STELLAR DISTRIBUTION FUNCTIONS IN THE SOLAR NEIGHBOURHOOD USING STAR SAMPLES SUCH AS THE HIPPARCOS DATABASE

K. Van Caelenberg[1,2], **H. Dejonghe**[1]

[1]Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281–S9,B-9000 Gent, Belgium
[2]Research Assistant FWO

ABSTRACT

The Hipparcos catalogue provides a large amount of high quality positional and kinematical data of stars in the solar neighbourhood. In principle, these data can be used to construct a distribution function for the orbits of the stars, assuming three integrals of the motion and an axisymmetric potential of Stäckel type. Since orbits are not local, the construction of a distribution function is the better conditioned the larger the volume covered by the data.

On galactic scales however, the Hipparcos data cover a fairly restricted volume. To assess to what extent the distribution function derived from these data will be useful to determine a (unknown) distribution function, we generated Hipparcos-like artificial data samples coming from a known (global) distribution function, and investigate how well this distribution function can be recovered. The distribution function is constructed by fitting an analytical, but flexible form of the distribution function to the data.

Key words: Milky Way; dynamics; solar neighbourhood.

1. INTRODUCTION

The construction of dynamical models for galaxies is not an easy task. Since most of the galaxies are found at large distances from us, only projected quantities can be observed, while observational limitations such as seeing and pixel size give rise to observational and interpretative difficulties.

Our Galaxy on the other hand offers the opportunity to study a member of the galaxy population in greater detail. It is possible to obtain data on individual stars, but the problems here are the quality of the data needed for dynamical modeling (positions, distances, velocities), the (relatively) small volume that can be surveyed, and the number of stars for which such data are available. Only high quality positional and kinematical data of large numbers of stars will be useful in the construction of a dynamical model, relevant for the solar neighbourhood. With the Hipparcos catalogue, data that fulfill these requirements are available for the first time.

The model for our Galaxy we are going to use, is based on a distribution function, which is the probability of finding a star (with given spectral properties) at a given position and with a given velocity, analogous to a Maxwell-Boltzmann equation for a gas. Since stars move on orbits, this distribution function is equivalent to the probability densities for the various orbits.

A distribution function derived from the Hipparcos data will give information on the orbits in the Galaxy that pass through the solar neighbourhood. We want to assess to what extent such distribution function will be useful to determine a (unknown) distribution function. We do this by generating Hipparcos-like artificial data samples, drawn from a known global distribution function, and investigate how well this distribution function can be recovered.

Section 2 first gives a short introduction to dynamical models in general, and then describes the applied fitting procedure. In Section 3, we give a description of the model used for the simulations. Section 4 discusses the generation of artificial samples, and the results of the fitting procedure. Conclusions are given in Section 5.

2. DYNAMICAL MODELS AND FITTING PROCEDURE

2.1. Dynamical Models

The 'elementary particle' of a dynamical model is an orbit. A general classification of orbits is not obvious, since we need quantities that can be attached to a particular orbit. Such quantities are integrals of the motion, which are functions of coordinates and velocities that, by definition, remain constant along an orbit. In an axisymmetric system, two of such integrals are the binding energy $E = -V(\varpi,z) - \frac{1}{2}v^2$, with $V(\varpi,z)$ the gravitational potential, and L_z, the z-component of the angular momentum. In general, we do not expect more than 3 such integrals. Stronger yet, from a theoretical point of view, there is no reason for the existence of a third integral. In reality

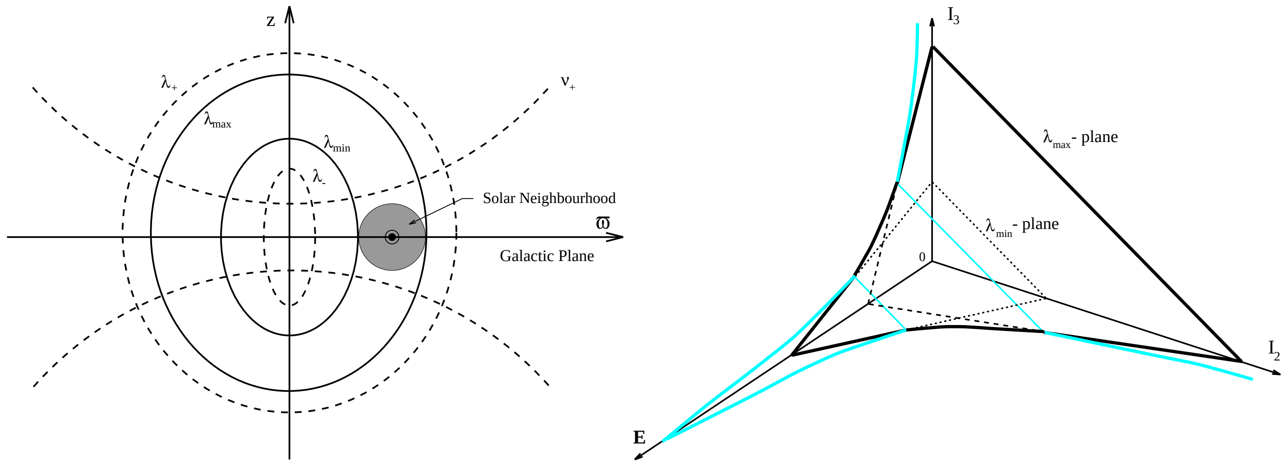

Figure 1. Left panel: Physical space. We consider the kinematics of stars in a sphere with radius 1 kpc, centered around the Sun, at 8 kpc from the Galactic Center (not drawn to scale!). The parameters λ_- and λ_+ denote the bounding spheroids, ν_+ the bounding hyperboloid (spheroidal coordinates). Orbits can only visit the solar neighbourhood if $\lambda_- \leq \lambda_{\max}$ and $\lambda_+ \geq \lambda_{\min}$. Right panel: Three-integral space. In all points 'behind' the black surface we can determine the distribution function when we have kinematic data in the grey volume (left panel). Such data cannot determine the distribution function in the volumes between the grey and the black surface.

however, there is strong evidence that most orbits obey a third integral, almost all the time.

Stäckel potentials (de Zeeuw 1985) have the advantage that all orbits have 3 integrals, and that the elusive third integral I_3 is a simple quadratic function of the velocities. All orbits are uniquely determined by their 3 integrals (E, I_2, I_3). Since all orbits are known in such a potential, we know exactly which orbits pass through the solar neighbourhood and which don't. Therefore, we can know in principle to what degree samples such as the Hipparcos sample determine the distribution function of stars of a particular type.

What happens, for instance, when we only consider the kinematics of stars, located within a certain distance from the Sun? In general, orbits in Stäckel potentials are bounded by 2 spheroids, indicated in spheroidal coordinates by λ_- and λ_+, and one hyperboloid of 2 sheets, indicated by ν_+ (Figure 1, left panel). Denoting the 2 spheroids bounding our Sun-centered volume as $\lambda_{\min}$ and $\lambda_{\max}$, only orbits with $\lambda_- \leq \lambda_{\max}$ and $\lambda_+ \geq \lambda_{\min}$ visit the solar neighbourhood. It is only for these orbits that the distribution functions that we will determine contain valid information, if the construction of it was based on kinematical data of stars in the (λ_-,λ_+)-volume. (The volume is also bounded by a hyperboloid, which could be indicated by $\nu_{\max}$, but this bound does not place any restrictions on the accessible orbits, because of the position of the Sun in the Galactic Plane.)

Figure 1 shows how the bounds $\lambda_{\min}$ and $\lambda_{\max}$ in physical space translate into planes in three integral space, thereby restricting the volume in three integral space inside which the distribution function can be recovered.

2.2. The Fitting Procedure

During the fitting procedure, the real distribution function will be approximated by a sum of modified Fricke components. These are an extension of simple powers in the integrals of the motion E, L_z (or I_2) and I_3, designed to reproduce certain features present in the data (bulge-like components, thin disks, etc. see Van Caelenberg & Dejonghe (1997)). The coefficients of the terms in the sum are determined by a Quadratic Programming method, that fits to the observed data by minimizing a χ^2-variable. During the process, the positivity of the resulting distribution function is assured. A full explanation of this technique can be found in Dejonghe (1989).

The data needed by the program are typically moments of the distribution function such as mass density, velocity components and velocity dispersions.

The advantages of this approach are threefold.

1. By using this procedure, useful results can be obtained from incomplete kinematical information: one does not necessarily need all three velocity components together with the positions, which would be necessary in order to construct the orbit. Of course, the more complete the data, the better! In our case, the kinematical information we used in the simulations consisted in proper motions, without radial velocities, and the 2 corresponding velocity dispersions (Section 4).

2. Axisymmetric distribution functions of the kind we consider work well for the older stellar populations (e.g. Dejonghe 1992, Durand et al. 1996, Sevenster et al. 1995).

3. Working in three integral space implies that the solution space can be 3-dimensional, instead of 6-dimensional (in phase space) if one uses direct orbit calculations.

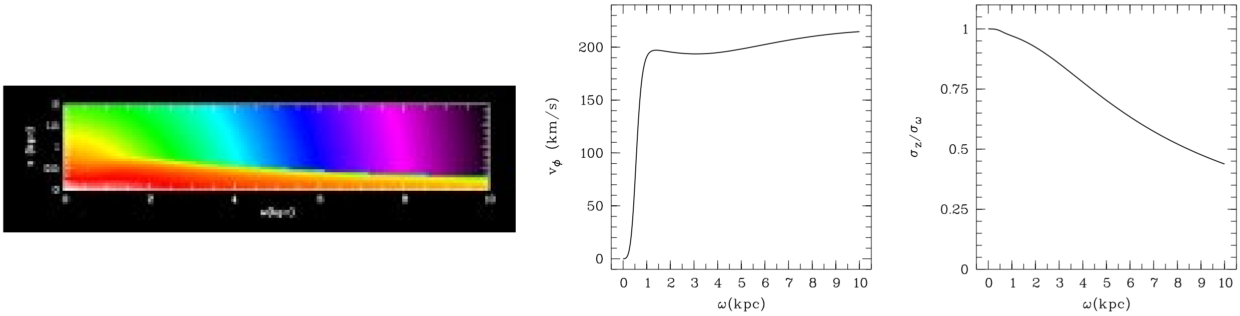

Figure 2. Some of the characteristics of the adopted distribution function, with (ϖ,ϕ,z) cylindrical coordinates and the z coordinate perpendicular to the Galactic plane. The Sun is at 8 kpc from the Galactic Center. Left panel: Log of the mass density in the meridional plane. Middle panel: Rotational velocity v_ϕ in the Galactic Plane. Right panel: Dispersions σ_ϖ and σ_z for the velocity components v_ϖ and v_z, with $\sigma_\varpi(0) = \sigma_z(0) = 90$ km s^{-1} and $\sigma_\varpi(10) = 9$ km s^{-1}.

3. THE MODEL USED FOR SIMULATIONS

For the potential we have chosen an axisymmetric Stäckel potential with a halo-disk structure, constructed to be compatible with all known parameters of the Galactic potential in the solar neighbourhood (Batsleer & Dejonghe 1994). The distribution function we used to generate our data points consists of two components with different characteristics. The first (and most important) one represents a thin 3-integral stellar disk, the second one is a model for the bulge. As can be seen from Figure 2, the distribution function reproduces (1) a bulge-disk galaxy, (2) a flat rotation curve, with a rotation of about 210 km s^{-1} in the solar neighbourhood, and (3) a ratio $\sigma_z/\sigma_\varpi = 1/2$ in the solar neighbourhood, with σ_z and σ_ϖ the velocity dispersions in the z and ϖ direction respectively, and (ϖ,ϕ,z) cylindrical coordinates.

4. SAMPLE GENERATION AND RESULTS OF THE FITTING PROCEDURE

In order to investigate to what degree stellar samples, extracted from the Hipparcos database, make it possible to recover (a part of) the distribution function of the Galaxy, we will have to generate samples with the same kind of data and the same characteristics as present in the Hipparcos data.

At any given position (specified by galactic longitude l, galactic latitude b and distance r from the sun), the adopted distribution function enables us to specify:

- the number density $\rho(l,b,r)$;
- the mean streaming in proper motion parallel to galactic longitude, $\mu_l(l,b,r)$, and galactic latitude $\mu_b(l,b,r)$;
- the corresponding velocity dispersions σ_l and σ_b.

A typical data sample consists of a set of positions, with the corresponding densities, proper motions and velocity dispersions. Any star sample drawn from such a distribution function will produce values that deviate from the exact ones because of statistical fluctuations. These fluctuations are taken to be Gaussian, and are accounted for in our samples. We suppose that $n_\star$ stars are available in the immediate neighbourhood of any given position. The obtained data were then used by the Quadratic Programming fitting routines to construct a fit.

As for the samples, we considered two different cases.

4.1. Case 1: Stars Uniformly Distributed in a Sphere Around the Sun

For the first series of samples, we suppose we are dealing with data from stars uniformly sampled in a sphere with radius 1 kpc around the Sun, and thus without any sampling bias. We investigated the difference between the original distribution function and the fit, as a function of the available number of stars $n_\star$ per position. We consider 40 positions, or about 1 every 460 pc, with $n_\star$ ranging from 20 to 1000. Because of the restriction of uniform coverage and homogeneity in astrophysical properties, the total number of stars was always kept much smaller than the number of stars available in the Hipparcos catalogue.

Figure 3 gives us an idea of the goodness of the fits, with increasing number of stars per position. The differences shown do only apply to the deviation between the original distribution function and the fit in the solar neigbourhood. Clearly, the fits become better with increasing number of stars per position, as could be expected. An unbiased and volume-limited sample will reproduce our distribution function when there are more than 100 sampled stars per position. This assumes that the velocity distribution does not vary too much with position.

4.2. Case 2: Stars Uniformly Distributed Within a Limited Area

For the second series of samples, the stars are uniformly distributed as well, but only within a limited area on the sky, centered around the direction of the Galactic Center. All stars are still within 1 kpc from

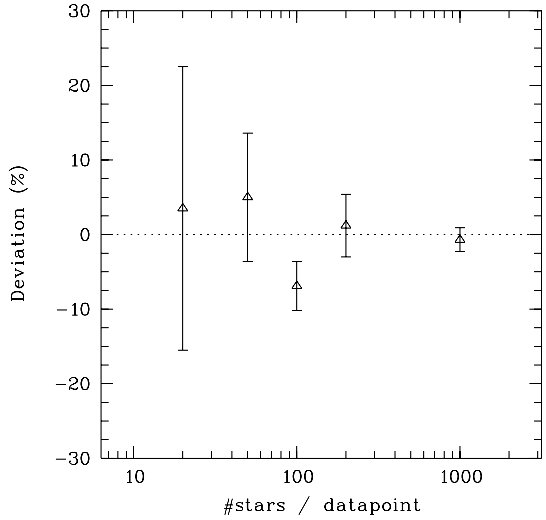

Figure 3. Results of the fitting procedure, for samples belonging to case 1 (an unbiased and volume-limited sample): the relative difference between the original distribution function and the fit (triangles), as a function of $n_\star$, the available number of stars per position. Error bars denote the dispersion on this value. Fits become better with increasing $n_\star$, and form a good approximation for the original distribution function when more than 100 stars per position are available.

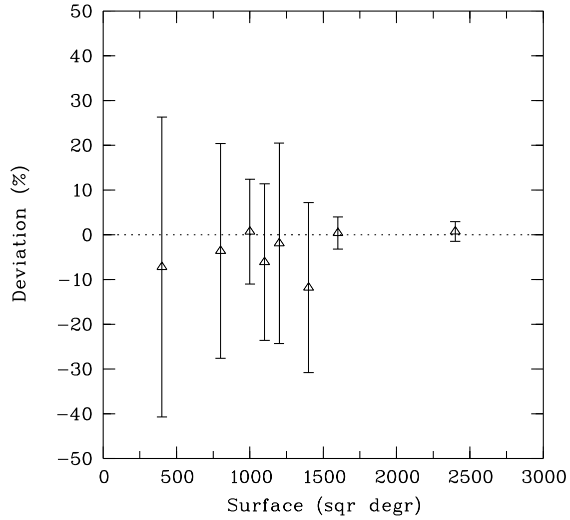

Figure 4. Results of the fitting procedure, for the samples from case 2 (unbiased and volume-limited over an area): the relative difference between the original distribution function and the fit (triangles), this time as a function of area on the sky. Error bars denote the dispersions on the deviations. The original distribution function is recovered if one uses an area of at least 1500 square degrees.

the Sun. Here we investigated the difference between the original distribution function and the fit, when the area was allowed to vary. We worked with a fixed number of stars per position and a sky density comparable to that of Hipparcos (i.e. about 3 stars per square degree), and tried to figure out what the minimum area is in order to get a reasonable fit. For this case we assumed 100 sampled stars per position. The adopted star density and total area then fixes the total number of stars to be considered, which in turn determines the number of positions to be sampled. The generation of the data points then is similar to case 1.

Figure 4 represents the results of the fits for star samples of case 2. The area on the sky in which the sample stars are located varies from 400 up to 2400 square degrees. Obviously, fits become better with larger areas on the sky. The results show that a sample, that is unbiased and volume-limited over an area, reproduces the distribution function sufficiently well, if the area is at least 1500 square degrees.

5. CONCLUSIONS

We investigated to what extent star samples such as the Hipparcos database can be used to construct a distribution function for the orbits of stars in the solar neighbourhood. We generated artificial data samples by using a galactic model, that consists of an axisymmetric Stäckel potential and a three integral distribution function, chosen to represent features observed in the solar neighbourhood. Quadratic Programming routines were used to perform the fitting procedures. For samples that are unbiased and volume limited in a sphere with radius 1 kpc around the sun, we were able to recover the original distribution function, when the number of stars per position is sufficiently high ($n_\star \geq 100$, when working with 40 positions). Unbiased samples that are volume limited toward a certain direction on the sky, within 1 kpc, also enable us to recover the distribution function, when the area on the sky in which the sample stars are distributed is large enough (at least 1500 square degrees, given $n_\star = 100$ and a Hipparcos-like density of 3 stars per square degree).

REFERENCES

Batsleer, P., Dejonghe, H., 1994, A&A, 287, 43

Dejonghe, H., 1989, ApJ, 343, 113

Dejonghe, H., 1992, in Galactic Bulges, IAU Symposium 153, eds. H. Dejonghe & H. Habing (Dordrecht: Reidel), p.73

Durand, S., Dejonghe, H., Acker, A., 1996, A&A, 310, 97

Sevenster, M.N., Dejonghe, H., Habing H.J., 1995, A&A, 299, 689

Van Caelenberg, K., Dejonghe, H. 1997, in preparation

de Zeeuw, T., 1985, MNRAS, 216, 273