

# The absolute magnitude of K0V stars from Hipparcos data using an analytical treatment of the Malmquist bias

A. G. Butkevich<sup>1</sup>, A. V. Berdyugin<sup>2</sup>, and P. Teerikorpi<sup>2</sup>

<sup>1</sup> Pulkovo Observatory, Pulkovskoye shosse 65, Saint-Petersburg 196140, Russia  
e-mail: butkevich@regal.spb.ru

<sup>2</sup> Tuorla Observatory, 21500 Piikkiö, Finland  
e-mail: andber@utu.fi; pekkatee@utu.fi

Received 1 November 2004 / Accepted 28 January 2005

**Abstract.** We calculate the average absolute magnitude for Hipparcos single K0V stars, using a theoretical curve for the distance-dependent Malmquist bias in the data. This method is shown to be well applicable to stellar data with good parallaxes and gives results in agreement with a previous study that used another treatment of the bias developed in extragalactic astronomy (finding the “unbiased plateau”). In particular, we point out that such a fit, which uses stars in the biased part of the sample, may be less vulnerable to the fluctuations in the unbiased plateau whose definition is also somewhat subjective. We found for K0V stars  $M_0 = 5.8$ , with the spread of the luminosity function being 0.3 mag. It is shown that inclusion of Hipparcos non-single stars may underestimate  $M_0$  by about 0.05–0.1 mag. During the study it was found that about 20% of the sample may be mis-classified K0IV stars.

**Key words.** methods: statistical – stars: distances – stars: fundamental parameters

## 1. Introduction

In recent years several studies have been made of the luminosity calibration of various stellar types using Hipparcos data (Reid 1999; Lebreton 2000). A number of statistical effects complicate this task (Arenou & Luri 2002; Feast 2002; Smith 2003). Our work focuses on one of them – the Malmquist bias.

Oudmaijer et al. (1998) pointed out the rough correlation between the derived absolute magnitude and parallax for about 2000 Hipparcos stars with high precision parallaxes. Later Oudmaijer et al. (1999, hereafter referred to as OGS) made a more extensive analysis of this effect and identified it as “the magnitude selection Malmquist bias”.

The Malmquist bias originates from a simple but subtle selection effect. It can be briefly expressed as follows: the more distant objects we gather, the brighter objects we get, and at larger distances there is more spatial volume. In a magnitude-limited star sample the bias at a fixed distance cuts off the stars from the faint side of the luminosity function. This effect influences magnitude-limited samples of any observed objects: stars, galaxies and even quasars, even though at high redshifts cosmological effects will interfere.

This type of bias has been known for a long time in extragalactic astronomy (for reviews, see Sandage 1995; Teerikorpi 1997). We shall refer to it as the distance-dependent Malmquist bias, although there is no generally adopted terminology in this field (see examples in Teerikorpi 1997). Our aim is to

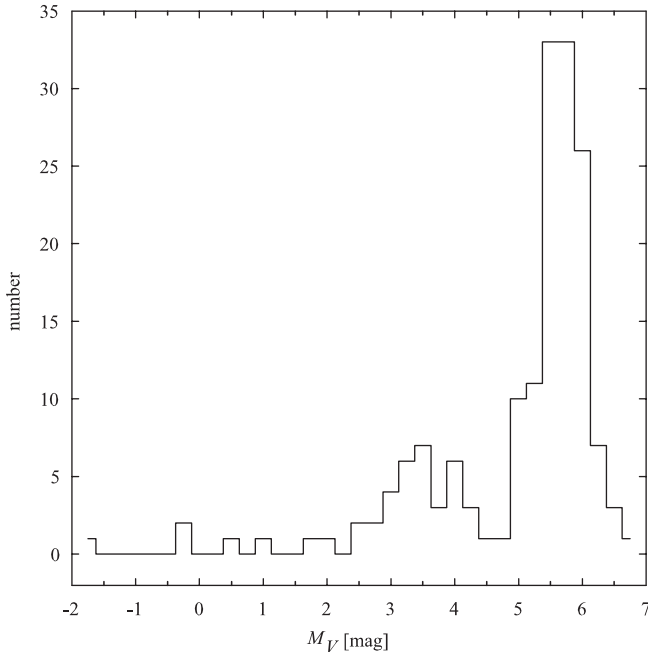
demonstrate that the treatment developed for galaxies may be fruitful for the problem of stellar luminosity calibration. We also have in mind future astrometric space projects which will provide a great number of accurate parallaxes. This paper is the first one in a series where we plan to study basic issues concerning Malmquist-like biases in stellar astronomy.

The paper is organized as follows. In Sect. 2 we describe the criteria used for selecting stars from the Hipparcos catalogue. In Sect. 3 we discuss the statistical properties of the sample. The treatment of the Malmquist bias is given in Sect. 4. We compare our results with those obtained by OGS in Sect. 5. Our conclusions are summarized in Sect. 6.

## 2. Selection of the sample

In order to make our results comparable with those of OGS, we consider the same type of objects as they did, namely K0V stars. The following criteria were used for the selection from the Hipparcos catalogue (ESA 1997):

1. Spectral type (Field H76) is strictly K0V. Stars with ambiguous types, such as G8/K0V or K0IV/V, were excluded.
2. Source of spectral type (Field H77) is Michigan Spectral Survey (hereafter MSS) only, for homogeneity of spectral type data. This restricts the sample to  $\delta \leq 12^\circ$ , because the Hipparcos catalogue only refers to MSS vols. 1–4.
3. Stars with non-positive parallaxes were excluded. No other criteria based on parallax or parallax error were specified.



**Fig. 1.** The distribution of  $M_V$  for 166 Hipparcos single K0V stars calculated from parallax and  $V$  magnitude neglecting the interstellar extinction, binned to 0.25 mag intervals.

- Field H59 is blank. This means that double stars, i.e. all the catalogue entries included in any part of the Double and Multiple Star Annex, were rejected.

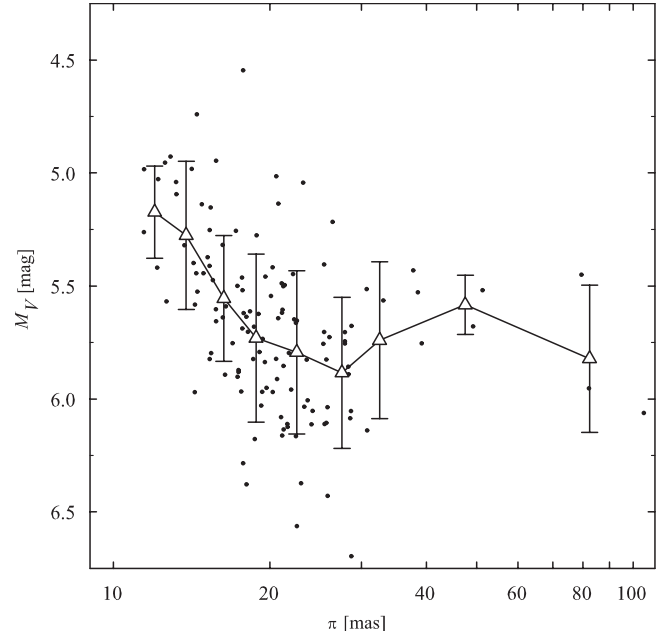
The above criteria slightly differ from those used by OGS. The first three items are essentially common. As to the last criterion, OGS did not make any selection by multiplicity. We did because magnitudes of these stars may be significantly altered by the presence of a companion. This effect is discussed in Sect. 5.

### 3. Properties of the sample

The sample selected by the above criteria contains 166 objects. Not all parallaxes in the sample are of high quality. The worst case HIP 114896 (HD 219493) has a large relative error  $\sigma_\pi/\pi = 2.94$ . Such objects, of course, should be avoided when deriving absolute magnitudes. To study the entire sample of Hipparcos K0V single stars, however, no object was excluded despite the inaccurate absolute magnitudes derived for stars with large  $\sigma_\pi/\pi$ . The mean value of the parallax is 19 mas, corresponding to a distance of 52 pc. Thus the objects are rather nearby, on average, and we ignore interstellar extinction in all our calculations. The distribution of  $M_V$  is plotted in Fig. 1.

The two peaks visible in the diagram suggest that the sample is populated by stars of at least two different classes. The faint one approximately corresponds to K0V stars, while the bright class, close to 3.5 mag, may be roughly identified as K0IV stars. This appearance of two stellar classes seems to reflect a misclassification in MSS, as was already pointed out by OGS. The separation between the classes is at about  $M_V = 4.5$ .

For further study, all stars brighter than  $M_V = 4.5$  were excluded. This rejects 40 of 166 objects, leaving 126 stars. The parallax related properties of this new sample are listed



**Fig. 2.** The dependence of  $M_V$  on parallax for 126 Hipparcos single K0V stars with  $M_V > 4.5$  mag. The mean and its scatter in parallax bins is shown with the triangles. See text for details.

**Table 1.** Parallax properties of 126 single K0V stars with  $M_V > 4.5$ .

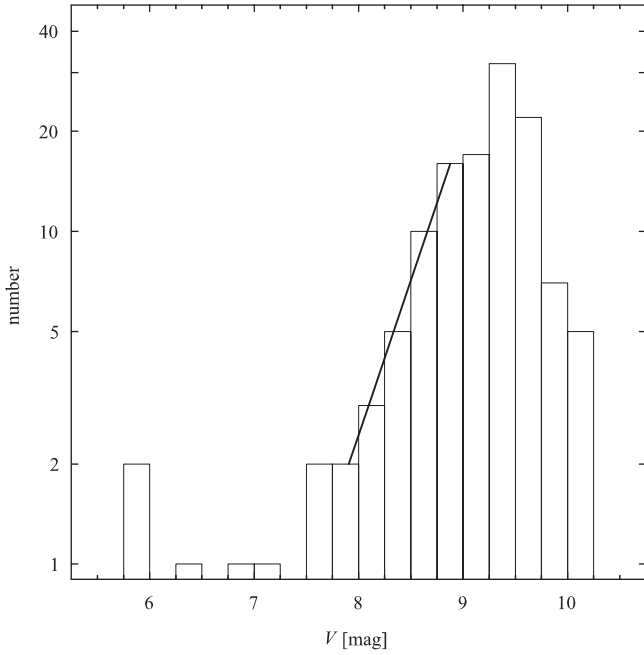
Parameter	Minimum	Maximum	Mean
$\pi$ [mas]	11.44	104.84	$22 \pm 12$
$\sigma_\pi$ [mas]	0.66	1.94	$1.3 \pm 0.3$
$\sigma_\pi/\pi$	0.01	0.15	$0.07 \pm 0.03$

in Table 1. It shows that the quality of parallaxes became much better. Now the bulk of the sample has  $\sigma_\pi/\pi \leq 0.1$  and the absolute magnitudes computed from observed parallaxes are almost unbiased for such small relative errors (Brown et al. 1997).

In order to study the dependence of the derived absolute magnitudes on the parallaxes for the sample of 126 stars with  $M_V > 4.5$  mag, we divide the range of parallaxes into relatively small bins and then calculate the mean  $M_V$  and its scatter in each bin. Ideally, all the bins should contain roughly equal numbers of objects in order to get mean values that are evenly representative in a statistical sense. However, we chose to use the same nine bins as OGS did, for easy comparison: the first  $<12.5$ , the last  $>60$  and the other seven separated at 15, 17.5, 20, 25, 30 and 35 mas. The numbers of stars in the bins are: 4, 15, 20, 23, 34, 19, 3, 5 and 3, in order of increasing  $\pi$ .

The derived  $M_V$  versus  $\pi$  plot is shown in Fig. 2. The most pronounced feature is the clearly visible correlation between absolute magnitude and parallax. The more distant stars appear to be more luminous. This is the familiar signature of the distance-dependent Malmquist bias (Sandage 1994; Teerikorpi 1997).

As tallying of objects into the bins implies selection according to parallax, the derived means may suffer from the parallax error-dependent Lutz-Kelker bias. To estimate this effect quantitatively, we calculate the mean value of the relative parallax



**Fig. 3.** Magnitude distribution of the 126 Hipparcos single KOV stars with  $M_V > 4.5$ . The scale of the vertical axis is logarithmic and the labels indicate the beginning of a bin 0.25 mag wide. The linear envelope of the bright side indicates that the sample is complete up to  $V = 9$ .

error,  $\langle \sigma_\pi / \pi \rangle$ , for each bin. We found comparatively high values, 0.13 and 0.11, only for the first two bins, with the rest of the bins all having  $\langle \sigma_\pi / \pi \rangle$  less than 0.09. Corresponding values of the bias (Lutz & Kelker 1973) are 0.20, 0.13 and less than 0.08 mag, respectively. These quantities are significantly lower than the rise in  $M_V$  due to the magnitude limit. A correction for the bias would change the pattern visible in Fig. 2 only by a small amount. We, therefore, decided to neglect the Lutz-Kelker bias in what follows. Some comments on this subject are, however, given in Sect. 6.

#### 4. Distance-dependent Malmquist bias

A quantitative treatment of the distance-dependent Malmquist bias developed by Teerikorpi (1975) implies that, under the following assumptions:

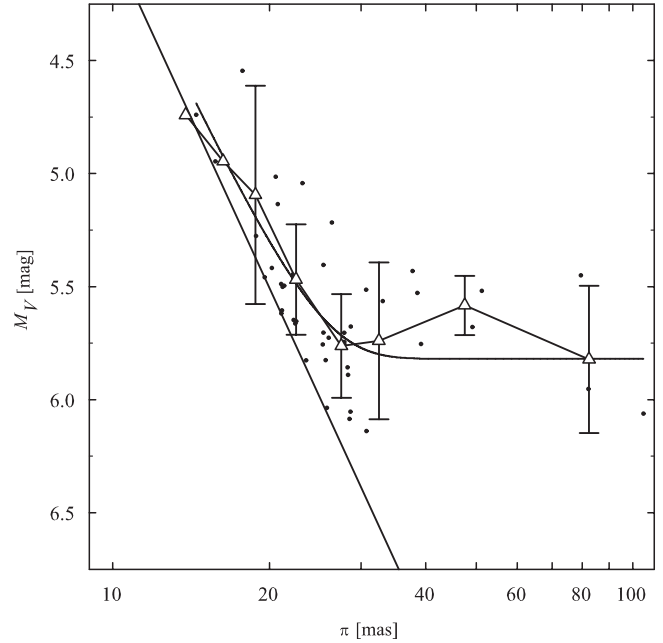
1. the luminosity function obeys a Gaussian law with a mean  $M_0$  and a scatter  $\sigma$ ;
2. the considered stellar sample is complete up to a definite apparent magnitude limit  $m_{\text{lim}}$ ;
3. parallax  $\pi$  gives the true distance as  $r = 1/\pi$ ;

the mean absolute magnitude at a fixed parallax  $\pi$  is given by

$$\bar{M}(\pi) = M_0 - \sigma \sqrt{\frac{2}{\pi}} \frac{\exp\left[-(M_{\text{lim}}(\pi) - M_0)^2 / (2\sigma^2)\right]}{1 + \text{erf}\left[(M_{\text{lim}}(\pi) - M_0) / (\sigma\sqrt{2})\right]}, \quad (1)$$

where the limiting absolute magnitude corresponding to a specific parallax is, with  $\pi$  in mas,

$$M_{\text{lim}}(\pi) = m_{\text{lim}} - 10 + 5 \lg \pi, \quad (2)$$



**Fig. 4.** The resulting sample of 45 KOV stars with  $M_V > 4.5$  and  $V < 9$ . The triangles indicate the mean  $M_V$  and its scatter in the same parallax bins as in Fig. 2. The solid line is the theoretical bias curve (1) fitted to the data. The magnitude limit at  $V = 9$  is shown by the straight line.

and erf is the error function defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \quad (3)$$

Our sample of KOV stars obviously meets none of the above requirements. First, the distribution of absolute magnitude, a practical realization of a real luminosity function, cannot be Gaussian because of the cut-off of stars at  $M_V = 4.5$ . This is not a serious disadvantage in our case. For the Gaussian luminosity function with  $M_0 \approx 5.7$ , and  $\sigma = 0.5$  (an overestimate of the true value), cutting off the bright wing of the distribution at 4.5 mag removes only 0.8% of the stars. Such a small portion cannot significantly affect the expected value of  $\bar{M}$ .

Another critical point is the completeness. Our sample was not constrained to be complete to any magnitude by selection. On the other hand, the above treatment of the Malmquist bias presupposes that the sample is complete up to a sharp magnitude limit. To approach this situation, one may make an artificial cut-off in the magnitude distribution, as has been done in extragalactic applications (e.g. Bottinelli et al. 1987). The distribution of the  $V$  magnitude is shown in Fig. 3. The diagram suggests that the sample is roughly complete up to about 9 mag. All the stars fainter than  $V = 9$  mag were excluded, leaving only 45 objects that are believed to constitute a magnitude-limited sample.

The third condition  $r = 1/\pi$  is, of course, only approximately valid, but the errors are small in the present case.

One can find the mean absolute magnitude,  $M_0$ , and its scatter,  $\sigma$ , for a stellar class in two different ways. The first is to select a region not affected by the bias (an “unbiased plateau” as it is often called in extragalactic astronomy). This region represents the volume-limited part of a sample (Sandage 1994),

**Table 2.** The mean absolute magnitude,  $M_{RP}$  and  $M_{UM}$ , and scatter  $\sigma$  as derived from the unbiased plateau for two samples of Hipparcos K0V stars.  $N$  is the number of objects in the unbiased bias region, for three positions of its boundary.

Boundary (mas)	All stars				Single stars only			
	$M_{UM}$	$\sigma$	$M_{RP}$	$N$	$M_{UM}$	$\sigma$	$M_{RP}$	$N$
20	5.68	0.40	5.69	90	5.80	0.33	5.74	63
25	5.71	0.35	5.72	44	5.81	0.32	5.76	30
30	5.63	0.27	5.73	17	5.69	0.25	5.75	11

which gives us an estimate of  $M_0$  unaffected by the Malmquist bias. It is the method used by OGS. We consider this approach in Sect. 5.

The second way, which we study here, is to fit the theoretical bias curve (1) to the data. A drawback is that the estimates are model-dependent. For our sample of 45 K0V stars, the least-squares fit yields  $M_0 = 5.81 \pm 0.01$  and  $\sigma = 0.33 \pm 0.02$ , with a correlation between them of 0.3 (see Fig. 4).

### 5. Model fit versus unbiased plateau

As we mentioned in Sect. 2, OGS used a different set of criteria for their sample based on the solution quality inferred from the Hipparcos data reduction. They accepted the K0V stars satisfying the following criteria: (1) percentage of rejected data (F1, Field H29) <10%; and (2) goodness-of-fit (F2, Field H30) <3. Moreover, they also rejected the stars brighter than  $M_V = 4.5$ , as we did. In order to compare with their study, we selected the objects meeting the above requirements and obtained a sample of 168, in contrast with 159 found by OSG. This small difference is unexplained but not very significant.

We excluded from this sample all non-single stars, giving 123 objects. The analogous sample considered above (see Sect. 3) contains 126 stars. The three extra stars are HIP 28452, 93579 and 108598 (HD 41327, 176520 and 208740). These objects do not satisfy the F2-criterion, having an F2 value from 3.10 to 3.45. They are not very close, with a parallax in the range from 15.95 to 22.95 mas, therefore will not affect the subsequent analysis.

Thus we have two samples of Hipparcos K0V stars. The first one of 168 entries contains both single and double/multiple stars. The second sample has 123 stars and consists of single stars only. Below we compare the mean absolute magnitudes derived from these samples.

OGS used two techniques to derive the mean absolute magnitude. They calculated an ordinary unweighted mean and so-called reduced parallax. To make a detailed comparison, we also apply both these methods and denote the values obtained from unweighted means and from reduced parallaxes with the subscripts “UM” and “RP”, respectively. Although the method of reduced parallaxes has been widely discussed in the recent literature especially in connection with Hipparcos parallax data (Arenou & Luri 1999, 2002; Feast 2002; Sandage & Saha 2002; Smith 2003), no generally adopted designation for the reduced parallax exists. Following Arenou & Luri (1999), who developed this approach by introducing the promising

concept of astrometric magnitudes, we denote this quantity as  $a$ :

$$a = 10^{M/5}. \tag{4}$$

OGS, unfortunately, did not mention the weighting scheme they used. We adopt the straightforward variant with the weights depending on parallax errors only (Sandage & Saha 2002), with the mean reduced parallax given by

$$\bar{a} = \frac{\sum_i \pi_i \alpha_i / w_i}{\sum_i 1 / w_i}, \tag{5}$$

where  $\alpha_i = 10^{(m_i - 10)/5}$  and the weight,  $w_i$ , is

$$w_i = \alpha_i^2 \sigma_{\pi i}^2. \tag{6}$$

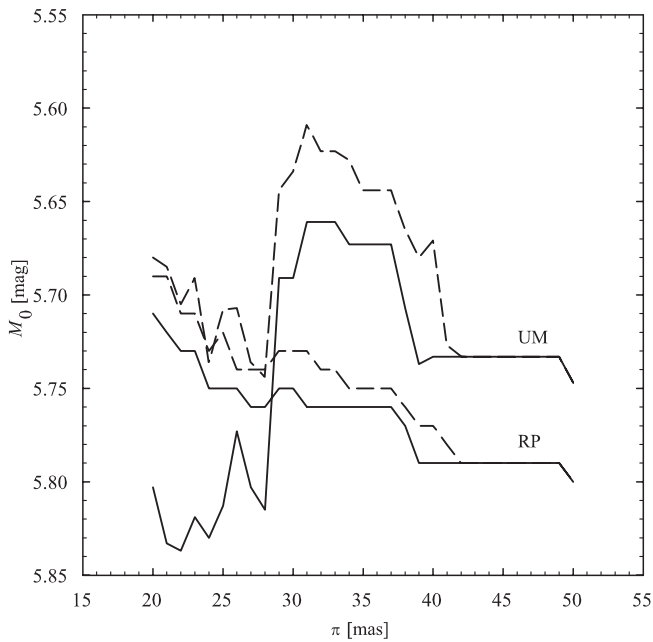
A disadvantage of the method of reduced parallaxes is the conversion from  $a$  to  $M$ . Let us introduce the quantity

$$M_{RP} = 5 \lg \bar{a}. \tag{7}$$

Two means,  $M_0$  and  $M_{RP}$  are in general not equivalent, with the difference,  $\Delta M_{RP} = M_{RP} - M_0$ , depending on a number of factors. For example, a Gaussian luminosity function and a volume-limited sample give  $\Delta M_{RP} = 0.23\sigma^2$ . Thus dispersion  $\sigma$  must be known in order to derive  $M_0$  by this method. We estimate  $\sigma$  with the unweighted means, the obtained values result in  $\Delta M_{RP}$  not exceeding 0.04 mag. We do not apply such a small correction and list  $M_{RP}$  “as it is”.

OGS derived  $M_0$  from an “unbiased plateau”. From an analysis of parallax distribution they concluded that the sample is complete to  $\pi \approx 20$  mas. We think it is hard to judge definitely where the region unaffected by the bias starts (see Fig. 2), therefore we tried three different trial boundaries of the plateau: 20, 25, and 30 mas. Two mean absolute magnitudes,  $M_{RP}$  and  $M_{UM}$ , derived for the stars with parallax higher than these limits, and other related quantities, are given in Table 2. There is a good agreement with the results of OGS, who found  $M_{RP} = 5.69$  and  $M_{UM} = 5.69 \pm 0.40$  for 85 stars with  $\pi > 20$  mas. In our opinion, this form of writing is not quite correct, because 0.40 mag is not an error in  $M_{UM}$ , rather its rms scatter. Although OGS did not give the error, it may be estimated as the scatter divided by  $\sqrt{N}$ . For  $N = 85$  we have 0.04 mag, thus the value they found can be written as  $M_{UM} = 5.69 \pm 0.04$ .

Another interesting thing in Table 2 is the variation in the mean values with the parallax limit. This prompted an additional test. A set of the unbiased plateau boundaries were probed in the range 20 to 50 mas, separated by 1 mas.



**Fig. 5.** Comparison of the mean absolute magnitudes,  $M_{RP}$  and  $M_{UM}$ , calculated from the unbiased plateau for the two samples of Hipparcos K0V stars. The means derived for single stars only are shown by the continuous lines. The dashed lines give the magnitudes computed for all stars.

Figure 5 shows much fluctuation, e.g. a significant decrease in  $M_{UM}$  for  $\pi > 30$  mas, due to the random occurrence of a few bright stars. Although the means roughly behave in the expected manner, the fluctuations may lead to erroneous conclusions. Clearly, when the numbers are small, the results obtained from the unbiased plateau are to be treated with caution. The clearly visible discrepancy between  $M_{UM}$  and  $M_{RP}$  suggests that the reduced parallaxes are less sensitive to outliers than the unweighted means.

Additional information may be gained from a comparison of the results obtained from the theoretical bias fitting in Sect. 4 and the data in Table 2. The value given by the fitting are in very good agreement with those derived from the “unbiased plateau” for  $\pi > 25$  mas. This coincidence suggests that the sample is complete to about 25 mas.

Let us further compare these two approaches. First, in order to estimate, by any method, the mean absolute magnitude from an “unbiased plateau”, this “plateau” must be available. Inspection of Fig. 2 suggests that a definite plateau can barely be detected from the mean values in our case. Moreover, a small number of intrinsically brighter (or fainter) stars may easily disturb the derived means, especially if the unweighted means are used, as can be seen in Fig. 5. The small fraction (say, 10 to 20%) of objects in the unbiased plateau is also well known in extragalactic astronomy. One solution is to have a sufficiently large sample, which not always attainable. An advantage of the fitting is that it uses information also from the incomplete part of the sample.

Another point is the derivation of the width of the luminosity function  $\sigma$ . With the unbiased plateau, one calculates an unweighted mean, with  $\sigma$  being the rms deviation around the

mean. Each derived  $M_V$  has its own error arising from the errors in apparent magnitude and parallax. All information about these errors is lost when the unweighted mean is used. The scatter in derived  $M_V$  reflects both intrinsic width and observational errors. These effects are mixed in the unweighted mean. On the other hand, the least squares fitting effectively utilizes individual errors and gives  $M_0$  and  $\sigma$  with their associated standard errors that reflect the accuracy of data.

## 6. Discussion and conclusions

We have studied a sample of 166 Hipparcos single K0V stars, with spectral type information taken from the Michigan Spectral Survey. The derived absolute magnitude distribution suggests that 40 stars (24%) have  $M_V < 4.5$  mag. These objects are probably mis-classified K0IV stars. Thus we confirm the conclusion drawn by OGS, who found a contamination of about 20%.

The treatment of the distance-dependent Malmquist bias developed by Teerikorpi (1975) is found to be well applicable to the magnitude-limited subsample of 45 stars brighter than  $V = 9$ . To take advantage of this fact, an alternative method to determine the luminosity function is offered.  $M_0$  and  $\sigma$  are usually derived from the volume complete portion of a sample. In our method they are obtained by fitting a theoretical bias curve to the observational data. We find a good agreement between the methods. The fitting gives  $M_0 = 5.81 \pm 0.01$  and  $\sigma = 0.33 \pm 0.02$ , while the values derived from the unbiased plateau with  $\pi > 25$  mas are  $M_{UM} = 5.81$  and  $\sigma = 0.32$ , with a slightly brighter value,  $M_{RP} = 5.76$ , obtained with the method of reduced parallaxes. Summarizing these results, it may be concluded that the mean absolute magnitude of K0V stars is 5.8 mag with a scatter of 0.3 mag.

Our mean absolute magnitude is systematically fainter than the value of OGS, 5.69 mag. The difference depends on the method used and is about 0.1 mag for the unweighted mean and 0.05 mag for the reduced parallaxes. To clarify this discrepancy, we selected a sample with the same criteria as they used. Then we determined  $M_0$  from the samples containing all stars and single stars only. The means derived with double/multiple stars included are systematically brighter than those calculated from single stars only by the same amounts as mentioned above. Thus, accepting non-single stars, OGS obtained an underestimated  $M_0$ . By excluding binaries we reject another source of contamination thus making the calibration more realistic. Moreover, with this rejection the sample should better meet the required assumption of a Gaussian luminosity function. This also confirms that “the Hipparcos photometry for double and multiple stars had to be treated with particular care” (ESA 1997, vol. 1, Sect. 1.3).

In our investigation we have neglected the Lutz-Kelker bias. A detailed study of its interaction with the Malmquist bias is beyond the scope of the present paper. Although the combined effect of the distance-dependent Malmquist bias and the Lutz-Kelker one has not been studied systematically, some information may be gained from the numerical simulations made by Smith (2003) and Sandage & Saha (2002). They indicate that the net effect may be rather complicated depending on

features such as the spatial distribution of objects, apparent magnitude limit, parallax error law and others. Fortunately, in the case of faint and close K0V stars with low relative parallax errors the Lutz-Kelker bias may be neglected because it is almost fully dominated by the Malmquist bias.

*Acknowledgements.* The authors are very much indebted to the anonymous referee for thoughtful remarks that have contributed to the clarity of our presentation. A.G.B. is grateful to Tuorla Observatory for warm hospitality during his short-term visit. This study has been supported by the Academy of Finland (the projects “Fundamental questions of observational cosmology” and “Study of circumstellar and interstellar medium with polarimetry”).

## References

- Arenou, F., & Luri, X. 1999, in Harmonizing Cosmic Distance Scales, ed. D. Egret, & A. Heck, ASP Conf. Ser., 167, 13
- Arenou, F., & Luri, X. 2002, in Highlights of Astronomy, ed. H. Rickman, 12, 661
- Bottinelli, L., Fouqué, P., Gouguenheim, L., Paturel, G., & Teerikorpi, P. 1987, A&A, 181, 1
- Brown, A. G. A., Arenou, F., van Leeuwen, F., Lindegren, L., & Luri, X. 1997, in Hipparcos-Venice '97 Proceedings, ESA SP-402, 63
- ESA 1997, The Hipparcos and Tycho catalogues, ESA SP-1200
- Feast, M. 2002, MNRAS, 337, 1035
- Lebreton, Y. 2000, ARA&A, 38, 35
- Lutz, T. E., & Kelker, D. H. 1973, PASP, 85, 573
- Oudmaijer, R. D., Groenewegen, M. A. T., & Schrijver, H. 1998, MNRAS, 294, L41
- Oudmaijer, R. D., Groenewegen, M. A. T., & Schrijver, H. 1999, A&A, 341, L55, (OGS)
- Reid, I. N. 1999, ARA&A, 37, 191
- Sandage, A. 1994, ApJ, 430, 1
- Sandage, A. 1995, in The deep universe, ed. B. Binggeli, & R. Buser (Berlin: Springer), 1
- Sandage, A., & Saha, A. 2002, AJ, 123, 2047
- Smith, H. 2003, MNRAS, 338, 891
- Teerikorpi, P. 1975, A&A, 45, 117
- Teerikorpi, P. 1997, ARA&A, 35, 101