaia EDR3 5350302097082779136	CD-59 3304	
314	Gaia EDR3 5350306082812792576	CPD-59 2673	
315	Gaia EDR3 5350308865923993600	Cl* Trumpler 16 MJ 568	
316	Gaia EDR3 5350310480858998656	V* V662 Car	
317	Gaia EDR3 5350310824456389504	HD 93343	
318	Gaia EDR3 5350310927535609728	V* V731 Car	
319	Gaia EDR3 5350311339852481792	CD-59 3310	
320	Gaia EDR3 5350311374212228736	CD-59 3312	
321	Gaia EDR3 5350347520660979840	HD 305523	
322	Gaia EDR3 5350349101209092224	HD 305524	
323	Gaia EDR3 5350350303799853056	V* V661 Car	
324	Gaia EDR3 5350353567958779392	HD 305518	
325	Gaia EDR3 5350355251602278528	HD 303316	
326	Gaia EDR3 5350356007521809664	CPD-59 2591	2006ApJ...648..580H
327	Gaia EDR3 5350356419833915904	CD-59 3300	2018A&A...613A..65H
328	Gaia EDR3 5350357205782177664	HD 93204	2018A&A...613A..65H
329	Gaia EDR3 5350357313186767104	V* V560 Car	1995A&AS..114..269D
330	Gaia EDR3 5350357519345171200	HD 93162	
331	Gaia EDR3 5350357519345176192	Cl Trumpler 16 244	
332	Gaia EDR3 5350357725503668096	CPD-59 2629	2011AJ....142..146W
333	Gaia EDR3 5350357725503681664	CPD-59 2626	2006ApJ...648..580H
334	Gaia EDR3 5350358069101053184	V* V572 Car	
335	Gaia EDR3 5350358343979094912	CPD-59 2627	2006ApJ...648..580H
336	Gaia EDR3 5350358481418098944	CPD-59 2628	
337	Gaia EDR3 5350358683250920704	HD 303308	2018A&A...613A..65H
338	Gaia EDR3 5350362638946360960	CPD-58 2625	2006ApJ...648..580H
339	Gaia EDR3 5350362982528878976	HD 93160	2018A&A...613A..65H
340	Gaia EDR3 5350362982543827456	HD 93161A	
341	Gaia EDR3 5350362982543828352	HD 93161B	
342	Gaia EDR3 5350363326141056768	CPD-58 2608A	
343	Gaia EDR3 5350363807162637696	HD 93128	2018A&A...613A..65H
344	Gaia EDR3 5350363841537246592	Cl Trumpler 14 21	
345	Gaia EDR3 5350363875896996480	CD-58 3526	
346	Gaia EDR3 5350363910241888768	Cl Trumpler 14 9	2018A&A...613A..65H
347	Gaia EDR3 5350363910256783488	HD 93129A	1979IAUS...30...57E

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
348	Gaia EDR3 5350363910256783744	HD 93129B	
349	Gaia EDR3 5350363944616553216	CD-58 3529	
350	Gaia EDR3 5350376004884719232	V* V725 Car	
351	Gaia EDR3 5350383460949215232	HD 93250	2018A&A...613A..65H
352	Gaia EDR3 5350383529668697472	HD 303311	2018A&A...613A..65H
353	Gaia EDR3 5350387549760231808	CPD-58 2627	
354	Gaia EDR3 5350388821068542848	TYC 8626-2506-1	
355	Gaia EDR3 5350395383778733568	HD 93249	2018A&A...613A..65H
356	Gaia EDR3 5350396620729366656	HD 93190	
357	Gaia EDR3 5350671494307479808	HD 92206A	
358	Gaia EDR3 5350671498620168448	HD 92206B	
359	Gaia EDR3 5350908374676231808	HD 94024	2018A&A...613A..65H
360	Gaia EDR3 5351209808330136832	HD 92504	2007AN....328..889K
361	Gaia EDR3 5351451529087848832	HD 91824	2018A&A...613A..65H
362	Gaia EDR3 5351457202768029952	HD 91572	2018A&A...613A..65H
363	Gaia EDR3 5351703390282380800	WR 21a	
364	Gaia EDR3 5351758846903542400	HD 90273	1972MNRAS.158...85C
365	Gaia EDR3 5358486586756759040	HD 89137	2006AstL...32..759G
366	Gaia EDR3 5522306019626566528	HD 74194	2006AstL...32..759G
367	Gaia EDR3 5523764178195480448	HD 75211	2018A&A...613A..65H
368	Gaia EDR3 5524337882746365952	HD 76341	2006AstL...32..759G
369	Gaia EDR3 5524580737379318016	HD 75759	1979IAUS...30...57E
370	Gaia EDR3 5525970966757975808	HD 71304	2018A&A...613A..65H
371	Gaia EDR3 5541462673279047168	HD 69106	2007AN....328..889K
372	Gaia EDR3 5541472465805985024	HD 68450	1977ApJ...214..759C
373	Gaia EDR3 5543181519189083776	HD 69464	2018A&A...613A..65H
374	Gaia EDR3 5546501254035205376	CD-34 4496	
375	Gaia EDR3 5601917842664443520	CD-26 5136	2011AJ....142..146W
376	Gaia EDR3 5602027755165929344	CPD-26 2704	
377	Gaia EDR3 5602033390154015744	HD 64568	2006AstL...32..759G
378	Gaia EDR3 5619742506693999488	HD 57236	2018A&A...613A..65H
379	Gaia EDR3 5625488726258364544	HD 75222	2007AN....328..889K
380	Gaia EDR3 5853581485795684992	HD 123008	2018A&A...613A..65H
381	Gaia EDR3 5854087364193927168	HD 123590	
382	Gaia EDR3 5856117406232328832	HD 105056	2018A&A...613A..65H
383	Gaia EDR3 5858915766471941248	* tet Mus B	
384	Gaia EDR3 5858981908977997312	V* V340 Mus	
385	Gaia EDR3 5862482685215601664	V* V961 Cen	2006AstL...32..759G
386	Gaia EDR3 5864726788439265024	HD 120678	
387	Gaia EDR3 5864992114393344128	HD 118198	

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
388	Gaia EDR3 5865350383360485120	HD 117797	1957MmRAS..68....1F
389	Gaia EDR3 5866059156028925312	HD 122313	
390	Gaia EDR3 5866651998931097472	HD 125241	2006AstL...32..759G
391	Gaia EDR3 5868478390840987776	HD 115455	2019AJ....158...46B
392	Gaia EDR3 5869057421174852480	HD 117490	
393	Gaia EDR3 5869791585712427776	HD 116282	1979IAUS...30...57E
394	Gaia EDR3 5870885427979725824	HD 120521	2006AstL...32..759G
395	Gaia EDR3 5875491419647525248	* del Cir	2004A&A...424..727P
396	Gaia EDR3 5875879684681370624	HD 135591	2018A&A...613A..65H
397	Gaia EDR3 5878555174390220416	CD-59 5309	
398	Gaia EDR3 5881100234598293248	HD 130298	2007AN....328..889K
399	Gaia EDR3 5885280745250903680	LS 3386	
400	Gaia EDR3 5885668499206748160	GEN# +6.20109123	
401	Gaia EDR3 5897456668300003584	HD 124979	1979IAUS...30...57E
402	Gaia EDR3 5939657509457702144	CD-46 11017	1979IAUS...30...57E
403	Gaia EDR3 5940106213259136000	HD 151018	2007AN....328..889K
404	Gaia EDR3 5940954177259978880	HD 150136	2004A&A...424..727P
405	Gaia EDR3 5940954898814487168	HD 150135	1977ApJ...214..759C
406	Gaia EDR3 5941151986262863104	HD 148937	1979IAUS...30...57E
407	Gaia EDR3 5942786616486903296	HD 149452	
408	Gaia EDR3 5943084201826281984	HD 150574	
409	Gaia EDR3 5944076481085129984	HD 328209	2006AstL...32..759G
410	Gaia EDR3 5950908193476521856	HD 154811	2018A&A...613A..65H
411	Gaia EDR3 5951407646637937280	HD 155756	
412	Gaia EDR3 5953634393855997824	HD 156292	
413	Gaia EDR3 5953699131931631232	HD 155913	
414	Gaia EDR3 5958243001205162112	HD 161807	2006AstL...32..759G
415	Gaia EDR3 5964025538960495744	HD 329100	1972MNRAS.158...85C
416	Gaia EDR3 5965009601862125056	HD 152147	2018A&A...613A..65H
417	Gaia EDR3 5965963252092372608	HD 326775	
418	Gaia EDR3 5966450786728068736	HD 152424	2018A&A...613A..65H
419	Gaia EDR3 5966503941246144512	CD-41 11027A	
420	Gaia EDR3 5966509022176616448	CD-41 11037	1974AJ.....79.1271H
421	Gaia EDR3 5966509095207261056	CD-41 11042	1983ApL....23..183L
422	Gaia EDR3 5966509129566636160	HD 326329	1983ApL....23..183L
423	Gaia EDR3 5966509267005564288	HD 326331	1983ApL....23..183L
424	Gaia EDR3 5966509438804267264	HD 152314	1983ApL....23..183L
425	Gaia EDR3 5966509748041919616	HD 152249	1983ApL....23..183L
426	Gaia EDR3 5966509885480902656	HD 152248	2004A&A...424..727P
427	Gaia EDR3 5966510229078336768	HD 152233	1953GCRV..C.....0W

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
428	Gaia EDR3 5966515829715666304	HD 152200	1983ApL....23..183L
429	Gaia EDR3 5966519162610481664	HD 152003	1977ApJ...214..759C
430	Gaia EDR3 5966522667303699072	HD 152218	1974AJ.....79.1271H
431	Gaia EDR3 5966524419650392832	HD 152247	1953GCRV..C.....0W
432	Gaia EDR3 5966563486675142912	HD 152408	2018yCat.1345....0G
433	Gaia EDR3 5967079569931393024	HD 322417	2018A&A...613A..65H
434	Gaia EDR3 5968004912078306560	HD 151515	2018A&A...613A..65H
435	Gaia EDR3 5968259964433588992	HD 151003	
436	Gaia EDR3 5968468321845736192	HD 149404	2004A&A...424..727P
437	Gaia EDR3 5969564157008128512	HD 152246	
438	Gaia EDR3 5969686958696979712	HD 152405	2018A&A...613A..65H
439	Gaia EDR3 5969715271123841280	V* V1297 Sco	2018A&A...613A..65H
440	Gaia EDR3 5970357660807157632	HD 153426	1979IAUS...30...57E
441	Gaia EDR3 5972658934309986560	Cl HM 1 18	
442	Gaia EDR3 5972659655864062464	Cl HM 1 9	
443	Gaia EDR3 5972659655864503296	Cl HM 1 8	
444	Gaia EDR3 5972660411778740352	Cl HM 1 19	
445	Gaia EDR3 5972660622267521152	Cl HM 1 6	
446	Gaia EDR3 5972812973334855424	HD 323110	
447	Gaia EDR3 5973477318881351168	V* V1012 Sco	2007AN....328..889K
448	Gaia EDR3 5975270794119902080	HD 156738	2018A&A...613A..65H
449	Gaia EDR3 5975275943790057472	VdBH 86c	
450	Gaia EDR3 5975278903017810560	VdBH 86a	1972MNRAS.158...85C
451	Gaia EDR3 5975286256001980416	HD 319702	2018A&A...613A..65H
452	Gaia EDR3 5975481144442221696	HD 319699	2018A&A...613A..65H
453	Gaia EDR3 5975542648373731072	HD 156154	2018A&A...613A..65H
454	Gaia EDR3 5976033305433336576	2MASS J17252916-3425157	
455	Gaia EDR3 5976050725820752512	2MASS J17242895-3414506	
456	Gaia EDR3 5976050966338916864	2MASS J17244229-3413213	
457	Gaia EDR3 5976051065119174144	2MASS J17243603-3414004	
458	Gaia EDR3 5976051206857087232	2MASS J17244328-3412439	
459	Gaia EDR3 5976057082372349568	Cl Pismis 24 17	
460	Gaia EDR3 5976057288530788224	2MASS J17244578-3409399	
461	Gaia EDR3 5976382915813535232	HD 153919	2018yCat.1345....0G
462	Gaia EDR3 5977233658593947520	HD 154368	2018A&A...613A..65H
463	Gaia EDR3 5977345710074220800	HD 154643	2018A&A...613A..65H
464	Gaia EDR3 5978778884750785280	TYC 7370-460-1	
465	Gaia EDR3 5979008205986483712	HD 155806	2006AstL...32..759G
466	Gaia EDR3 5983259364592718080	HD 144647	1979IAUS...30...57E
467	Gaia EDR3 6018555337108830080	HD 148546	2006AstL...32..759G

Continued on next page

Table A.1 – continued from previous page

Index	GAIA EDR3 ID	Name	Bibcode rv source
468	Gaia EDR3 6056534942662447872	HD 110360	1979IAUS...30...57E
469	Gaia EDR3 6057486703068161792	HD 105627	2007AN....328..889K
470	Gaia EDR3 6071508298145817344	V* AB Cru	2004A&A...424..727P
471	Gaia EDR3 6072058878595295488	HD 104565	2018A&A...613A..65H

B. Code

GAIA ADQL code to retrieve the data on the O-stars:

Listing B.1: ADQL example

```
select ra,dec
from gaiaedr3.gaia_source
where source_id in
(...) and parallax is not null and parallax_over_error>10
```

where "(...)" represents a list of GAIA EDR3 ID's.