

On the use of the Hipparcos intermediate astrometric data^{*,**}

F. van Leeuwen and D.W. Evans

Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ, UK

Received April 30; accepted October 24, 1997

Abstract. With the publication of the Hipparcos catalogue, astrometric data of unprecedented quality and quantity has become available, providing parallaxes, positions and proper motions free from systematic errors down to a level of at least 0.1 mas and 0.1 mas y^{-1} . The Hipparcos catalogue, however, contains not only these “ready-to-use” positions, parallaxes and proper motions, but also intermediate astrometric data or abscissa residuals, which are the data from which the astrometric solutions were obtained. These data allow alternative solutions to be made for the astrometric parameters, for example, through the use of additional information. When combining data from stars in a small area on the sky, it becomes possible to account for correlations that exist between the abscissa residuals for stars measured on the same great circle. This is relevant for stars in open clusters and the Magellanic Clouds, where such correlations will be very frequent. The intermediate data also provide the possibility to add external constraints to an astrometric solution, such as an approximate but small parallax value, one that would have been too small to measure with Hipparcos. In that case the parallax can be fixed at the estimated small value, giving a better constrained solution for the proper motion. Similarly, when for a group of stars the absolute magnitudes are linked through a period-luminosity relation or by being all closely the same, as for RR Lyrae stars, such a condition can be superimposed on the parallax solution for all stars in this group, providing a distance scale calibration well beyond the range of direct parallax measurements. An example of how to use the data for solar system objects, which are provided only in the form of intermediate astrometric data, is shown.

Key words: methods: numerical — astrometry

1. Introduction

A full description of the Hipparcos data products and the reductions that led to these products has been presented in the Hipparcos and Tycho catalogues (ESA 1997). References made in the present paper to “Volumes” refer to volumes of that publication. In brief, Volume 1 provides a description of the data contents of the catalogue and Volume 3 descriptions of the Hipparcos data reduction and data verification processes. Volume 4 does the same for the Tycho data reductions. Volume 2 provides a description of the operations and is of little relevance to the present paper. A review of the data reduction procedures and their connections with the published data has been presented by van Leeuwen (1997), and will be referred to as VL97.

The Hipparcos catalogue is more than a “simple” catalogue of positions, parallaxes and proper motions. It also contains the data that were used to construct these quantities: the Intermediate Astrometric Data or abscissa residuals (Vol. 1, Sect. 2.8) – for further treatment of single stars; the Transit Data (Vol. 1, Sect. 2.9) – for further analysis of double stars; Solar System Objects abscissae (Vol. 1, Sect. 2.7); and the Hipparcos Epoch Photometry Annex (Vol. 1, Sect. 2.5) (records of the fully calibrated mean signal intensity and modulation amplitude). All of these are available as ASCII files on the Hipparcos CD-ROM set and can be used either on their own – for a better understanding of the astrometric data – or used in connection with ground-based observations to derive alternative or general solutions for individual objects or for groups of objects.

The abscissa data formed the basic astrometric measurements for single stars, where we define a single star as a star for which the modulated signal was not significantly disturbed by the presence of a component. The abscissa data were collected over intervals set by the orbital period of the satellite (10.6 hours) and referred to a reference great circle, defined for each orbital period by the position of its pole in the final International Celestial Reference

Send offprint requests to: F. van Leeuwen

* Based on observations made with the ESA Hipparcos satellite.

** Table 2 in its entirety is available at the CDS via anonymous ftp 130.79.128.5 or via <http://cdsweb.u-strasbg.fr/Abstract.html>

System (ICRS). The abscissae are one dimensional positional measurements, meridians of the great circle, where the position of the meridian is very well established (to an accuracy of a few mas), but the position of the object along the meridian is not known (except of course for the a priori knowledge of the positions of objects which were used in obtaining the observations). The abscissae, as observed over the length of the mission, were fitted with a set of astrometric parameters (usually position, parallax and a linear proper motion). Differences between observed and the predicted abscissae based on such models are referred to as abscissa residuals, and were made available in the *Intermediate Astrometric Data* file on the Hipparcos CD-ROM. For solar system objects the data were only presented in the form of abscissa data. The present paper focusses on when and how to use these intermediate astrometric data, for individual as well as groups of single stars.

Two reasons led to the decision to release the abscissa data: firstly, a small number of solutions provided in the final catalogue are not satisfactory, in particular when proper motions were fitted with a first or even second order function of time, or when a so-called stochastic solution had to be applied (see Sect. 2.4); secondly, correlations between errors on the abscissa estimates exist between measurements on the same reference great circle, most importantly when separations are less than 2-3 degrees (see Vol. 3, Chap. 17 and Sect. 2.2). The publication of the intermediate astrometric data has made it possible to incorporate the original Hipparcos abscissa data in a new solution when, at a later stage, additional observations clarify or constrain the nature of the object. Their publication also allows a proper determination of mean proper motions and parallaxes for stars in clusters and in the Magellanic Clouds, by allowing to fully take into account existing correlations.

The abscissa data allow for a complete reconstruction of the Hipparcos solutions for single stars (see Sect. 4). Thus, if for some reason at a later stage a particular abscissa measurement is considered suspect, then the solution can be repeated without this measurement. Similarly, abscissae rejected for the published solution are still included in the data file and can be recognized from the records and reintroduced if appropriate. If for some reason the parallax or the proper motion of an object is more accurately known already, a new solution can be obtained incorporating this information. In nearby star clusters, ground-based studies have over the years provided differential proper motions with precisions ten times higher than Hipparcos and could be used as external constraints on a new solution. Incorporating this kind of data is, however, not simple due to the distortions that appear to be present in such proper motion determinations (see e.g. Tian et al. 1996).

The present paper shows examples of how the intermediate astrometric data can be used in alternative mod-

els for the astrometric parameters of individual objects or groups of objects. These models are of two kinds: constrained solutions, where one or more parameters are supplemented with or replaced by information obtained from other sources than the Hipparcos Catalogue (Sect. 4); and “combined solutions”, where parameters for a selection of stars are solved for in one solution while, when necessary, removing correlations between abscissa residuals (Sect. 6). The latter method may also include constrained parameters, when part of the astrometric parameter solutions for a group of stars is replaced by the solution for a few common parameters, shared by all stars in that group. A simple example is the stars in the LMC. Instead of solving for each star a parallax and a proper motion, it becomes possible to eliminate the parallax from the solution (a parallax of 0.02 mas could not be detected by Hipparcos), and solve for a collective proper motion for all LMC stars together (see e.g. Kroupa & Bastian 1997). The possibility to combine astrometric solutions for groups of stars, provides, in addition, the means to detect intrinsically weaker signals in luminosity calibrations.

A brief review of those aspects of the Hipparcos data that have an immediate bearing on the use and understanding of the intermediate astrometric data is presented in Sect. 2. This is followed by Sect. 3 explaining the information provided in the intermediate astrometric data file, and the proper implementation of these data. Sect. 5 presents methods for analyzing data on the solar system objects. A more extensive review is presented by VL98, and in Volume 3.

2. Overview of the Hipparcos Astrometric data

This section provides a summary of data reduction aspects more fully described in Volume 3, primarily in Chapters 5, 9, 11, 14, 16 and 17; in fact, almost any chapter in Volume 3 has some bearing on the results used here. Many details can also be found in VL98. The emphasis will be on understanding the properties of the *Hipparcos Intermediate Astrometric Data* or abscissa records. There are four main aspects:

1. the reduction of the photon-counts (Sect. 2.1);
2. the great-circle reduction (Sect. 2.2);
3. the sphere reconstruction (Sect. 2.3);
4. the merging and determination of the astrometric parameters (Sect. 2.4).

In addition, Sect. 5 provides some background to the way the solar system data were obtained and are presented in the catalogue.

A very important aspect of the Hipparcos data reductions was the use of two independent data reduction consortia, NDAC and FAST (see Perryman et al. 1997), each providing what they considered their best final results. The results from the two consortia were merged to form the final catalogue.

There were two kinds of errors affecting the consortia results: errors due to photon noise on the original measurements, and errors due to inadequacies in the calibration methods (instrument modelling) applied by the two groups. The first of these was the same for both sets of consortia results and showed as a correlation between residuals with respect to the accepted solution. Errors resulting from instrument modelling were considered largely uncorrelated due to the different reduction methods used by the two groups. However, some correlation may be expected for these errors too. By combining the results from the two consortia the influence of the uncorrelated (instrument modelling) errors was reduced. This was clearly observed from the parallax results obtained in the merged solution, which showed an improvement relative to the individual consortia results.

Another important aspect of the Hipparcos instrument was the scanning law and the two entrance pupils, projecting images on the same focal plane, observing only objects selected from the pre-defined Hipparcos Input Catalogue (ESA 1992). These provisions made it possible to measure very precisely large angular distances on the sky and allowed for the determination of a rigid optical reference frame and the measurement of absolute parallaxes. This all-sky rigidity and reliability makes it possible to combine the intermediate astrometric data from all parts of the sky in a single solution into an essentially distortion free all sky (or small field) solution. The two entrance pupils did imply, however, that occasionally observations were disturbed by an image from the other field of view.

2.1. The photon-count reductions

The main signal of the Hipparcos observations was obtained from the sampling of the light of stellar images passing over a modulating grid of 2688 lines. The sampling used a photon counting image dissector tube (IDT), which used a small sensitive area (30 arcsec diameter, the instantaneous field of view) that could follow a stellar image during its transit through the 0.9 by 0.9 degrees field of view and also be moved very quickly from one object to another. Thus, almost simultaneous observations were obtained for up to 10 stars at any one time. The modulated signal for a single stellar image could be accurately described by a zero-level and first and second harmonic modulation, with well calibrated relations between the amplitude ratio and phase difference of the first and second harmonic. The phase v of the modulated signal (either based on the first harmonic only, as was done by NDAC, or based on the first and second harmonic, as was done by FAST) provided a transit time estimate across a reference slit: the fundamental input for the astrometric measurements. Differences between the two methods were largely eliminated from the final results through the calibrations of instrument parameters, but do reflect in differences of accuracies between FAST and NDAC abscissae

(see VL98). The transit times were related to positions on the sky through the reconstructed satellite attitude. The position of the reference slit with respect to which the phase was determined was derived from the satellite attitude and the a-priori or updated catalogue position. The astrometric data was further reduced by the Great Circle Reduction process to one measurement (abscissa) per orbital period (see next section).

The mean signal level and modulation amplitude were processed in the photometric reductions and provided the data for the $H_{p_{dc}}$ and $H_{p_{ac}}$ magnitudes respectively. The reduced photometric data were combined in field transit magnitudes, which are presented in the *Hipparcos Epoch Photometry Annex* (HEPA) and the *Hipparcos Epoch Photometry Annex Extension* (HEPAE).

The modulated signal could be affected by images close enough to the target image to be visible at the same time by the IDT. Such images could be either due to duplicity of the star or to accidental superimposition of an image from the other field of view. The composite image would still fit the same modulation model, but the relations between the mean intensity level and the modulation parameters were altered. The modulation phases were no longer directly linked to the transit time of an image, and special processing was required to handle observations of the double and multiple stars, while data associated with accidental superimpositions had to be discarded. A related source of signal disturbance was caused by stray light, resulting from very bright stars at larger distances (up to a few hundred arcsec). These effects were corrected for approximately. Exact corrections were impossible due to the very limited knowledge of the sensitivity of the IDT instantaneous field of view at larger distances from its centre.

Information on recognized accidental superimpositions by one or more images from the other field of view is provided in the *Hipparcos Epoch Photometry Annex Extension* file, HEPAE. This information can be related to the data in the *Intermediate Astrometric Data* file through a comparison of epochs: the astrometric reference epoch 1991.25 corresponds to photometric epoch JD 2448349.0625. It has to be realized, however, that in combining field transit data to abscissae individual data points that were affected by spurious images from the other field of view were in many cases rejected.

A further source of signal disturbance, although for only a very small number of objects, was due to the presence of planetary nebulae around some stellar images. These could disturb the signal depending on the scan direction in a way that is difficult to reconstruct or interpret. The average effect was a relative decrease of the modulation amplitudes of the signal, which can be recognized from the HEPA/HEPAE files by comparing the magnitudes derived from the zero-level intensities (dc -magnitudes in the HEPA file) with the simultaneously derived magnitudes from the modulation amplitude (ac -magnitudes in the HEPAE file). Disturbance by

a planetary nebula leads to too bright *dc*-magnitudes in comparison with the *ac*-magnitudes. Deviations from circular symmetry led to distortions on the phase estimates, and few of these objects have reliable astrometric solutions.

Detailed analysis of the first and second harmonics (phases and amplitudes) in the modulated signal led to the discovery of several thousands of double stars. Signals for double stars were processed separately by both NDAC and FAST, but only FAST carried all these signals along into the great-circle reductions. For this reason, only FAST abscissae are available for most of the double stars. It should be realized, however, that, depending on the magnitude difference and the separation, the interpretation of these double star abscissae can often be ambiguous due to the complexity of the signal. The Transit Data file (Vol. 1, Sect. 2.9) has preserved the case history files for 35 535 known or suspected double or multiple stars as obtained by NDAC, permitting a revised interpretation of these data too (Vol. 1, Sect. 2.9).

2.2. The great-circle reductions

The aim of the great-circle reductions was to obtain from the modulation phases v_i , obtained from the IDT signals over a period of 4 to 8 hours, precise abscissae on a reference great circle (van der Marel & Petersen 1992). In the process the instrument parameters, describing the relation between a position on the sky and a position on the modulating grid, were calibrated. The most noticeable of the instrument parameters was the basic angle between the two fields of view. The great-circle reduction process used star positions, initially taken from the Hipparcos Input Catalogue (ESA 1992) and later from preliminary mission results, together with orbital parameters for the Earth and the satellite and the reconstructed attitude of the satellite. This information was used to transform the phases v to the proper slit positions on the modulating grid, thus obtaining preliminary abscissae from the phase measurements. The great-circle reduction process determined the scan phase of the instrument as a function of time, and relative to this scan-phase the averaged star abscissae. Between 5 and 90 measurements could contribute to a single abscissa determination. The great-circle reduction process can be summarized by the following equation (which was applied to every single scan-phase determination):

$$\Delta G_{ik} = \frac{\partial G_{ik}}{\partial v_i} \Delta v_i + \frac{\partial G_{ik}}{\partial \psi_k} \Delta \psi_k + \frac{\partial G_{ik}}{\partial \mathbf{d}'} \Delta \mathbf{d}' + \epsilon_{ik}, \quad (1)$$

where G_{ik} is the grid coordinate of the star (the mean position on the grid during the observation as derived from its apparent position, the scan phase estimate and the reconstructed satellite attitude). ψ_k is the along-scan attitude correction and \mathbf{d}' is the vector of instrument parameters. The very smooth motions of the satellite (except at times of thruster firings) allowed for the use of

cubic splines to fit locally the corrections $\Delta \psi_k$ to the original star-mapper-based attitude reconstruction, and thus to reconstruct very precisely the abscissae along the great circle. However, the attitude corrections used the same abscissa data, and as a result there are correlations between the errors on the final abscissae and the attitude corrections. This propagated into correlations of abscissae errors for stars affected by the same attitude errors. Due to the two fields of view, abscissae errors for stars very close together on the sky, as well as for stars separated by 58 degrees (the basic angle) and multiples thereof, are found to be correlated (see Fig. 1). A preliminary study of these correlations was presented in Volume 3, Chapter 17. The correlations were re-investigated at a higher spatial resolution and taking into account the projection of the stellar separation to an abscissa difference. Also investigated was the influence of the length of the time interval covered by the data included in each great-circle reduction run. It was expected that correlations would be much stronger for short sets than for long sets. As the actual length of the data stretches was not available, the number of stars per great circle was used instead as an indicator of long and short sets. There were other aspects too, that affected the quality of the great-circle results, but these are difficult to reconstruct from the published data. They concern gaps in the data due to occultations (a major problem for great circles with small inclinations with respect to the ecliptic), and problems with the attitude reconstruction due to high background levels. Most of these problems reflect in individual abscissa accuracies.

Only stars with a standard 5-parameter solution were used in the determination of the correlations. On each great circle there are mostly between 900 and 2000 such stars (extremes run from 27 to 2110 stars for NDAC, and 295 to 2027 stars for FAST). Only in one situation were these correlations both significant and able to accumulate and affect a discussion of Hipparcos astrometric data: for stars in a small field (a few degrees diameter, like an open cluster or the Magellanic Clouds). For any other separation the correlation between measurements for a pair of stars seldom repeated themselves over the mission, and the cumulative effect was very small (stars at a separation of 180° also accumulated a correlation, but at that separation the correlations were rather small). The strength of the correlations diminished when the time span covered by the data became longer. The increase in data decreased the degrees of freedom for the along-scan attitude improvements. The actual time span covered by each RGC is not recorded in the data files, but reflects in the number of stars included in each RGC. Figure 2 shows the correlations for short separations and for different ranges in dataset length. In particular for NDAC the increase in the correlations was strong for shorter datasets, reflecting one of the differences in the data reduction approach. The correlations were fitted with a polynomial in even powers of s , the abscissa separation measured in units of 4 degrees.

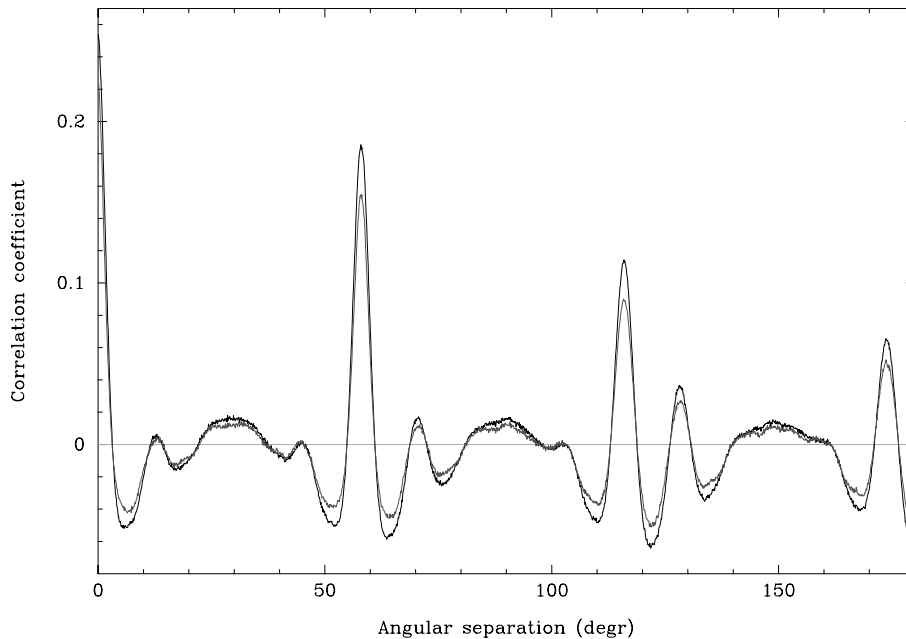


Fig. 1. The correlation coefficient of the abscissa residuals, as a function of the separation on the reference great circle, for all datasets. The NDAC curve can be distinguished from the FAST curve by its lower minima, and higher maxima, differences that become more pronounced as datasets get shorter

Table 1. Functional representations of the correlation coefficients at short abscissa distances for different lengths of datasets. The lengths of the sets are indicated by n_5 , which is the number of abscissae residuals accepted from 5-parameter solutions (Col. I in Table 2)

$\langle n_5 \rangle$	Consort.	1	s^2	s^4	s^6	s^8	s^{10}	s^{12}
600	FAST	0.5931	-0.7801	0.7846	-0.4452	0.1313	-0.0189	0.00105
	NDAC	0.5678	-0.7292	0.5518	-0.2545	0.0653	-0.0085	0.00044
900	FAST	0.3290	-0.6231	0.5589	-0.2817	0.0771	-0.0106	0.00058
	NDAC	0.4152	-0.7237	0.6163	-0.3050	0.0836	-0.0117	0.00064
1200	FAST	0.2423	-0.5591	0.5266	-0.2689	0.0741	-0.0103	0.00056
	NDAC	0.2962	-0.5792	0.4721	-0.2229	0.0594	-0.0081	0.00044
1500	FAST	0.2041	-0.4931	0.4652	-0.2343	0.0636	-0.0087	0.00047
	NDAC	0.2362	-0.4962	0.4103	-0.1925	0.0506	-0.0068	0.00037
1800	FAST	0.1860	-0.4747	0.4742	-0.2500	0.0701	-0.0099	0.00054
	NDAC	0.2006	-0.4518	0.3961	-0.1920	0.0512	-0.0070	0.00038
2100	FAST	0.1790	-0.4837	0.4964	-0.2632	0.0735	-0.0102	0.00056
	NDAC	0.1759	-0.3913	0.3703	-0.1928	0.0538	-0.0075	0.00041

The fits only cover the separation range 0 to 6 degrees, i.e. s ranging from 0 to 1.5. The results of those fits are summarized in Table 1. Table 2 (here only represented by an extract of the complete file, which is available electronically via the CDS) provides for each reference great circle the numbers of accepted and rejected abscissae. These data can also be used as an indicator of (the very few) generally unreliable reference great circles, by comparing the numbers of accepted and rejected observations. Section 6 shows how these correlations can be incorporated in a determination of a common proper motion or parallax for a group of stars with small separations on the sky.

The result of the great-circle reductions was a set of 2341 great circles. They cover a time-span of 2768 orbits or 1230 days. Not every great circle was reduced by both consortia. Due to a tape delivery problem that was detected too late, the NDAC reductions are not available for 4 RGCs towards the end of the mission, while in a few cases an RGC is missing in the FAST reductions due to problems with the data reductions. In most cases this concerned RGCs with small numbers of stars. Instrument parameters were not solved for when numbers of stars were low. They were interpolated from neighbouring, better determined solutions. For 2247 RGCs data is available from both consortia; for 15 RGCs data is only available from

Table 2. Numbers of abscissa residuals per orbit, split into three types of solutions: (I) type 5; (II) types 7, 9 and X; and (III) types C, V, O and $-$. In the first two cases the numbers of accepted and rejected residuals are given. For the third case only the number of abscissae was available. Only an extract of the Table is presented here. The full version is available in electronic form through the CDS

Orbit	NDAC					FAST				
	I		II		III	I		II		III
	accep.	rejec.	accep.	rejec.	accep.	accep.	rejec.	accep.	rejec.	accep.
	(n_5)					(n_5)				
1	1364	6	48	2	121	0	0	0	0	0
48	1270	6	45	0	91	1169	3	42	0	129
50	1393	0	64	2	104	789	0	41	0	100
51	958	1	38	0	74	965	1	40	0	142
52	1207	9	64	0	110	1051	3	58	0	161
53	1160	2	57	0	90	944	2	50	0	128
54	1200	6	54	0	93	1156	4	56	0	163
55	1245	7	56	0	102	1218	3	56	0	169
56	1282	7	73	1	114	1175	4	72	0	180
57	1400	3	70	0	112	896	4	51	0	119
59	1166	27	50	1	93	860	174	41	5	130
60	959	3	36	0	83	0	0	0	0	0
63	1314	8	64	0	129	1325	3	66	0	195
64	1475	4	59	0	115	1461	4	60	0	192
65	1497	5	51	0	136	981	3	35	0	162

the FAST consortium; while for 79 RGCs data is only available from the NDAC consortium.

2.3. The sphere solution

The main task for the sphere solution (Vol. 3, Chap. 16) was to establish reference zero points for all reference great circles, and to remove or calibrate any features left behind by the preceding processing. Although, as part of the sphere reconstruction, astrometric parameters were calculated, these are not the parameters presented in the catalogue. They were used to check the consistency between the solutions of the two consortia and to detect any grid-step ambiguities left over from the great-circle reduction. The result of the sphere reconstruction was, therefore, the original great-circle reduction data, with calibrated zero points and corrected systematic defects.

A comparison between the final Hipparcos and Tycho results seems to indicate the presence of grid-step ambiguities for 57 stars in the final catalogue (Vol. 4, Chap. 11). These stars can be solved for again by using the Tycho data as starting points and allowing corrections of multiples of ± 1.2074 arcsec on some or all of the abscissae.

2.4. Merging and astrometric parameter determination

Before any merging of data took place, the results from the two consortia had to be rotated to a common reference frame. This was done through the use of orthogonal rotations in positions and proper motions. As a first

step, the formal errors on the FAST and NDAC data were investigated as functions of magnitude and quoted errors. The quoted errors were adjusted statistically to give the expected unit weight variances. Next, the correlation between the FAST and NDAC abscissa residuals were determined and applied. Astrometric solutions were made using the abscissae obtained by both consortia by incorporating the correlation coefficients. All solutions were tested for the necessity to allow a non-linear proper motion. In this process apparently outlying residuals or pairs of residuals were removed, and these can be recognized as such in the abscissa records. Solutions were accepted as either the standard 5-parameter model (two positional parameters, parallax and two proper motion parameters), the 7-parameter model (proper motion changing linearly with time) or in exceptional cases the 9-parameter model (proper motion changing quadratically with time). When none of these models provided an acceptable solution, and the star was not recognized as a double star, a so-called stochastic solution (indicated by “X”) was applied. In this solution, the 5-parameter model was implemented to the observed abscissae, but with the estimated errors on these abscissae artificially increased by adding quadratically “cosmic noise” until a satisfactory solution was obtained. The level of “cosmic noise” added is preserved in the DMSA part X, described in Sect. 2.3 of Volume 1. Any such solution has to be treated with great care. Likely interferences causing this “cosmic noise” are orbital motion (Bastian & Bernstein 1995; Bernstein 1997) and the presence of a planetary nebula. In all these cases the information provided in the Intermediate Astrometric

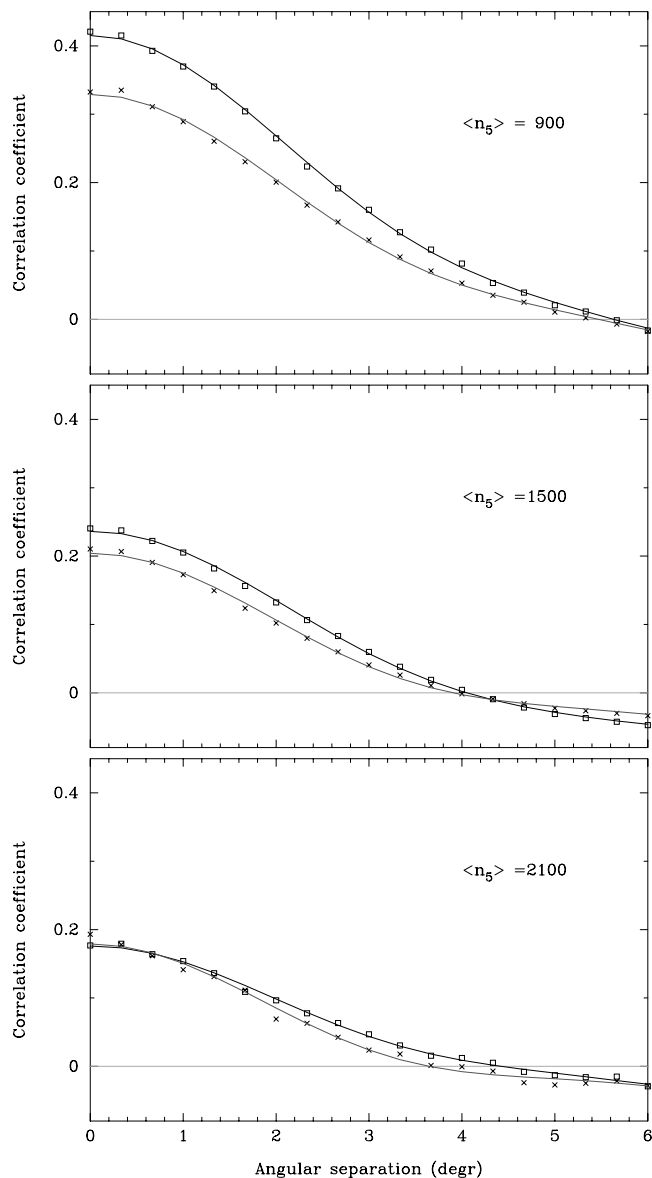


Fig. 2. The abscissa residuals correlation coefficient for small separations in data sets of different lengths. Top: short datasets; middle: medium length data sets; bottom: long datasets, as defined in Table 1. Crosses represent FAST data, open squares represent NDAC data. Also shown are the fits as given in Table 1

Data file allows for a full reconstruction of the solution and its covariance matrix through the mechanism described in Sect. 3. Stars with solutions of type “O” (orbital solutions) or “-” (no astrometric solution) may also use the abscissae records. This is not the case for two other types of solutions, indicated with “C” and “V”. These represent a component solution and a “variability induced mover” respectively. The latter type stands for a small number of objects where duplicity was inferred by a photocentric motion caused by the variability of one of the components.

Finally, all results were transformed to the International Celestial Reference System (ICRS). This transformation was based mainly on very high accuracy radio positions and proper motions for a small set of radio stars (see Vol. 3, Chap. 18 and Kovalevsky et al. 1997).

3. The intermediate astrometric data or abscissa residuals

3.1. The general solution

The details of the intermediate astrometric data file are described in Volume 1, Sect. 2.8, the file itself is available on disc 5 of the ASCII CD-ROM set. The file contains the data from which the astrometric parameters for single stars, as presented in the main catalogue, were derived through the fitting of one of the models described in Sect. 2.4. The model used is indicated for each star, both in the main catalogue and the intermediate astrometric data file.

A model fitted to the abscissae is described in the form of the expected changes in the abscissa positions, v_i , as a result of changes in each of the model parameters a_j . Collecting all abscissa residuals in a vector \mathbf{v} , the following relation is obtained:

$$\Delta \mathbf{v} = \sum_{j=1}^n \frac{\partial \mathbf{v}}{\partial a_j} \Delta a_j + \epsilon, \quad (2)$$

which is solved through minimizing the χ^2 , the sum of the squares of the residuals:

$$\chi^2 = \left(\Delta \mathbf{v} - \sum_{j=1}^n \frac{\partial \mathbf{v}}{\partial a_j} \Delta a_j \right)' \mathbf{V}^{-1} \left(\Delta \mathbf{v} - \sum_{j=1}^n \frac{\partial \mathbf{v}}{\partial a_j} \Delta a_j \right), \quad (3)$$

as a function of the parameter corrections Δa_j . \mathbf{V} is the covariance matrix of the abscissa residuals. In the Hipparcos Catalogue the standard errors on the parameters have been calculated assuming the predicted χ^2 value, i.e. using an assumed standard error for the solution of 1.0 rather than the observed value, which may be larger or smaller than 1.0. In the presence of the variance matrix $\mathbf{U} = \mathbf{V}^{-1}$, the minimization of χ^2 results in $k = 1, n$ equations of the type:

$$\sum_{j=1}^n \left(\frac{\partial \mathbf{v}}{\partial a_k} \right)' \mathbf{U} \left(\frac{\partial \mathbf{v}}{\partial a_j} \right) \Delta a_j = \left(\frac{\partial \mathbf{v}}{\partial a_k} \right)' \mathbf{U} \Delta \mathbf{v}, \quad (4)$$

which constitute the so-called normal equations. The matrix \mathbf{U} is symmetric and can be factorized as:

$$\mathbf{U} = \mathbf{T}'\mathbf{T}, \quad (5)$$

where \mathbf{T} is a lower triangular matrix (see e.g. Bierman 1977). The matrix \mathbf{T} is called the square root of \mathbf{U} , and can be obtained using e.g. the Cholesky decomposition algorithm, as described by Bierman. After replacing \mathbf{U} by

$\mathbf{T}'\mathbf{T}$ and some reorganization the following equations are obtained ($k = 1, n$):

$$\sum_{j=1}^n \left(\mathbf{T} \frac{\partial \mathbf{v}}{\partial a_k} \right)' \left(\mathbf{T} \frac{\partial \mathbf{v}}{\partial a_j} \right) \Delta a_j = \left(\mathbf{T} \frac{\partial \mathbf{v}}{\partial a_k} \right)' (\mathbf{T} \Delta \mathbf{v}). \quad (6)$$

Thus, by multiplying the left- and right-hand sides of the observation equations, as defined by Eq. (2), by the matrix \mathbf{T} , which is the Cholesky square root of the inverse of the covariance matrix \mathbf{V} , a set of de-correlated and properly weighted observation equations is obtained, ready to be incorporated in the traditional least-squares solution. Note that the square root of a matrix is not unique, and therefore different ways exist to de-correlate the same set of observation equations. Also, when the covariance matrix is diagonal, the square root is too, and the multiplication of the observation equations reduces to a straight forward weighting by the inverse of its standard error for each observation.

3.2. The covariance matrix

The observation equations for a single star consist in fact of two sets of correlated observation equations, one set from each consortium. Data obtained for observations on reference great circle k by FAST will be indicated by F_k , and for NDAC by N_k . The counter i (or j) refers to an observation equation independent of its origin. The covariance matrix \mathbf{V} for a single star solution is built up as follows. The diagonal elements are given by:

$$V_{ii} = \sigma_{N_k}^2, \quad V_{jj} = \sigma_{F_k}^2 \quad (7)$$

for observation i obtained by NDAC and observation j obtained by FAST. If two observations i and j originate from NDAC and FAST reductions on the same reference great circle k , additional off-diagonal elements have to be added to the covariance matrix, representing the correlation between these observations:

$$V_{ij} = V_{ji} = q_{NF_k} \sigma_{N_k} \sigma_{F_k}, \quad (8)$$

where q_{NF_k} is the correlation coefficient between the FAST and NDAC reduction results for orbit number k . The values of q_{NF_k} have been determined empirically from the FAST and NDAC abscissa residuals as function of magnitude, time and the estimated standard errors on the abscissa by FAST and NDAC and are provided in the abscissa records.

When only these consortia correlations are considered, the correlations between observations can be incorporated in the observation equations by multiplying pairs of correlated observations equations with the Cholesky square root of the inverse of their covariance matrix:

$$\mathbf{T} = \begin{pmatrix} \frac{1}{\sigma_{F_k}} & 0 \\ \frac{-q_{NF_k}}{\sigma_{F_k} \sqrt{1-q_{NF_k}^2}} & \frac{1}{\sigma_{N_k} \sqrt{1-q_{NF_k}^2}} \end{pmatrix}, \quad (9)$$

where the first equation is assumed to originate from FAST, and the second from NDAC. The resulting pair

of observations are uncorrelated and properly weighted, and can be treated as any other observation equations in a least squares solution. It is clear that, if the correlation coefficient q_{NF_k} is equal to zero, Eq. (9) reduces to a simple scaling of each equation by the inverse of the square root of the variance of the standard error on the observation it represents.

3.3. Astrometric parameter models

The abscissa file provides $\frac{\partial v}{\partial a_i}$ for the solution of the five astrometric parameters: corrections to the mean position ($d\alpha \cos \delta$, $d\delta$), the assumed parallax ($d\pi$) and the assumed linear proper motions ($d\mu_\alpha \cos \delta$, $d\mu_\delta$). Provisions have also been made for variable proper motions, the so-called 7- and 9-parameter solutions (see Sect. 2.4 and Vol. 1, Sect. 2.8). The information for the construction of the covariance matrix \mathbf{V} (standard errors and correlation coefficients) is included in the abscissa records (fields IA9 and IA10 in Table 2.8.3 of Vol. 1).

The abscissa residuals presented in the intermediate astrometric data file are given relative to an implementation of the five astrometric parameters given in the header records. These five parameters may be part of a more complicated solution, such as a 7- or 9-parameter solution or an orbital motion solution. In these cases the abscissa residuals were obtained through implementing only the first five parameters of the solution, as given in each header record. In the 7- and 9-parameter solutions the first- and second-order time dependence of the proper motion were referred to a reference time such that the effect of these parameters on the basic 5-parameter model would be very small. This leads to the following results:

- In case the original solution was a 5-parameter solution, the proper application of a 5-parameter solution to the residuals must give negligible corrections to the astrometric parameters (only rounding-off errors), and reproduce the standard errors on those parameters and their covariances as given in the main catalogue;
- In the case of 7- or 9-parameter or orbital motion solutions, the application of the original solution must produce negligible corrections for the first 5-parameters, and reproduce the remaining parameters, together with all standard errors and the full covariance matrix. When applying a 5-parameter solution to the residuals of the partially implemented 7- or 9-parameter solution, the expected results are small corrections to the 5-parameters and an increase in the standard error of the solution.
- In the case of a stochastic solution (“X” in field H59 of the main catalogue, and in field IH8 in the header records of the intermediate astrometric data), the original solution can be recovered by adding in quadrature a “cosmic noise” to the standard errors of the abscissae, and applying the standard 5-parameter solution to

the result. The level of the “cosmic noise” was adjusted such as to give the expected χ^2 for the solution.

In the case of double star solutions (types “C” and “V”) the abscissa residuals (often only given for the FAST data) are mostly of little meaning, and should not be used. For these stars the Transit Data (Vol. 1, Sect. 2.9) should be used instead.

4. Individual solutions

Most individual solutions of the 5-parameter type are unlikely to be improved upon. Exceptions are stars for which it can safely be assumed that their parallax is much too small to be measured by Hipparcos, in which case the parallax can be set to zero. This could be relevant when a relatively large correlation exists between the parallax and proper motion determination. In order to do so, all abscissa residuals are to be corrected for the difference between the determined parallax and the assumed parallax:

$$\Delta v' = \Delta v - \frac{\partial v}{\partial \pi} \Delta \pi, \quad (10)$$

where $\Delta \pi = \pi_{\text{new}} - \pi_{\text{old}}$. The new set of residuals are then solved using the mechanism described in Sect. 3 for only four parameters a_i : position and proper motion.

The situation is much more complicated when a very precise proper motion is available from an external source. The problem is that before such a constraint can be incorporated, it needs to be represented in exactly the same reference frame as the Hipparcos data. The same applies when positions are obtained on photographic plates at a different epoch than the Hipparcos data. The transformation to the Hipparcos reference frame can only be obtained if a sufficiently large number of objects is available for the determination of the transformation parameters, e.g. when the measurement has been obtained in a sufficiently well determined reference frame, and has been corrected for any systematic differences between that reference frame and the Hipparcos reference frame, both in positions and in proper motions. In general, the uncertainty of the systematic errors will inhibit incorporation of proper motion data. Here, however, the Tycho catalogue can be of assistance in some cases. With a much higher density of stars than the Hipparcos catalogue, the Tycho positions can provide a sufficient number of reference points to determine accurate plate transformations if plates were obtained at an epoch close to the mean Tycho epoch. For clusters like the Pleiades and Praesepe one can use the fact that the internal proper motions of the cluster members are very small, and that by using only cluster members, larger epoch differences can be allowed.

Another possible application is for stars with orbital motions, where prolonged observations can provide additional measurements. It is likely that some of the 7- and 9-parameter solutions refer to orbital motions over a time-span much longer than the 3.5 years of the Hipparcos

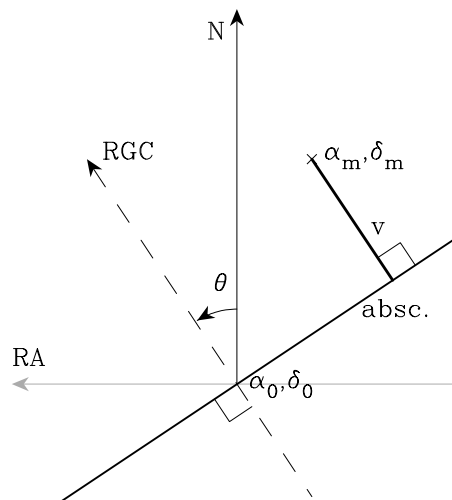


Fig. 3. Definition of the abscissa offset v for a predicted position (α_m, δ_m) , with respect to the catalogue position (α_0, δ_0) and RGC orientation θ

mission. The same applies to some stochastic solutions, where the orbital motion time-scale could be shorter than 3.5 years. Further observations may constrain the possibilities of non-linear proper motion fits. It will be difficult, however, to obtain positional measurements with accuracies comparable with the Hipparcos data. In such cases, the addition of radial velocity data may sometimes help resolve the solution of the Hipparcos data.

Important in the construction of orbital parameter solutions is the definition of the parameters a_i and their derivatives. This problem is very similar to that of the solar system objects, described in the next section, except that the angle θ (Fig. 3), defining the local direction to the equatorial pole, still needs to be calculated. Although it can be derived from the positions of the star and the pole of the reference great circle, it is far easier to obtain the angle θ directly from fields IA3 and IA4 (Vol. 1, Table 2.8.3) in the intermediate astrometric data file:

$$\tan \theta = \text{IA3/IA4}, \quad (11)$$

while the epoch of observation is taken from the Reference Great Circle file (fields IR2 and IR5 in Vol. 1, Table 2.8.1) or from:

$$\text{epoch} - 1991.25 = \text{IA6/IA3} = \text{IA7/IA4}. \quad (12)$$

Using the same mechanism as employed for solar system objects, described in the next section, further models can be developed and implemented.

5. Solar system objects

The solar system objects observed by Hipparcos (48 minor planets, two of Jupiter’s moons, Callisto and Europa, and one of Saturn’s moons, Titan), do not lend themselves for a simple parametrized representation. They could only be presented in the form of the abscissa results, with one

observation per field of view crossing. Although an actual position is given, the only value of this position is to serve as a reference point for the abscissa measurement. The actual data-point could be found anywhere close to this line. Figure 3 shows how from the published data (described in Vol. 1, Sect. 2.7) and a predicted position (α_m, δ_m) the abscissa residual v can be derived. Assuming that the distances between the published and the predicted positions are small (generally less than 1 arcsec), the Euclidian approximations can be made:

$$v = (\alpha_m - \alpha_0) \cos \delta \sin \theta + (\delta_m - \delta_0) \cos \theta. \quad (13)$$

The handling of these measurements is very similar to the handling of the star abscissae: a model predicts abscissa positions and their dependence on the model parameters. We take e.g. the best available ephemerides of an object to calculate predicted positions $(\alpha_{m|i}, \delta_{m|i})$. The abscissa residuals v_i for these predicted positions (Eq. 13) can be represented in various ways: either as a function of some critical parameters in the model, or simply as observed offsets in coordinates as a function of time:

$$dv = d\alpha \cos \delta \sin \theta + d\delta \cos \theta + \epsilon, \quad (14)$$

with $\langle \epsilon^2 \rangle^{\frac{1}{2}} = \sigma_v$. By representing $d\alpha$ and $d\delta$ as functions of parameters a_i , such as time or orbital elements, Eqs. (14) are solved through minimizing $\sum_i \left(\frac{v_i - dv_i}{\sigma_{v_i}} \right)^2$. Here, as in all other similar solutions, no correlations are assumed to exist between data on different reference great circles.

It should be realized, however, that for the determination of both $d\alpha$ and $d\delta$, measurements are required from different reference great circles, providing the different angles θ needed to remove their correlation.

6. Combined solutions

As was explained in Sect. 2.2, correlations exist between abscissa residuals on the same reference great circle, in particular for observations with differences between the abscissa values of less than 4 degrees. These correlations can only be fully accounted for through the use of the abscissa data in combined solutions for the astrometric parameters of groups of stars. Combined solutions are essential for solving mean parallax and proper motion values in a relatively small area of the sky (density of objects more than 0.2 per square degree). In addition, combined solutions offer the best possibilities for incorporating constraints obtained from other data.

In a combined solution one starts by collecting all the abscissa data for the stars involved, and sort these on orbit number (there is one reference great circle, or RGC, per orbit). Covariance matrices are determined and applied per reference great circle. The reference great circle data file (Table 2.8.1 in Vol. 1) provides the pole positions for the NDAC and FAST RGCs, which are transformed to unit

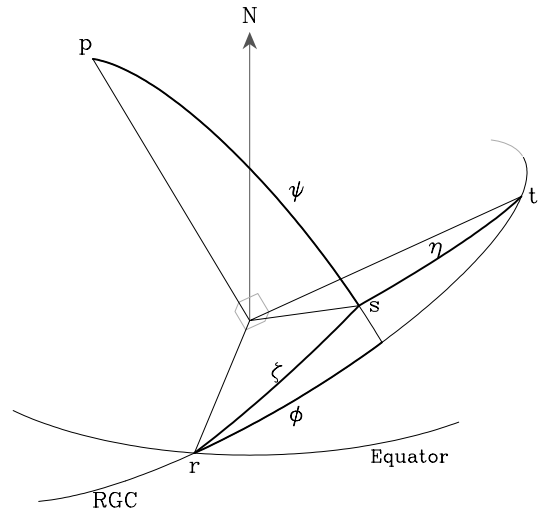


Fig. 4. Definition of the angles and directions used to calculate the relative abscissae along the reference great circle

vectors in the equatorial reference system, \mathbf{p}_N and \mathbf{p}_F . A unit length reference direction \mathbf{r} (Fig. 4) on each RGC can be obtained e.g. from its crossing with the Equatorial plane:

$$\mathbf{p} \equiv \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \quad (15)$$

from which:

$$\mathbf{r} = \frac{1}{\sqrt{p_1^2 + p_2^2}} \begin{pmatrix} p_2 \\ -p_1 \\ 0 \end{pmatrix}. \quad (16)$$

A vector \mathbf{t} completes the triad $(\mathbf{r}, \mathbf{t}, \mathbf{p})$:

$$\mathbf{t} = \frac{1}{\sqrt{p_1^2 + p_2^2}} \begin{pmatrix} -p_1 p_3 \\ -p_2 p_3 \\ p_1^2 + p_2^2 \end{pmatrix}. \quad (17)$$

Given the unit vector \mathbf{s} directed towards the reference position of a star (the effect of aberration and other smaller effects have been removed in the great-circle reduction process and would have been almost identical for neighbouring stars), we calculate the angles ζ and η (see Fig. 4):

$$\cos \zeta_j = \mathbf{s}' \mathbf{r}_j \quad (18)$$

$$\cos \eta_j = \mathbf{s}' \mathbf{t}_j \quad (19)$$

and ψ :

$$\cos \psi_j = \mathbf{s}' \mathbf{p}_j, \quad (20)$$

where j equals N or F , depending on the origin of the observation. The abscissa ϕ for this star is then derived from:

$$\begin{aligned} \cos \phi_j &= \cos \zeta_j / \sin \psi_j \\ \sin \phi_j &= \cos \eta_j / \sin \psi_j. \end{aligned} \quad (21)$$

Although for large distances between stars the difference between the abscissae will be almost equal to the actual

distance between those stars on the sky, this is not necessarily the case for small distances (less than a few degrees), where the correlations are the strongest.

The abscissa separation between two stars, $\Delta\phi_{jk}$, measured by the same consortium on the same RGC, is translated into a correlation coefficient, using the functions given in Table 1 (with n_5 as defined in Sect. 2.2 and F_c a flag indicating the relevant consortium):

$$Q_{jk} = f(F_c, n_5, \Delta\phi_{jk}), \quad (22)$$

which produces a coefficient in the covariance matrix:

$$V_{jk} = Q_{jk}\sigma_j\sigma_k. \quad (23)$$

Although no direct correlations exist between data for different stars on the FAST and NDAC RGCs, secondary correlations or covariances do occur as a result of the correlations between measurements of the same stars by NDAC and FAST. One of the differences between a natural correlation and a covariance due to a secondary correlation is that the first kind affects the total amount of information (“total weight”) of the observations, while the second kind does not.

The exact values of these covariances are difficult to estimate due to possible small correlations between attitude errors, but their approximate values can be derived as follows. Assume three sets of unit weight residuals, $\epsilon_{k,i}$, ($k = F, N$) and $\epsilon_{k,j}$, ($k = F$). Say that there are natural correlations between $\epsilon_{F,i}$ and $\epsilon_{N,i}$ (same observation, different reductions), such that $\langle \epsilon_{F,i}\epsilon_{N,i} \rangle = Q_{FN_{ii}}$. Similarly, a natural correlation exist between $\epsilon_{F,i}$ and $\epsilon_{F,j}$ (same reference great circle, different stars), given by $Q_{FF_{ij}}$. No natural correlation exist between $\epsilon_{N,i}$ and $\epsilon_{F,j}$, but due to the other correlations, a covariance will occur, which can be approximated by substituting $\epsilon_{F,j} = Q_{FF_{ij}}\epsilon_{F,i} + \nu_{F,j}$, which results from the first correlation. Thus, as $\epsilon_{N,i}$ and $\nu_{F,j}$ are uncorrelated, we find for the covariance $\langle \epsilon_{N,i}\epsilon_{F,j} \rangle \approx Q_{FF_{ij}}Q_{FN_{ii}}$, giving the following element in the covariance matrix:

$$V_{FN_{ij}} = Q_{FF_{ij}}Q_{FN_{ii}}\sigma_{F_i}\sigma_{N_j}. \quad (24)$$

This covariance is only an approximation; in fact a slightly different value can be obtained by using a different link between the variables:

$$V_{FN_{ij}} = Q_{NN_{ii}}Q_{FN_{ij}}\sigma_{F_j}\sigma_{N_i}. \quad (25)$$

The effect of introducing these covariances in the covariance matrix is to preserve the proper weight reduction caused by the natural correlations: they cancel out when inverting the covariance matrix, leaving at the diagonal elements only the effects of the natural correlations. In the example given above the covariance matrix looks like:

$$\begin{pmatrix} 1 & q_{12} & q_{13} \\ q_{12} & 1 & q_{12}q_{13} \\ q_{13} & q_{12}q_{13} & 1 \end{pmatrix}, \quad (26)$$

where we substituted $Q_{FF_{ij}} = q_{12}$, $Q_{FN_{ii}} = q_{13}$, and assumed all variances to be equal to 1. A Gauss elimination produces:

$$\begin{pmatrix} 1 & q_{12} & q_{13} \\ 0 & 1 - q_{12}^2 & 0 \\ 0 & 0 & 1 - q_{13}^2 \end{pmatrix}. \quad (27)$$

The inverted, variance matrix then reads as follows:

$$\begin{pmatrix} 1 + \frac{q_{12}^2}{1 - q_{12}^2} + \frac{q_{13}^2}{1 - q_{13}^2} & \frac{-q_{12}}{1 - q_{12}^2} & \frac{-q_{13}}{1 - q_{13}^2} \\ \frac{-q_{12}}{1 - q_{12}^2} & \frac{1}{1 - q_{12}^2} & 0 \\ \frac{-q_{13}}{1 - q_{13}^2} & 0 & \frac{1}{1 - q_{13}^2} \end{pmatrix}, \quad (28)$$

from which is obtained the Cholesky square root, as defined by Eq. (5), and by means of which the observations are weighted:

$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{-q_{12}}{\sqrt{1 - q_{12}^2}} & \frac{1}{\sqrt{1 - q_{12}^2}} & 0 \\ \frac{-q_{13}}{\sqrt{1 - q_{13}^2}} & 0 & \frac{1}{\sqrt{1 - q_{13}^2}} \end{pmatrix}. \quad (29)$$

Thus, in the final weighting of the observations, the covariances as produced by the secondary correlations (products of q_{12} and q_{13}) have disappeared, and only the influences of the natural correlations remain.

After inverting \mathbf{V} and taking its Cholesky square root, it can be applied to the original observations to obtain a set of uncorrelated and properly weighted observation equations. This is done for the data in each RGC, covering the observations of both consortia (when available). These de-correlated observations can then be incorporated in a classical least-squares solution.

The observational equations used for each star can either be the original equations, in which case corrections to the individual parameters are found, and a set of uncorrelated parameters is determined, or they can contain common parameters such as a common proper motion and/or a common parallax. When solving for a common parameter, it is essential that all abscissa residuals are corrected to represent a reference solution referring the abscissa residuals for all stars involved in the combined solution to the same parallax and/or proper motion values. These corrections are obtained using Eq. (10). Thus, for the LMC-stars example, the residuals are corrected to reference values of zero proper motion and a parallax of 0.02 mas, and then solved together for a common proper motion. The combined solution provides corrections to the reference values. It is also possible to use all observations, ignoring rejections from the standard processing, and determine new rejections under the new conditions, based on the residuals relative to the combined solution. When determining the mean proper motion of a star cluster, it is possible to incorporate precise differential proper motions obtained on the ground as constraints for the solution. In all cases, the degrees of freedom available will be strongly reduced, thus improving the reliability of the solution.

In the solutions for a star cluster the space velocity rather than the proper motion should be considered constant. Without the presence of internal motions this condition can be used to determine individual distances and

radial velocities of cluster members, as was shown for the Hyades cluster by Dravins et al. (1997). For a cluster like the Pleiades, however, with a higher internal velocity dispersion but a larger distance and a small radial velocity, the differential distance variations and the projection of the radial velocity can be ignored, and the shared space velocity of cluster members can be expressed as a function of the proper motion of the centre of the cluster (indicated by the subscript “c”) and the position of each object on the sky, relative to the cluster centre (see VL98 for a derivation):

$$\begin{aligned}\mu_\alpha \cos \delta &\approx (\mu_\alpha \cos \delta)_c + (\mu_\delta)_c \sin \delta \sin(\alpha - \alpha_c) \\ \mu_\delta &\approx (\mu_\delta)_c - (\mu_\alpha \cos \delta)_c \sin \delta \sin(\alpha - \alpha_c).\end{aligned}\quad (30)$$

The effects are, as one would expect, most noticeable close to the equatorial poles ($\sin \delta \approx 1$ and relatively large variations in α for relatively small angular separations) and for clusters covering large parts of the sky. There are different ways to solve for $((\mu_\alpha \cos \delta)_c, (\mu_\delta)_c)$. One way is to first transform all abscissae residuals back to zero proper motion, and then implement Eqs. (30) in the solution for each star. This then gives, as the proper motion part of the solution, the following equation (where $a_4 = \mu_\alpha \cos \delta$ and $a_5 = \mu_\delta$):

$$\begin{aligned}d\mathbf{v} &= (\mu_\alpha \cos \delta)_c \left(\frac{\partial \mathbf{v}}{\partial a_4} - \sin \delta \sin(\alpha - \alpha_c) \frac{\partial \mathbf{v}}{\partial a_5} \right) \\ &+ (\mu_\delta)_c \left(\frac{\partial \mathbf{v}}{\partial a_5} + \sin \delta \sin(\alpha - \alpha_c) \frac{\partial \mathbf{v}}{\partial a_4} \right).\end{aligned}\quad (31)$$

Alternatively, a first-order approximation is made for the cluster proper motion, from which the local projection corrections are calculated. In a subsequent iteration, further adjustments of the cluster proper motion will have negligible influence on these projection corrections. Only when the cluster is spread out over a large part of the sky, like the Hyades, must these and other corrections be fully taken into account.

Other possibilities of combined and constrained solutions may not involve small fields, and can be simpler to apply as they would not require de-correlating observations of different stars. Accidental correlations that could occur because of two stars appearing on the same reference great circle will be small and rather rare, and therefore of very little influence. Correlations between the FAST and NDAC reductions of the same observations are removed using Eq. (9). This kind of solution can be used for luminosity calibrations.

As an example, consider determining the mean absolute magnitudes of RR Lyrae stars in a single solution, using the parallax information contained in the abscissa residuals and the reddening-corrected apparent mean magnitudes. The procedure goes as follows. First assume a reasonable value for the absolute magnitude $\hat{M}_V(RR)$, related in a well defined manner (passband) to the reddening corrected mean apparent magnitude

$< m_v >$ of each star. Translate the difference into a predicted parallax:

$$\hat{\pi} = 100 \times e^{-0.4605(m_v - \hat{M}_V(RR))}, \quad (32)$$

where $\hat{\pi}$ is measured in mas. All abscissae are corrected for the difference $\Delta\pi = \hat{\pi} - \bar{\pi}$ between the published ($\bar{\pi}$) and the calculated predicted parallax (Eq. 10). The estimated absolute magnitude $\hat{M}_V(RR)$ will differ by a small amount from the best estimated value indicated by the observations:

$$\tilde{M}_V(RR) = \hat{M}_V(RR) + \Delta M_V(RR). \quad (33)$$

This correction to the absolute magnitude will in the first approximation result in a scaling correction of the assumed parallaxes:

$$\tilde{\pi} = \hat{\pi} \times (1 - 0.4605 \Delta M_V(RR)) \equiv \hat{\pi}(1 - \nu), \quad (34)$$

giving corrections to the assumed individual parallaxes of $-\hat{\pi}\nu$, where $\hat{\pi}$ is the estimated parallax of an individual RR Lyrae star, and ν a correction factor which is the same for all RR Lyrae, if our basic assumption that the absolute magnitudes of these stars are equal is indeed correct. Thus, we can now create one solution for all RR Lyrae, solving for each star the position and proper motion corrections, and solving for one overall parallax correction factor ν . The solution provides an accuracy of the estimated value $\tilde{\nu}$ and thus for $\tilde{M}_V(RR)$, and will tell how good the model is (at the available accuracy of the observations) through its χ^2 value. By using this method rather than the direct interpretation of the published parallaxes, a more reliable and better to interpret solution is obtained, involving fewer degrees of freedom and deriving the critical quantity of the distance scale correction directly from the actual observations: the abscissa data. However, as the correction factor ν is multiplied by the estimated parallax, most of the weight in the solution will still come from the nearest stars. As long as the accuracy of the corrected apparent magnitudes is high, any bias on the derived parallaxes will be very small, and no bias is expected from the solution. This was confirmed by a simple simulation. A more general description of luminosity calibrations using this method can be found in VL98. A different method for obtaining unbiased luminosity estimates has been presented by Luri et al. (1996), and is based on a maximum likelihood estimate for a set of probability distributions.

An estimate of the expected accuracy of $\tilde{\nu}$ can be obtained from:

$$\sigma_\nu \approx \sqrt{\frac{1}{\sum_i \hat{\pi}_i^2 (\mathbf{T}_i \frac{\partial \mathbf{v}_i}{\partial \pi})' (\mathbf{T}_i \frac{\partial \mathbf{v}_i}{\partial \pi})}}, \quad (35)$$

where the index i represents the different stars, and \mathbf{T}_i is the Cholesky square root of the variance matrix for star i . This can be approximated with the following simplification to:

$$\sigma_\nu \approx \sqrt{\frac{1}{\sum_i \left(\frac{\hat{\pi}_i}{\sigma_{\pi_i}} \right)^2}}, \quad (36)$$

where σ_{π_i} is the parallax accuracy given in the final catalogue, but $\hat{\pi}_i$ is the estimated parallax value as defined above, and NOT the value given in the catalogue.

7. Application examples

The examples presented in this section are intended to show how the methods presented in the present paper can be incorporated in various studies. They do NOT intend to be exhaustive presentations of these studies. The emphasis is always on how the Hipparcos intermediate astrometric data can be used, or in some cases, should be used.

7.1. The absolute magnitude of RR Lyrae stars

The Hipparcos variability annex gives 38 RRc and 146 RRab stars. Of the latter there is one which was newly defined as RRab, and which now can be rejected as such on the basis of its parallax: HIP 100859, which has a parallax of 26 mas, giving it an absolute magnitude of 2.9. Furthermore, RRc types with periods below 0.3 days were found to show often fainter absolute magnitude (see also VL98 on the classification of A and F type periodic variables). Leaving out those cases, 155 RRab and RRc stars were left. These stars provide 9345 individual abscissae measurements, from which, without taking into consideration the effects of reddening, a parallax based absolute magnitude of $M_{\text{Hp}} = 0.94 \pm 0.19$ was obtained. This is not significantly different from kinematically derived parallax estimates by Fernley et al. (1997), who note, however, that some of the stars in the selection used here are not confirmed independently as RR Lyrae stars. The unit weight standard deviation of the solution is 1.021, which, given the number of observations, is close to significantly different from the expected value of 1.00. The consistency between the magnitude based parallaxes and the parallaxes obtained from the individual solutions is shown in Fig. 5.

7.2. The PL relation of Mira stars

The Mira stars in the Magellanic Clouds show in the K band a period luminosity relation. The slope of this relation was used by van Leeuwen et al. (1997b) in combination with Hipparcos parallaxes to derive a zero point for this relation. A major problem with the Mira stars is the effect of the very large brightness and colour variations on the astrometric measurements. Most of the brighter Miras received so-called “V” type solutions, implementing a brightness related correction to the abscissae residuals.

We started off with the same selection of 16 stars as used by van Leeuwen et al. (1997b), and removed from this list two stars suspected of being fundamental mode pulsators, a symbiotic, a C-type and a double mode Mira,

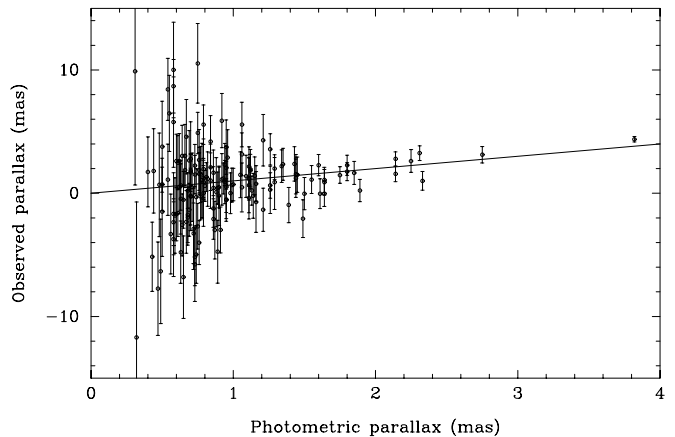


Fig. 5. The relation between the photometric parallaxes, calibrated using Hipparcos abscissa data, and the observed individual parallaxes with their estimated errors for RRab and RRc stars. The solid line represents the expected one-to-one relation

and tested all solutions for a standard 5-parameter solution. Three stars showed very bad fits according to the unit weight variances of their residuals, and were removed. There remained 8 stars, for which the reddening corrected “K” magnitudes were fitted using:

$$M_K = -3.47 \log P + \beta_1. \quad (37)$$

Predicted parallaxes were calculated for assumed values of β_1 . Corrections to this assumed value reflected in a scaling parameter for the parallaxes, which was solved for all 8 stars simultaneously using their 465 abscissae records. No corrections for brightness variations were introduced. The value thus obtained is $\beta_1 = 0.65 \pm 0.16$, slightly smaller than what was given by van Leeuwen et al. (1997b), but this may be related to ignoring the above mentioned corrections. The need for these corrections was reflected in the unit weight standard error, which measured 1.35.

7.3. The Magellanic Clouds

A preliminary calculation of the proper motions of the Large and Small Magellanic Clouds was carried out for the paper by Kroupa & Bastian (1997). The covariance values have since been recalculated in more detail (see Sect. 2.2), and were in most cases found to be smaller than shown in Figs. 17.11 and 17.12 of Volume 3, Chapter 17.

The situation for the LMC and SMC is simple. The Hipparcos Catalogue contains 31 members of the LMC and 8 members of the SMC for which a single star solution of type “5” or “7” was obtained. Given the distances of the Magellanic Cloud stars, it was assumed that solutions of type “7” were spurious. The astrometric solutions for these stars were obtained from in total 2229 abscissa measurements in the LMC and 514 in the SMC. Each star had a parallax and proper motion derived in its individual solution. The first task was to remove the parallax

Table 3. Preliminary results for common proper motion solutions for stars in the LMC and SMC. The correlation coefficients ρ between the proper motion in declination and right ascension are also given

	LMC		SMC	
Number of stars	31		8	
Number of abscissae	2229		514	
$\mu_{\alpha \cos \delta}$ (mas/yr)	1.60	± 0.29	1.13	± 0.77
μ_{δ} (mas/yr)	0.19	± 0.37	-1.17	± 0.66
$\rho_{\mu_{\alpha \cos \delta}}^{\mu_{\delta}}$	-0.116		0.135	
uw standard error	1.090	± 0.016	1.056	± 0.033

and proper motions and to obtain abscissae relative to a zero proper motion and parallax (using Eq. 10). The abscissae were sorted on orbit number and for every orbit a set of de-correlated observation equations was obtained (corrected for the consortia correlations and the correlations described in Sect. 2.2). The observation equations described corrections to individual star positions and to a common LMC or SMC proper motion, but did not allow for a parallax solution. These de-correlated observation equations were accumulated in a single least squares solution, (using Eq. (31) for the proper motions) providing the results shown in Table 3. From the unit weight standard errors it is clear that the condition of a single common proper motion over the LMC may not be entirely satisfied, producing a χ^2 significantly above the expected value. The situation is better for the SMC. The differences between the values presented here and those presented earlier by Kroupa & Bastian have no significant effect on the discussions presented in that paper.

7.4. The Pleiades and Praesepe

The determination of the Pleiades parallax and proper motion will be dealt with in detail by van Leeuwen and Hansen (in preparation). Here we present the preliminary results of that study and a similar application to stars in the Praesepe cluster.

The Hipparcos catalogue contains 60 members of the Pleiades cluster. Of these, 54 were selected as single stars with solution type “5”, providing a total of 2182 abscissae (see van Leeuwen & Hansen-Ruiz 1997). All abscissae were incorporated, including those rejected in the individual star solutions. In an iteration over the combined solution a total of 4 abscissa residuals were rejected. All abscissae were corrected to a reference parallax and proper motion, and sorted on orbit number. For each orbit a set of de-correlated observation equations was created, which were combined in one least-squares solution. The observation equations described positional corrections for all 54 stars and one parallax and proper motion for the cluster centre, using Eq. (31). Thus, the degrees of freedom were reduced

Table 4. Preliminary results for common proper motion and parallax solution for stars in the Pleiades and Praesepe clusters. The various correlation coefficients ρ are also given

	Pleiades		Praesepe	
Number of stars	54		22	
Number of absc.	2182		940	
π (mas)	8.61	± 0.23	5.04	± 0.36
Distance modulus	5.32	± 0.05	6.49	± 0.15
$\mu_{\alpha \cos \delta}$ (mas/yr)	19.67	± 0.24	-35.61	± 0.40
μ_{δ} (mas/yr)	-45.55	± 0.19	-12.54	± 0.26
$\rho_{\mu_{\delta}}^{\pi}$	-0.100		-0.192	
$\rho_{\mu_{\alpha \cos \delta}}^{\pi}$	-0.042		-0.165	
$\rho_{\mu_{\delta}}^{\mu_{\alpha \cos \delta}}$	0.244		-0.078	
uw standard error	1.058	± 0.016	1.053	± 0.024

from $54 \times 5 = 270$ to $3 + 54 \times 2 = 111$. The results are summarized in Table 4. From the unit weight standard error and its uncertainty it appears that there remained unmodelled effects in the parallax and proper motion. This is most likely the internal proper motion dispersion in the cluster, which has a dispersion in the projected centre of the cluster of around 1 mas s^{-1} . The correlation coefficient of $\rho_{\mu_{\alpha \cos \delta}}^{\mu_{\delta}} = 0.244$ reflects the proximity of the Pleiades to the ecliptic and limited range of scan-directions resulting from this situation.

Table 4 also provides similar results for the Praesepe cluster. Here there is no significant contribution from an internal proper motion dispersion.

8. Conclusions

By providing the users with the intermediate astrometric data in the form of the actually observed abscissa residuals, a range of uses of the Hipparcos data has become possible that allows proper determination of statistical quantities, as well as the combination of Hipparcos data with other information, providing general determinations with improved reliability of quantities like cluster proper motions and parallaxes, as well as for luminosity calibrations.

Acknowledgements. The main reason that we have availability of the intermediate astrometric data was an initiative by C.A. Murray around 1993, when he explained that merging of the astrometric data could only be done reliably when using these abscissae data. C.A. Murray and F. Arenou developed and implemented the algorithms for the merging of the data, and F. Arenou provided some samples and further explanations. Discussions with L. Lindegren about some of the ideas presented here are also gratefully acknowledged. Discussions with U. Bastian and the implementation of the methods presented here to the Magellanic Clouds measurements were very beneficial. L.V. Morisson read the manuscript and indicated various possible small improvements. A number of mistakes were pointed out by Anthony Brown. Finally, the implementation in the distance determination of the Pleiades and the

involvement of Cristina S. Hansen Ruiz in that project, also greatly contributed to the present paper. The referee provided constructive suggestions and comments which were much appreciated.

Appendix A: Summary of the implementation method

In the following appendices a set of algorithms is presented as they were used in various applications of the methods described in the present paper. The algorithms are built using a vectorial storage of symmetric and upper or lower triangular arrays, implicitly using the knowledge of symmetry or zero elements. This reduces storage needs and in some cases facilitates implementation. All algorithms are given in Fortran 77.

As a first step, the abscissae for all stars to which a combined and/or constrained solution is to be applied are extracted from the intermediate astrometric data file, where they are sorted on star number and for each star on orbit number. The abscissa file thus created is then sorted on orbit number and within each orbit on star number.

Appendix B: Vector storage of arrays

The relation between the one-dimensional vectors (here symbolized by \mathbf{v}) and the matrices (\mathbf{M}) they represent is as follows (i represents the rows, j the columns):

$$v(k) = M(i, j) \quad \text{for all } j \geq i \text{ and } k = j(j-1)/2 + i \quad (\text{B1})$$

In a symmetric matrix the additional condition

$$M(j, i) = M(i, j) \quad (\text{B2})$$

applies, and in a lower triangular matrix:

$$M(i, j) = 0 \quad \text{for all } j < i. \quad (\text{B3})$$

Appendix C: Building up of the covariance matrix

A covariance matrix is created for the data contained in each orbit. To each abscissa record is assigned a flag $F(J)$, which tests as `.TRUE.` for data from FAST, and as `.FALSE.` for data from NDAC. Abscissae are calculated using the mechanism described in Sect. 6, giving for each record a value $ABSC(J)$. Each record is in addition identified by its star number, $HIP(J)$. The covariance matrix is called CV and is stored in vector form as described above. The standard error on the abscissa residual is given by $SRES(J)$ and the correlation coefficient with the same measurement as reduced by the other consortium by $CCOR(J)$. The function $ABSCORR(F(J), LE, DIS)$ returns the correlation coefficient $ACOR$, which applies to data obtained within one consortium ($F(J)$) with a given dataset length of $N5$ and an abscissae distance of DIS between two stars. Then, given N abscissa records for current orbit:

```

I1 = 0
DO K = 1, N
C
C diagonal element
C
I1 = I1 + K
CV(I1) = SRES(K) * SRES(K)
DO L = K-1, 1, -1
L1 = K*(K-1)/2 + L
IF(HIP(L) .EQ. HIP(K)) THEN
C
C same observation, different consortium
C
CV(L1) = CCOR(K)*SRES(K)*SRES(L)
ELSE
DIS = ABS( ABSC(K) - ABSC(L) )
ACOR = ABSCORR(F(L),N5,DIS)
IF(F(L) .EQ. F(K)) THEN
C
C same consortium
C
CV(L1) = ACOR*SRES(K)*SRES(L)
ELSE
C
C different consortia, secondary correlation
C
CV(L1) = CCOR(K)*ACOR*SRES(K)*SRES(L)
END IF
END IF
END DO

```

Appendix D: The square root of the inverse of the covariance matrix

There are various ways in which the square root of the inverse of the covariance matrix can be obtained. A numerically stable and simple method is by taking the square root of the covariance matrix first, and then inverting this lower triangular matrix. Alternatively, one could invert the covariance matrix, and then take the square root (but this could involve less numerically stable calculations).

The algorithm for the Cholesky factorization of the symmetric matrix CV is as follows:

```

M = N*(N+1)/2
DO I = 1, M
LT(I) = CV(I)
END DO
DO I = N, 2, -1
LT(M) = SQRT(LT(M))
K = 0
M = M - I
L = M
DO J1 = 1, I-1
L = L+1
LT(L) = LT(L)/LT(M)
DO J2 = 1, J1
K = K+1
LT(K) = LT(K) - LT(M+J2)*LT(L)
END DO

```

```

      END DO
    END DO
    LT(1) = SQRT(LT(1))

```

from which is obtained the lower (or upper) triangular matrix LT. The determinant of the covariance matrix CV, which provides a measure of the total amount of information available in the observations, can be obtained from the product of the squared diagonal elements of LT. The matrix LT is inverted using:

```

      K = 0
C
C   Creating unit matrix
C
      DO I = 1, N
        K = K + 1
        LI(K) = 1.0
        DO J = 2, I
          K = K + 1
          LI(K) = 0.0
        END DO
      END DO
C
C   Starting inversion process
C
      M = N*(N+1)/2
      DO I = N, 1, -1
        L = M
        DO J1 = I, 1, -1
          LI(M) = LI(M) / LT(L)
          K1 = M
          DO J2 = J1-1, 1, -1
            L = L-1
            K1 = K1-1
            LI(K1) = LI(K1) - LI(M)*LT(L)
          END DO
          M = M-1
          L = L-1
        END DO
      END DO

```

giving the square root LI of the inverse of the covariance matrix CV.

Appendix E: Application of the square root matrix

Given an array P(N,NPAR) and a vector O(N) representing the left and right hand side of N observation equations with NPAR parameters, then the multiplication of the N by N matrix LI with P and O produces a set of decorrelated observation equations (PO,OO):

```

      L=1
      DO J=1,N
C
C   initialization of PO, OO
C
      DO I = 1, NPAR
        PO(J,I) = 0.0DO
      END DO
      OO(J) = 0.0DO

```

```

C
C   de-correlation processing
C
      DO K = 1, J
        DO I = 1,NPAR
          PO(J,I) = PO(J,I) + LI(L) * P(K,I)
        END DO
        OO(J) = OO(J) + LI(L) * O(K)
        L = L+1
      END DO
    END DO

```

The new set of observation equations can now be used in e.g. a Householder transformations based least squares solution (see Bierman 1977).

References

- Bernstein H.-H., Bastian U., 1995, in: Future Possibilities for Astrometry in Space, Perryman M.A.C., van Leeuwen F. (eds.) ESA SP-379, p. 55
- Bernstein H.-H., 1997, in: Hipparcos, Venice'97, Perryman M.A.C., Bernacca P.L. (eds.) ESA SP-402, p. 705
- Bierman G.J., 1977, Factorization Methods for Discrete Sequential Estimation, Mathematics in Science and Engineering, Vol. 128. Academic Press, New York
- Dravins D., Lindegren L., Madsen S., Holmberg J., 1997, in: Hipparcos Venice'97, Perryman M.A.C., Bernacca P.L. (eds.) ESA SP-402, p. 733
- ESA 1992, The Hipparcos Input Catalogue, ESA SP-1136
- ESA 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200
- Fernley J., Barnes T.G., Skillen I., et al., 1997, in: Hipparcos Venice'97, Perryman M.A.C., Bernacca P.L. (eds.) ESA SP-402, p. 653
- Kovalevsky J., Lindegren L., Perryman M.A.C., et al., 1997, A&A 323, 620
- van Leeuwen F., Evans D.W., Grenon M., et al., 1997a, A&A 323, L61
- van Leeuwen F., Feast M.W., Whitelock P.A., Yudin B., 1997b, MNRAS 387, 955
- van Leeuwen F., Hansen-Ruiz C.S., 1997, in: Hipparcos Venice'97, Perryman M.A.C., Bernacca P.L. (eds.) ESA SP-402, p. 689
- van Leeuwen F., 1997, Space Sci. Rev. 81, 201
- Luri X., Mennessier M.-O., Torra J., Figueras F., 1996, A&AS 117, 405
- van der Marel H., Petersen C.S., 1992, A&A 258, 60
- Kroupa P., Bastian U., 1997, New Astron. 2, 77
- Perryman M.A.C., Lindegren L., Kovalevsky J., Høg E., et al., 1997, A&A 323, L49
- Tian K.P., van Leeuwen F., Zhao J.L., Su C.G., 1996, A&AS 118, 503