THE ASTRONOMICAL JOURNAL, 121:2737–2742, 2001 May
© 2001. The American Astronomical Society. All rights reserved. Printed in U.S.A.

THE SPATIAL DISTRIBUTION OF O–B5 STARS IN THE SOLAR NEIGHBORHOOD AS MEASURED BY *HIPPARCOS*[1]

JESÚS MAÍZ-APELLÁNIZ
Space Telescope Science Institute,[2] 3700 San Martin Drive, Baltimore, MD 21218
Received 2000 November 13; accepted 2001 January 17

ABSTRACT

We have developed a method to calculate the fundamental parameters of the vertical structure of the Galaxy in the solar neighborhood from trigonometric parallaxes alone. The method takes into account Lutz-Kelker–type biases in a self-consistent way and has been applied to a sample of O–B5 stars obtained from the *Hipparcos* catalog. We find that the Sun is located 24.2 ± 1.7 (random) $\pm$ 0.4 (systematic) pc above the Galactic plane and that the disk O–B5 stellar population is distributed with a scale height of 34.2 ± 0.8 (random) $\pm$ 2.5 (systematic) pc and an integrated surface density of $[1.62 \pm 0.04$ (random) $\pm$ 0.14 (systematic)] $\times 10^{-3}$ stars pc^{-2}. A halo component is also detected in the distribution and constitutes at least $\approx$5% of the total O–B5 population. The O–B5 stellar population within $\sim$100 pc of the Sun has an anomalous spatial distribution, with a less than average number density. This local disturbance is probably associated with the expansion of the Gould Belt.

Key words: astrometry — catalogs — Galaxy: structure — solar neighborhood — stars: early-type

1. INTRODUCTION

Studies based on the spatial distribution of massive stars (Blaauw 1960) and interstellar hydrogen (Gum, Kerr, & Westerhout 1960) established some time ago that the Sun is located above the Galactic plane. This has been confirmed using a variety of tracers (see Humphreys & Larsen 1995 for a review) with somewhat incompatible values of the Sun's distance above the Galactic plane, $z_\odot$, ranging from 9.5 pc to 42 pc. Some of the discrepancies may be due to real differences in the vertical distribution of different stellar or gaseous components, but others must be due to observational errors, since they appear among studies of the same type of objects. This is seen in Table 1, where the results on early-type stars are shown. The use of early-type stars to determine $z_\odot$ has the advantage that their scale height, h, with respect to the Galactic plane is of the same order of magnitude as $z_\odot$, which makes the determination simpler than for objects with larger scale heights (e.g., late-type stars). However, early-type stars are scarce, and to obtain a statistically reliable sample previous studies had to resort to objects close to the Galactic plane, where extinction can introduce severe biases.

The measurement of $z_\odot$ using early-type stars is closely related to the vertical distribution of the number density of those stars, $\rho(z)$. We consider here two scenarios. First, if the analyzed sample consists of a single population that is responsible for the gravity field in which the stars are moving we have a single-component, self-gravitating, isothermal disk (e.g., King 1996), which can be described by

$$\rho(z) = \rho_o \,\mathrm{sech}^2\left(\frac{z + z_\odot}{2h_s}\right), \tag{1}$$

where z is the vertical coordinate (with its origin at the Sun's position), h_s is the corresponding scale height, and ρ_o is the density at the Galactic plane. The factor of 2 is introduced to have $\rho \propto \exp(-|z + z_\odot|/h_s)$ for $|z + z_\odot| \gg h_s$. A second possible scenario can be introduced by assuming the same conditions as before but now establishing that the gravity field is caused by a different population (e.g., low-mass stars or dark matter) with constant density. In that case we have a Gaussian vertical distribution

$$\rho(z) = \rho_o \exp\left[-\frac{1}{2}\left(\frac{z + z_\odot}{h_g}\right)^2\right], \tag{2}$$

where h_g is the Gaussian half-width.

The stellar surface density, σ_*, can be obtained by integrating $\rho(z)$ from $-\infty$ to $+\infty$, and for the models described by equations (1) and (2) the results are $4h_s\rho_o$ and $(2\pi)^{1/2}h_g\rho_o$, respectively. Measured values of h and σ_* have been included in Table 1. In each case it is indicated whether the model used is that of equation (1) or an exponential one: $\rho(z) = \rho_o \exp(-|z + z_\odot|/h_e)$, in which case $\sigma_* = 2h_e\rho_o$.

The profiles defined by equations (1) or (2) are actually not extremely different from one another and either one of them is expected to be a reasonably good approximation to the vertical profile of the disk population of early-type stars. However, we must bear in mind that there are also early-type stars in the halo (see, e.g., Rolleston et al. 1999 and references therein) that are expected to dominate the O–B5 population at distances larger than several hundred parsec from the Galactic plane and, thus, add another component to the observed disk population. Most of these objects are runaway stars (Rolleston et al. 1999; Hoogerwerf, de Bruijne, & de Zeeuw 2000), but some appear to have been formed in situ (Conlon et al. 1992).

The availability of the *Hipparcos* catalog (ESA 1997) has revolutionized the quality of astrometric data and has produced for the first time good-quality trigonometric parallaxes for a large sample of early-type stars. However, even with these data one has to take into account a fundamental problem discovered by Lutz & Kelker (1973) in the use of trigonometric parallaxes for the determination of distances. Assuming a uniform spatial distribution of stars and a Gaussian distribution of observed parallaxes π_o about the true parallax π with a width σ_π, they discovered that the distribution of true distances $r = 1/\pi$ about the observed

[1] Based on data from the *Hipparcos* astrometry satellite.
[2] The Space Telescope Science Institute is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

TABLE 1

EARLY-TYPE STELLAR POPULATION PARAMETERS

Source	h (pc)	$z_\odot$ (pc)	Range[a]	Model[b]	$\sigma_{*,\,\mathrm{meas}}$[c] (stars pc^{-2})	Range[d]	$\sigma_{*,\,\mathrm{equiv}}$[e] (stars pc^{-2})
Allen 1973	60	...	O+B	Exponential	1.2×10^{-2}	O+B	2.6×10^{-3}
Stothers & Frogel 1974	46 ± 7[f]	24 ± 3	O–B5	Exponential	5.3×10^{-4}	O–B3.5	0.85×10^{-3}
Stothers & Frogel 1974	70 ± 3[g]	...	...	...	...	...	...
Conti & Vacca 1990	45 ± 5	15 ± 3	WR	Self-gravitating	...	...	...
Reed 1997, 2000	45 ± 20	9.5 ± 3.5	O–B2	Exponential	$(0.6\text{–}8.0) \times 10^{-4}$	O–B2	$(0.15\text{–}1.95) \times 10^{-3}$

Previous determinations of parameters associated with the vertical distribution of the early-type stellar population. The last column was calculated by multiplying the measured surface density by a factor derived from an unbiased O + B sample.

[a] Spectral subtypes used for the scale height and Sun's distance above the Galactic plane determinations.

[b] Model used for the scale height determination.

[c] Measured surface density.

[d] Spectral subtypes used for the surface density determination.

[e] Equivalent surface density for O–B5 stars.

[f] Within 200 pc of the Sun.

[g] Within 800 pc of the Sun.

one $r_o = 1/\pi_o$ was asymmetrical and even divergent for $r \to \infty$. The effect is caused by the inverse relationship between π and r and by the fact that the available volume between r and $r + dr$ is proportional to r^2. The effect is unimportant for parallaxes with small relative errors ($\varepsilon_\pi \equiv \sigma_\pi / \pi_o \lesssim 0.05$) and can be corrected for cases with $\varepsilon_\pi \approx 0.10 - 0.15$, but for values with $\varepsilon_\pi \gtrsim 0.175$ it becomes so large that an individual trigonometric parallax by itself provides little information about the distance to the object unless it is combined with additional information (e.g., the spectroscopic parallax). Lutz & Kelker (1973) provided a table to correct the effect of the bias for $\varepsilon_\pi < 0.175$ (see Smith 1987; Kovalevsky 1998 for updated analyses) but for larger values no simple meaningful correction exists.

A large fraction of the *Hipparcos* trigonometric parallaxes have $\varepsilon_\pi > 0.175$ and a nonnegligible number have negative parallaxes. Those values cannot be discarded when statistical studies of distances are undertaken without introducing severe biases. One approach to the problem is that of Ratnatunga & Casertano (1991), who use spectroscopic parallaxes to constrain the possible range of distances without discarding any stars with large values of ε_π or negative ones of π_o. Another approach is that followed by Feast & Catchpole (1997) to evaluate the zero point of the Cepheid period-luminosity relation. Those authors combine trigonometric parallaxes with ε_π up to ≈ 1.0 with the apparent magnitude and period of the stars to reduce all Cepheids to a single distance and produce an equivalent parallax with $\varepsilon_\pi \approx 0.05$.

In this paper we introduce a new method to correct for Lutz-Kelker–type biases (§ 2) and apply it to a sample of stars obtained from the *Hipparcos* catalog (§ 3) to derive the parameters that determine the vertical distribution of O–B5 stars in the solar neighborhood (§ 4).

2. DESCRIPTION OF THE METHOD

The number of stars per unit distance, $n_{*,r}$, between r and $r + dr$ inside a solid angle Ω, $n_{*,r}\,dr$, is given by

$$n_{*,r}\,dr = \rho(r)\Omega r^2 dr\ , \tag{3}$$

where $\rho(r)$ is the star number density. The equivalent expression per unit parallax is

$$n_{*,\pi}\,d\pi = \rho(1/\pi)\Omega\pi^{-4}d\pi\ . \tag{4}$$

If we assume that the distribution of π_o about π is Gaussian with a width σ_π, then the distribution of observed parallaxes will be the convolution of the two functions

$$f(\pi_o) = \int_0^\infty Ae^{-(1/2)[(\pi-\pi_o)/\sigma_\pi]^2}\rho(1/\pi)\Omega\pi^{-4}d\pi\ , \tag{5}$$

where A is a normalizing factor. Note that equation (4) may be singular at $\pi = 0$ for some choices of $\rho(r)$ and that this may cause equation (5) to be nonnormalizable. This is the reason why no simple correction can be obtained for the standard case of Lutz-Kelker bias (which has $\rho(r) = \rho_o$) when $\varepsilon_\pi \gtrsim 0.175$. A generalization of that standard case can be achieved by allowing for nonconstant density profiles as previously suggested by Hanson (1979). This is motivated by the finite size of the Galaxy, which forces the introduction of a cutoff in any choice of $\rho(r)$ based on a realistic model. Such a cutoff would eliminate any convergence problems in equation (5) and allow for finite Lutz-Kelker–type corrections.[3]

We consider one such model based on a self-gravitating isothermal vertical distribution (eq. [1]) that for early-type stars should be valid for large absolute values of the Galactic latitude b, where the main limitation imposed on sample completeness is distance (as opposed to extinction):

$$\rho(1/\pi) = \rho_o\,\mathrm{sech}^2\left(\frac{\sin b/\pi + z_\odot}{2h_s}\right) = \rho_o\,\mathrm{sech}^2\left(\frac{1/\pi \pm z'_\odot}{2h'_s}\right), \tag{6}$$

where $h'_s = h_s|\csc b|$ and $z'_\odot = z_\odot|\csc b|$, and the sign in the last expression is positive for $b > 0$ and negative for $b < 0$. A least-squares fit to the observed distribution of parallaxes using equations (5) and (6) should then yield the three free parameters h_s, $z_\odot$, and σ_* (which is a function of the constants in the two equations). Unfortunately, two

[3] Note, however, that such corrections will depend on $\rho(r)$ itself and that for a nonspherically symmetric density distribution (as expected for a Galactic disk) $\rho(r)$ will be different for each direction. Thus, a tabulated version of Lutz-Kelker–type corrections becomes clearly impractical because of the increase in the number of degrees of freedom, and the best strategy to estimate the distance is to compute an individual correction for each star.

problems arise when this is attempted. First, a large range in b needs to be used to have a large enough sample, which implies that a mean value of $|\csc b|$ will have to be obtained to derive h_s and $z_\odot$ from h_s' and $z_\odot'$. However, for $|b| \lesssim 25°$ and reasonable values of h_s and $z_\odot$, $|\csc b|$ becomes so large that we are forced to go to distances larger than 1 kpc to have reasonably complete samples, and this is beyond the volume well-sampled by *Hipparcos*. Therefore, this method cannot be applied to the data currently available. The second problem is more subtle: Suppose we analyze all the stars contained within a double cone defined by $|b| > b_{\rm min}$ and that the vertical spatial distribution of early-type stars is given by the sum of either equation (1) or equation (2) and a halo component that is unimportant near the Galactic plane but dominates at high $|z|$. Then, since the sampled volume is proportional to z^2, the halo component will be overrepresented in our sample, and a strong bias could be introduced in our measurements of h_s and $z_\odot$ unless the proper linear combination of $\rho_{\rm disk}$ and $\rho_{\rm halo}$ is used.

An alternative method that avoids these two problems can be designed in the following way. Suppose that we can make an educated guess for the parameters that describe the O–B5 disk population (using either eq. [1] or eq. [2]) and that we also have a reasonably good approximation for the halo population (which is expected to be only a small contaminant in our sample if we avoid the bias mentioned in the previous paragraph). Then, for a single star with measured π_o, σ_π, and b we can calculate the probability distribution as a function of distance as

$$p(r) = Ar^2 e^{-(1/2)[(1-r\pi_o)/r\sigma_\pi]^2}\rho(r)\,, \qquad (7)$$

where A is a normalizing factor and the latitude dependence is embedded in $\rho(r) = \rho_{\rm disk}(r) + \rho_{\rm halo}(r)$. We can now project $\rho(r)$ into the vertical coordinate to obtain $p(z)$ and sum over all observed stars to obtain $f(z)$, the distribution of stars as a function of the vertical coordinate. Doing a least-squares fit of $f(z)$ with a vertical profile leads to a new estimate of the parameters. The process is then repeated until convergence is achieved.

Several precautions have to be taken to ensure the correctness of this iterative process. First, a maximum distance to the Sun vertical line, $\varpi_{\rm max}$, has to be specified to ensure completeness and to avoid the halo-disk bias previously discussed. This can be achieved by truncating each individual $f(z)$ at $z_{\rm max} = \varpi_{\rm max}\tan b$. Second, stars at very low Galactic latitudes ($|b| < b_{\rm min}$) have to be excluded from the sample since for lines of sight close to the Galactic equator the spatial structure is dominated by radial (not vertical) variations and extinction. To prevent inducing any biases due to this exclusion, the fit to the vertical profile should not include the range $(-\varpi_{\rm max}\tan b_{\rm min}, \varpi_{\rm max}\tan b_{\rm min})$. As shown in Figure 1, the resulting volume that will be used for our fit to the vertical profile is a double semi-infinite cylinder. For any given star with $|b| > b_{\rm min}$, three regions are then defined: $0 < |z| < \varpi_{\rm max}\tan b_{\rm min}$, $\varpi_{\rm max}\tan b_{\rm min} \le |z| \le \varpi_{\rm max}\tan|b|$, and $\varpi_{\rm max}\tan|b| < |z| < \infty$, and only the second one is used for our calculation. A star with a large value of ε_π will have nonnegligible values of $f(z)$ in the three regions, while the probability distribution for one with a small value of ε_π will be concentrated in one of the three (unless its value of π_o is close to one of the two critical values that correspond to the two boundaries between regions). In any case, in the second region each of those stars will have an integrated probability between 0 and 1, leading to a nonintegral estimated number of stars in the double semi-infinite cylinder. A third precaution to be considered is the introduction of biases in the obtained parameters by the fitting procedure. This problem can be solved by using a twofold strategy: First, biases can be minimized using a variable bin size determined by the criteria specified by D'Agostino & Stephens (1986). Second, a population with the same parameters as those derived in the fitting process can be generated numerically and fitted using the same procedure. Repeating this a large number of times allows us to measure the numerical biases by obtaining the mean displacements between the real and the derived parameters and to correct our results accordingly.

3. THE SAMPLE

We have extracted from the *Hipparcos* catalog the data for all stars classified therein as O–B5 with $|b| > b_{\rm min} = 5°$ (3531 objects). That value of $b_{\rm min}$ was chosen to eliminate the highly obscured regions near the Galactic plane while at the same time retaining most of the O–B5 stars in the immediate solar neighborhood. Also, for $h \approx 45$ pc (a typical

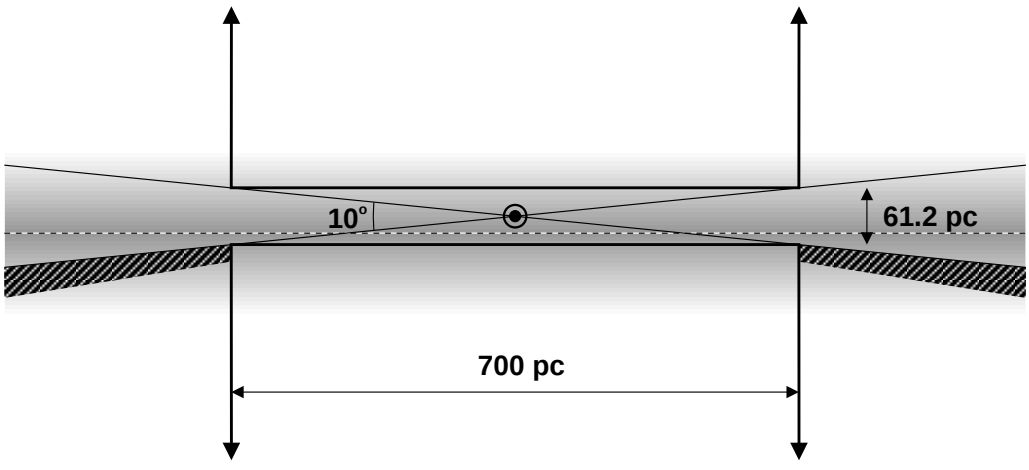


FIG. 1.—Diagram of the double semi-infinite cylinder (*two regions inside the thick lines*) used in this paper to study the vertical distribution of O–B5 stars, showing the Galactic disk (*shaded area*), the Galactic plane (*dashed line*), and our star (*Sun symbol*). The values of the parameters used are $\varpi_{\rm max} = 350$ pc and $b_{\rm min} = 5°$. The hatched regions mark the area excluded to estimate the bias induced by extinction (see the text for details).

TABLE 2

FREQUENCY OF *HIPPARCOS* RELATIVE PARALLAX ERRORS

ε_π Range	Number
0.00–0.05	6
0.05–0.15	237
0.15–0.25	380
0.25–0.50	848
> 0.50	1 444
Negative parallax.....	467

Frequency of *Hipparcos* relative parallax errors for the selected subsample of 3382 O–B5 stars used in this work.

value chosen from Table 1), a vertical distance of three scale heights corresponds to $\approx$1.5 kpc in ϖ for $|b| = 5°$, so we would not expect to sample regions with large variations in the early-type star population due to radial Galactic gradients. From that sample we have selected a subsample (3382 objects) with good astrometric solutions by eliminating those stars with more than 5% of data points rejected (ESA 1997, col. 29) and with a goodness-of-fit statistic g for ν degrees of freedom

$$g = \left(\frac{9\nu}{2}\right)^{1/2}\left[\left(\frac{\chi^2}{\nu}\right)^{1/3} + \frac{2}{9\nu} - 1\right], \qquad (8)$$

less than 3.0 (col. 30).[4]

Since we expect most stars in our subsample to be located at distances greater than 100 pc and since for *Hipparcos* data typically $\sigma_\pi \approx 1$ mas, Lutz-Kelker–type biases should be important. Indeed, that is what can be deduced from Table 2, where the distribution of ε_π values is shown.

The other parameter that has to be selected to define the double semi-infinite cylinder that will be used in this work is $\varpi_{\max}$. To do that, we have to consider that the *Hipparcos* completeness limit for early-type stars is determined by $m_V = 7.9 + 1.1|\sin b|$, which corresponds to $m_V = 8.0$ for $|b| = 5°$. A very small fraction ($\ll$1%) of stars below the limit is actually not included and the catalog is not complete between that limit and $m_V \approx$ 12–13 (ESA 1997). A B5 V star has $M_V \approx -1.2$; at a distance of 350 pc such a star with $A_V = 1.0$ mag (obtained from a very conservative estimate of 3.0 mag kpc^{-1} in V) would have $m_V = 7.5$, still within the completeness limit. If such a star was located at 135 pc (approximately 3 scale heights) from the Galactic plane, with $\varpi = 350$ pc, its apparent magnitude would also be $\approx$7.5 if we estimate A_V as $\approx$0.5 (again, a very conservative estimate at that latitude). We can then conclude that a choice of $\varpi_{\max} = 350$ pc leads to a nearly complete sample for the O–B5 disk stellar population.

4. RESULTS

We have carried out least-squares fits to the data described in § 3 by using either a self-gravitating isothermal or a Gaussian disk model and applying the method described in § 2. Since the halo population in our sample is clearly incomplete due to the *Hipparcos* magnitude limits, to describe it we used a simple reasonable model, a parabola. That choice allows us to discriminate between the disk and halo contributions within a few hundred parsecs of the Sun, but it can produce only a lower limit for $\sigma_{*,\mathrm{halo}}$. The data and fits are displayed in Figures 2 and 3 and the obtained parameters are shown in Table 3.

As previously discussed, a variable bin size was used to minimize numerical biases. For each of the two models we

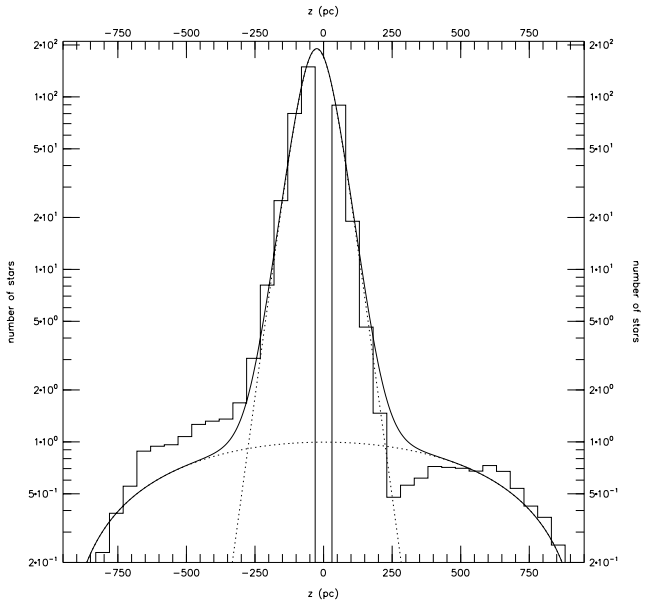


FIG. 2.—Observed (histogram) and model-fit frequency distributions as a function of the vertical coordinate for O–B5 stars, assuming a self-gravitating isothermal disk and parabolic halo distribution. The dotted lines represent individual components and the continuous one represents the sum of the two. The displayed bin size is uniform and equal to 50 pc but is not used for the fit (see text). The region immediately surrounding $z = 0$ is not considered for the fit, as it corresponds to the space between the two semi-infinite cylinders. Note that the observed distribution shows "fractional stars" due to the procedure used to derive the histogram.

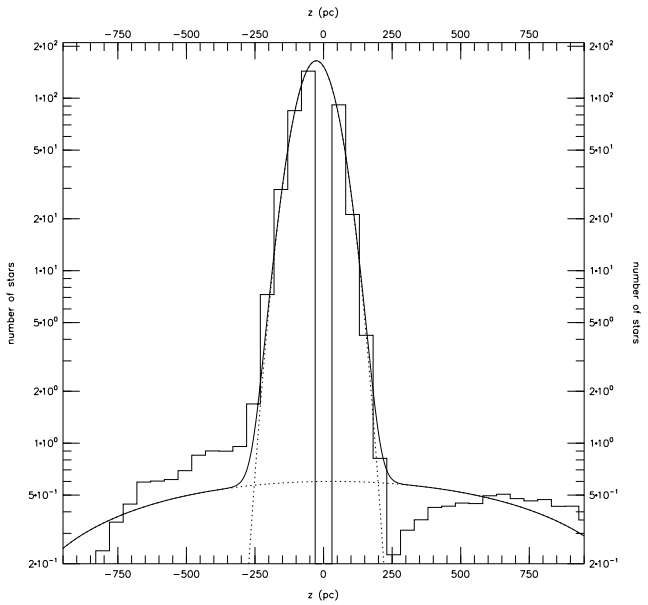


FIG. 3.—Same as Fig. 2, but using a Gaussian disk model instead of a self-gravitating isothermal one.

[4] The selected subsample is available at http://www.stsci.edu/~maiz/data/Hipparcos.

TABLE 3

RESULTS OF THE LEAST-SQUARES FITS

Model	h (pc)	$z_\odot$ (pc)	$\sigma_{*,\,\mathrm{disk}}$ (10^{-3} O–B5 stars pc^{-2})	$\sigma_{*,\,\mathrm{halo}}/\sigma_{*,\,\mathrm{disk}}$	χ^2/ν	ν	g
1.....	$34.2 \pm 0.8 \pm 2.5$	$24.2 \pm 1.7 \pm 0.4$	$1.62 \pm 0.04 \pm 0.14$	0.058	0.88	41	−0.50
2.....	$62.8 \pm 1.7 \pm 4.7$	$25.2 \pm 1.5 \pm 0.5$	$1.67 \pm 0.04 \pm 0.24$	0.037	0.87	41	−0.56

The least-squares fits used either a self-gravitating isothermal (first row) or a Gaussian (second row) disk with parabolic halo distribution. The values have been bias-corrected, and both the random (first) and systematic (second) uncertainties are shown in each case. The second column refers to the scale height (h_s) in the first row and to the half-width (h_g) in the second row. No uncertainties are given for $\sigma_{*,\,\mathrm{halo}}/\sigma_{*,\,\mathrm{disk}}$ because the fitted values are only lower limits to the real distributions.

generated 50 populations with the same parameters derived from the real population, and we estimated the biases by calculating the difference between the mean values of the derived parameters and the real ones. As expected, those biases were very small, typically 0.5 pc for h and $z_\odot$. However, one possible bias due to extinction still remains. Suppose we are observing a star with $b = -5°$ located inside the double semi-infinite cylinder but with a high value of ε_π. This star will be counted only as a "fractional star" (i.e., its integrated probability inside the volume of interest will not be close to one). If the sample were complete both inside and outside the double semi-infinite cylinder this would not be a problem, since the lost probability would be compensated for by that contributed by stars located outside the cylinder. However, we expect that extinction (and, to a lesser degree, distance) will remove from our sample some of the stars outside the cylinder, especially those closer to the Galactic plane. To simulate this effect we consider an extreme case by placing a wall of infinite extinction at $\varpi = \varpi_{\max}$ that extends 1 scale height (for the self-gravitating isothermal case) or 1 half-width (for the Gaussian case) north and south of the Galactic plane. The volume excluded by this wall is represented by the hatched area in Figure 1 for the self-gravitating isothermal case. Note that only stars in the southern Galactic hemisphere are eliminated by this wall since the position of the Sun above the Galactic plane locates any star with $b = 5°$ above the wall.[5] Under these conditions, we eliminated the stars in the hatched volume from our generated 50 populations and calculated the corresponding biases by calculating new least-squares fits. As expected, the new biases turned out to be more relevant than the ones obtained before, but even then they never induced changes larger than 15% in the values of the parameters. Since the real situation must be between the two extreme cases discussed here ([1] all stars in the possibly excluded area partially contributing to the probability inside the cylinder and [2] none of them doing so), we considered the real bias to be the average of the two and we indicated the two extremes as a systematic error.

The results in Table 3 indicate that both the self-gravitating isothermal and the Gaussian models are very good descriptions of the disk population, since g is close to 0 in both cases and the fitted distribution near the Galactic plane closely resembles the observed one, as shown in Figures 2 and 3. The parabolic fit to the halo population is of worse quality, as expected from the much smaller number of stars in our sample belonging to it and from its incompleteness. Nevertheless, the existence of a halo population in our sample is beyond doubt, as is clearly shown in Figures 2 and 3. It is also well known that this halo population extends farther than the distance-induced limits in our sample, even beyond a few kiloparsec (Rolleston et al. 1999). Thus, the values of $\sigma_{*,\mathrm{halo}}/\sigma_{*,\mathrm{disk}}$ that appear there must be understood only as lower limits to the real ones.

The values of $z_\odot$ and $\sigma_{*,\mathrm{disk}}$ obtained using the two different models are clearly compatible. Since we are unable to determine which of the two models provides a better fit to the disk population, either result can be used. Probably the ones obtained with the self-gravitating isothermal profile are preferred because of the better quality of the parabolic fit to the halo population, which should introduce less numerical noise into the result than in the Gaussian case.

Our value for h_s is somewhat lower than the ones obtained by other authors. The discrepancies are at only the 2 σ level with the values obtained by Conti & Vacca (1990) and by Stothers & Frogel (1974) (within 200 pc of the Sun) and even lower in comparison with the Reed (2000) result. Allen (1973) does not quote an uncertainty, while the value of Stothers & Frogel (1974) within 800 pc of the Sun is quite probably affected by heavy extinction. Some of the discrepancies may be attributed to contamination by halo stars (a nonnegligible effect, as we have seen) and others to the use of a purely exponential model instead of a self-gravitating isothermal one. Our value of $z_\odot$ is clearly compatible with the one obtained by Stothers & Frogel (1974) but is higher than the ones obtained by Conti & Vacca (1990) and Reed (1997), in this last case by a large amount. We believe that the differences may be caused by the difficulty in accounting for complex extinction variations at low Galactic latitudes, especially when one has to rely on spectroscopic parallaxes, a problem which does not affect our results since they are based on unbiased trigonometric distances. This can be seen in the fact that the result more similar to ours is the one that is obtained from the nearest population. Finally, our value for σ_* is within the range obtained by Reed (2000) but is clearly higher than the one obtained by Stothers & Frogel (1974) and lower than the one obtained by Allen (1973). Once again, we believe that extinction may be responsible for some of the discrepancies, but we must also take into account that σ_* is expected to suffer significant variations depending on the maximum radius used, since different regions of the disk (close to or far from spiral arms) are expected to be sampled.

One last point to discuss is the existence of spatial variations within the region discussed in this paper. Stothers & Frogel (1974) discovered that a 100 pc–radius hole exists in the distribution of O–B5 stars at low Galactic latitudes. Also, it has been known for a long time that a large fraction of the early-type stars in the solar neighborhood belong to

[5] For the Gaussian case a small volume is also eliminated in the northern Galactic hemisphere.

the Gould Belt, a structure inclined 15°–20° with respect to the Galactic plane that is expanding with respect to a point located close to the Sun's position (see Torra, Fernández, & Figueras 2000 for a recent analysis). Is this expected minimum in the number density of early-type stars detected in the *Hipparcos* data? It indeed is. The parameters derived from our self-gravitating isothermal model predict that there should be 11.6 O–B5 stars within 66.67 pc of the Sun (i.e., with $\pi \geq 15$ mas). However, only three[6] O–B5 stars are found in that volume: η UMa, α Eri, and α Pav. The difference between the expected and the measured number appears to be too large to be explained as a statistical fluctuation, so we conclude that the existence of a local "hole" in the distribution of early-type stars is indeed real.

We conclude that our method provides a solid measurement of the vertical distribution of O–B5 stars in the solar neighborhood by using trigonometric parallaxes alone and by taking into account all relevant biases.

The author would like to thank Ivan King for introducing him some years ago to the problems caused by the Lutz-Kelker bias and Stefano Casertano for very fruitful discussions regarding the topics in this article. Useful comments for improving this paper were also provided by an anonymous referee and by Nolan R. Walborn. Support for this work was provided by NASA through grant GO-8163.01-97A from the Space Telescope Science Institute, Inc., under NASA contract NAS 5-26555, and by the STScI DDRF.

[6] The three have small values of ε_π so their position can be determined with certainty to be within that volume. Two other stars (HIP 23342 and HIP 74368) have $\pi_o \geq 15$ mas but with large values of ε_π and with spectroscopic parallaxes that clearly indicate that they are much farther away. Finally, a few stars have values of π_o slightly below 15 mas and could actually be closer to us than 66.67 pc. However, the probabilistic estimation yields a number lower than 1, and even if there were an additional star within the given distance our arguments would remain unchanged.

REFERENCES

Allen, C. W. 1973, Astrophysical Quantities (3d. ed.; London: Athlone Press)
Blaauw, A. 1960, MNRAS, 121, 164
Conlon, E. S., Dufton, P. L., Keenan, F. P., McCausland, R. J. H., & Holmgren, D. 1992, ApJ, 400, 273
Conti, P. S., & Vacca, W. D. 1990, AJ, 100, 431
D'Agostino, R. B., & Stephens, M. A. 1986, Goodness-of-Fit Techniques (New York: M. Dekker)
ESA. 1997, The *Hipparcos* and Tycho Catalogues (ESA SP-1200) (Noordwijk: ESA)
Feast, M. W., & Catchpole, R. M. 1997, MNRAS, 286, L1
Gum, C. S., Kerr, F. J., & Westerhout, G. 1960, MNRAS, 121, 132
Hanson, R. B. 1979, MNRAS, 186, 875
Hoogerwerf, R., de Bruijne, J. H. J., & de Zeeuw, P. T. 2000, ApJ, 544, L133
Humphreys, R. M., & Larsen, J. A. 1995, AJ, 110, 2183
King, I. R. 1996, Stellar Dynamics (Mill Valley: University Science Books)
Kovalevsky, J. 1998, A&A, 340, L35
Lutz, T. E., & Kelker, D. H. 1973, PASP, 85, 573
Ratnatunga, K. U., & Casertano, S. 1991, AJ, 101, 1075
Reed, B. C. 1997, PASP, 109, 1145
———. 2000, AJ, 120, 314
Rolleston, W. R. J., Hambly, N. C., Keenan, F. P., Dufton, P. L., & Saffer, R. A. 1999, A&A, 347, 69
Smith, H., Jr. 1987, A&A, 188, 233
Stothers, R., & Frogel, J. A. 1974, AJ, 79, 456
Torra, J., Fernández, D., & Figueras, F. 2000, A&A, 359, 82