17. ASTROMETRIC CATALOGUE MERGING

Once the FAST and NDAC final sphere solutions were obtained, the results for each star had to be merged in order to provide the final astrometric parameters together with their associated covariance matrix. Instead of a simple weighted average of the astrometric parameters, this merging was done at the great-circle level. The astrometric parameters of each consortium were first transformed to a common reference frame. Then the FAST and NDAC abscissa residuals with respect to a reference astrometric solution were studied in order to calibrate empirically their correlation coefficient. Finally, for each star, the abscissae and the calibrated covariance matrix between them were used to obtain a merged solution. This merging was applied to the standard solutions of apparently single stars, using five astrometric parameters, but also to accelerated solutions, orbital astrometric binaries and stochastic solutions.

17.1. Introduction

The series of successive sphere solutions described in Chapter 16 ended with two catalogues, F37.3 and N37.5, obtained respectively by the FAST and NDAC Consortia. These catalogues were final in the sense that little improvement was expected from further iterations.

Figures 16.23–16.28 illustrate how the correlation between the FAST and NDAC results increased with the amount of satellite data (from 12 to 37 months of data) and the number of iterations. This can be understood with the following rough model: the random errors on the obtained astrometric parameters consisted of one component attributable mainly to the photon noise of the raw data, and which was therefore common to FAST and NDAC, and another component including the 'modelling errors' proper to each reduction procedure. The second component decreased with the improvement brought by each iteration; it did not vanish however, so that the final correlation between the consortia results was less than one, or about 0.7 on the average.

For this reason, a statistical improvement of the astrometric parameters could be expected from a merging of FAST and NDAC results. This was indeed the case with the merged H18 and H30 Catalogues (Figure 16.22); however these preliminary catalogues were obtained with a simple average of the FAST and NDAC astrometric parameters, with an equal weight for both. Clearly, better results were foreseen using a more adequate weighting, depending both on the standard errors of the FAST and NDAC

parameters and on the correlation between them. The optimal weight would be different for each star, and possibly also different for each of the five parameters. Assuming that it were possible to compute these weights optimally, there would however still remain another problem, namely how to compute the complete covariance matrix of the averaged parameters.

A better scheme was proposed by C.A. Murray. Going back one step in the data reductions, the astrometric parameters of a given star were estimated in each consortium by a least-squares solution in which the abscissae determined by that consortium were regarded as independent 'observations' of the star. If now the FAST and NDAC abscissae were taken together, and considered as correlated observations in a new least-squares solution for the astrometric parameters, this would give not only the optimally combined parameters, but also the correct covariance matrix for these data.

From this general principle, successive merged catalogues were created, with various improvements brought at each step: in order to combine the FAST and NDAC abscissae, the weights of the abscissae had to be revised, and the correlation between the abscissae was empirically calibrated using an unbiased estimator; finally the least-squares procedure which combined the abscissae was adapted for robustness and in order to produce the different types of solutions required (the standard five-parameter solution and the G, O and X type solutions described in Volume 1, Section 2.3). The last catalogue created by this process was called H37C and was used for the link to the extragalactic system (Chapter 18); after rotation to the final Hipparcos reference frame this became the main body of the Hipparcos Catalogue.

17.2. Astrometric Parameters and Abscissa Residuals

The data provided by FAST and NDAC for the catalogue merging consisted of a superset of the final sphere solutions F37.3 and N37.5 described in Chapter 16. This superset contained values of the astrometric parameters $\mathbf{a} = (a_1 \ \ldots \ a_5)'$ for every star, even when no accepted solution had been found. In addition, the complete set of abscissa residuals Δv_j with respect to the given parameters was provided, together with the partial derivatives $\partial v_j / \partial a_i$ of the computed abscissae with respect to the astrometric parameters. (Recall that the star abscissa v is the angle, as seen from the reference great-circle pole, from the ascending node of the reference great circle on the equator to the satellitocentric coordinate direction of the star at the epoch of the reference great circle; see Figure 11.1.) The standard errors on the abscissae had been computed, as described in Chapter 9, from the formal propagation of the errors through the great-circle reduction procedure, but with empirical corrections derived in connection with the sphere solutions (Chapter 11).

In the following, subscripts $_{\mathrm{F}}$ and $_{\mathrm{N}}$ will be used to designate data referring to the individual consortia, and subscript $_{\mathrm{H}}$ will be used for the corresponding merged quantities.

Table 17.1 shows the number of stars, observations (abscissae), and reference great circles among the initial data given by the consortia, and among the data used for the astrometric merging. For NDAC, the data collected during one orbit defined a single reference great circle. In FAST, this was also generally the case, but a few orbits were split into two great circles. For the merging procedure, a one-to-one correspondence between FAST and NDAC data was needed. Consequently, in the case of split orbits,

Table 17.1 Statistics of the data used for the astrometric merging.

Data	Stars	Abscissae	Great circles
FAST	118 160	3 668 140	2 262
NDAC	111 708	3 570 643	2 326
Merge	118 300	7 227 401	2 341

a single FAST abscissa residual was computed from the weighted average of the two great-circle abscissa residuals, and the partial derivatives were simply averaged. Thus, for the merging, each star had at most one FAST and one NDAC residual on each orbit or reference great circle.

The discrepancy between the total number of stars from FAST and NDAC is due to the fact that the known double stars were not reduced using the main NDAC reduction chain. For this reason, the great-circle abscissae for some of the stars were not available for the merging procedure. The number of great circles used differ because the consortia applied different criteria for accepting a great-circle reduction.

In order to combine the FAST and NDAC results, F37.3 (after transformation to the equatorial frame) and N37.5 had to be rotated first to a common, provisional reference frame approximately aligned with the FK5 Catalogue. This frame was in practice defined by the H30 catalogue, so the orientation and rotation differences F37.3–H30 and N37.5–H30 were used for this purpose. After these corrections of $\mathbf{a}_\mathrm{F}$ and $\mathbf{a}_\mathrm{N}$, an initial set of merged parameters was computed as the simple average:

$$\mathbf{a}_\mathrm{H} = \tfrac{1}{2}(\mathbf{a}_\mathrm{F} + \mathbf{a}_\mathrm{N}) \qquad [17.1]$$

Successive merged solutions replaced the parameters $\mathbf{a}_\mathrm{H}$ by improved values.

The abscissa residual Δv is the difference between the observed abscissa and the abscissa calculated from a given set of astrometric parameters. The original abscissa residuals provided by the consortia could thus be written:

$$\begin{aligned} \Delta v_\mathrm{F}(\mathbf{a}_\mathrm{F}) &= v_\mathrm{F}^\mathrm{obs} - v_\mathrm{F}^\mathrm{calc}(\mathbf{a}_\mathrm{F}) \\ \Delta v_\mathrm{N}(\mathbf{a}_\mathrm{N}) &= v_\mathrm{N}^\mathrm{obs} - v_\mathrm{N}^\mathrm{calc}(\mathbf{a}_\mathrm{N}) \end{aligned} \qquad [17.2]$$

A direct comparison or averaging of these residuals would not be meaningful, as the reference parameter vectors $\mathbf{a}_\mathrm{F}$ and $\mathbf{a}_\mathrm{N}$ were always different. It was therefore necessary to compute new abscissa residuals with respect to the parameters $\mathbf{a}_\mathrm{H}$. Around each estimate, a linear expansion was used:

$$v^\mathrm{calc}(\mathbf{a} + \Delta\mathbf{a}) = v^\mathrm{calc}(\mathbf{a}) + \sum_{i=1}^{5} \frac{\partial v}{\partial a_i}\Delta a_i \qquad [17.3]$$

resulting in the new abscissa residuals:

$$\begin{aligned} \Delta v_\mathrm{F}(\mathbf{a}_\mathrm{H}) &= \Delta v_\mathrm{F}(\mathbf{a}_\mathrm{F}) - \sum_{i=1}^{5} \frac{\partial v_\mathrm{F}}{\partial a_i}(a_{\mathrm{H}i} - a_{\mathrm{F}i}) \\ \Delta v_\mathrm{N}(\mathbf{a}_\mathrm{H}) &= \Delta v_\mathrm{N}(\mathbf{a}_\mathrm{N}) - \sum_{i=1}^{5} \frac{\partial v_\mathrm{N}}{\partial a_i}(a_{\mathrm{H}i} - a_{\mathrm{N}i}) \end{aligned} \qquad [17.4]$$

For 21 nearby, high-velocity stars, the FAST residuals were also modified to take into account the secular acceleration (Volume 1, Section 1.2.8) which was already included in the NDAC residuals of these stars.

The end result of these transformations was a set of n_F residuals Δv_{Fj}, with associated formal standard errors σ_{Fj}, and a corresponding set of n_N residuals Δv_{Nj} with standard errors σ_{Fj}, both sets now referring to the same astrometric parameters $\mathbf{a}_H$.

As the residuals of a given star came from a previous five-parameter astrometric solution, they had lost five degrees of freedom from the total number of observations, $n = n_F + n_N$. When computing the unit-weight error, i.e. the sample standard deviation of the normalised residuals $\Delta v_{Fj}/\sigma_{Fj}$, etc., the multiplicative factor $\sqrt{n/(n-5)}$ had to be applied in order to obtain an unbiased estimate of the dispersion of the residuals. For the same reason, a correction $5/n$ should be added to the sample correlation coefficient between the normalised residuals.

In fact, the exact degree of freedom of statistics computed from the abscissae was not known: other general parameters had been determined by each consortium, such as the great-circle origins, and the reference parameters for the residuals were different from the initial solution obtained by the consortia. For this reason, the statistics used in the merging were, as far as possible, based on the differences of the residuals, which were equal to the difference between observed abscissae, $\Delta v_{Fj} - \Delta v_{Nj} = v_F^{obs} - v_N^{obs}$. This relation holds for all but the first merged solution, where it was only approximate.

17.3. Scaling Corrections of Consortia Formal Errors

Unbiased formal errors were needed both for the computation of the correlation coefficients, described in the next section, and more generally for the weight matrix of the astrometric least-squares solution. The residuals were systematically plotted as functions of all available data: orbit number, position, proper motion, magnitude, colour. For this purpose, only stars considered previously as single stars, with good solutions obtained both by FAST and NDAC, were kept (about 96 000 stars); the observations rejected during the astrometric solutions of the consortia were also rejected for this calibration (but not necessarily in the final merging). Finally, great circles with less than 10 observations common to FAST and NDAC were not used for the calibration (2242 orbits remained).

No systematic trend of the mean residuals was found as a function of astrometric or photometric data. However, the unit-weight errors of the abscissa residuals, $u_F = [\mathrm{Var}(\Delta v_F/\sigma_F)]^{1/2}$, $u_N = [\mathrm{Var}(\Delta v_N/\sigma_N)]^{1/2}$, computed from the formal standard errors supplied by the consortia, showed some significant variations around unity as functions of time, as may be seen in Figure 17.1.

Small standard errors were also found to have been overestimated (Figure 17.2). For this reason, a scaling of the standard errors had to be introduced. To take into account the variations in Figure 17.2, u_F and u_N were approximated by polynomials of degree four for formal errors below 10 mas, and by constants for the larger standard errors. This calibration obviously induced a magnitude effect, and the unit-weight errors were thereafter calibrated also against magnitude, where linear corrections were sufficient. This procedure was iterated once. A slight colour correction was also applied for stars

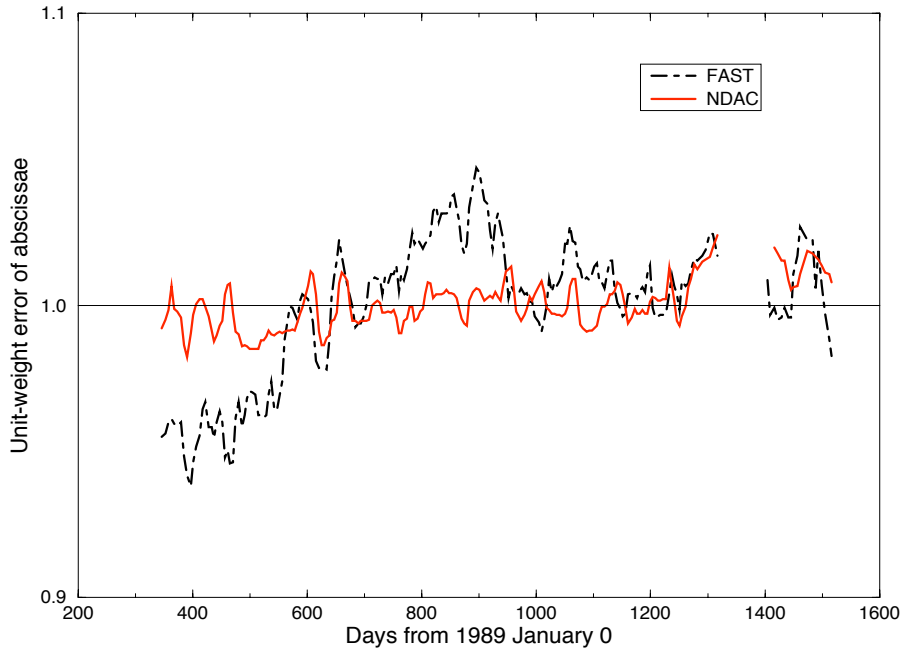

Figure 17.1. *Evolution of the unit-weight errors (u_{F} and u_{N}) of the abscissa residuals, based on the formal abscissa errors supplied by the reduction consortia. Each curve is a running median over 50 orbits ($\simeq$ 22 days).*

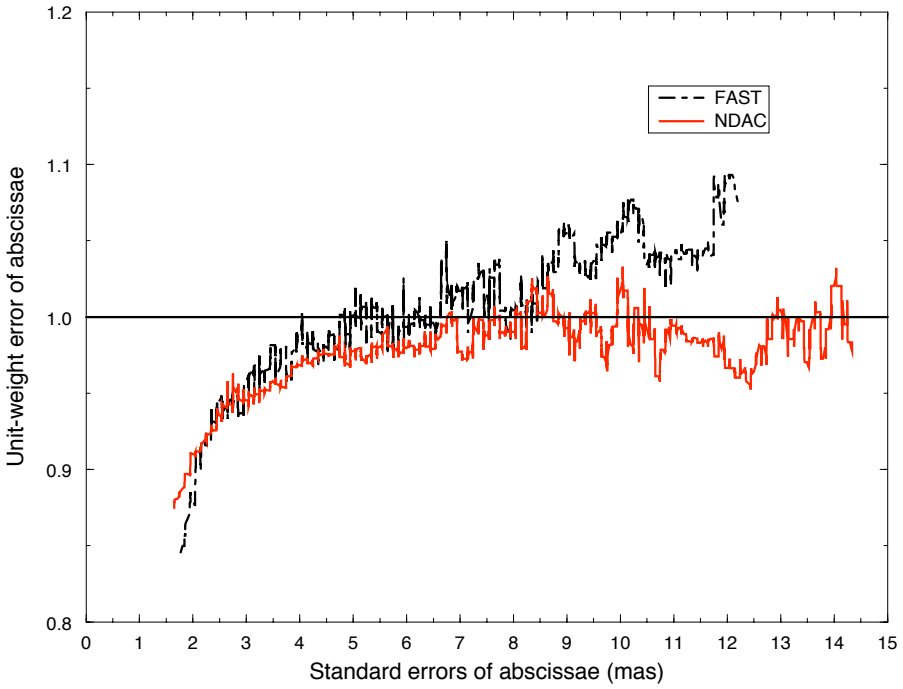

Figure 17.2. *Unit-weight errors (u_{F} and u_{N}) of abscissa residuals as functions of the formal standard errors.*

redder than $V-I=2$ mag, namely $1+0.035(V-I-2)$ for FAST and $1+0.016(V-I-2)$ for NDAC.

Having fitted the calibration functions $f_F(\sigma_F, Hp, V-I)$ and $f_N(\sigma_N, Hp, V-I)$ to the unit-weight errors, the standard errors could then be corrected through multiplication by these functions. These corrections were smaller than 30 per cent for FAST data (3 per cent on the average), and not more than 20 per cent for NDAC (1 per cent on the average). The corrected unit-weight errors became approximately constant (except for the very small formal errors), with no significant magnitude or colour effects. There was however still a significant variation with time, from one orbit to the next, for which an *ad hoc* correction factor was finally applied.

Astrometric solutions using only FAST or NDAC data were performed in order to verify the usefulness of the corrections which had been done. As may be seen in Figure 17.7, the standard errors on the FAST parallaxes decreased by about 4 per cent on the average. That this was not just a reduction of the formal errors is demonstrated by the fact that the number of negative parallaxes decreased by 2 per cent. For the NDAC data the overall changes brought by the scaling corrections were much smaller.

17.4. Correlation Between Abscissae

The correlation coefficient between FAST and NDAC abscissae varied essentially with magnitude, although there were also some variations with ecliptic longitude and orbit number. It would then have been logical to calibrate the correlation against magnitude. However, this would have created a practical problem for variable stars (where the correlation should then take into account the magnitude at each observation), since epoch photometry was not available at the time of the merging. For this reason it was decided to calibrate the correlation primarily against the abscissa standard errors (which are of course related to the epoch magnitude).

Let ρ be the statistical correlation between the residuals Δv_F and Δv_N, with standard deviations σ_F and σ_N. The variance of the abscissa difference $\Delta v_F - \Delta v_N$ is given by $\sigma_F^2 - 2\rho\sigma_F\sigma_N + \sigma_N^2$. Using the calibrated standard errors described in the previous section, the sample correlation coefficient was computed with the following formula:

$$\rho = \frac{\sigma_F^2 + \sigma_N^2 - \mathrm{Var}(\Delta v_F - \Delta v_N)}{2\sigma_F\sigma_N} \qquad [17.5]$$

This was done in bins of (σ_F, σ_N). Assuming that ρ depends only on the standard errors, the residuals in a given bin belonged to the same population and an unbiased estimate of ρ could be obtained, provided that the standard errors were also unbiased. Results are shown in Figure 17.3.

As can be expected from the rough model outlined in Section 17.1, the correlation increases with the standard errors and a relation $\rho \simeq 1 - \mathrm{const}/(\sigma_F\sigma_N)$ was expected. The correlation also decreases with the difference between the variances of the consortia abscissae, producing a rather sharp ridge along the diagonal $\sigma_N = k\sigma_F$, where $k \simeq 1.2$. It should be noted that the number of abscissa pairs on which the calculation of ρ was based drops quickly when going away from the diagonal, causing large statistical fluctuations in these areas of the diagram. The 'ridge' behaviour can also be understood in terms of the model mentioned above, as being due to increased modelling errors

on some abscissae in one or the other consortium. Based on these considerations, the following empirical form was chosen for the calibration of ρ as a function of the standard errors:

$$\rho_0(\sigma_F, \sigma_N) = 1 - \frac{a + \frac{1}{2}|b + k^2\sigma_F^2 - \sigma_N^2|}{\sigma_F\sigma_N} p_1\left(\frac{1}{\sigma_F\sigma_N}\right) \quad [17.6]$$

where $a \simeq 9.978$ mas^2, $b \simeq -0.160$ mas^2 and $k \simeq 1.168$ are constants, and p_1 is the polynomial:

$$\begin{aligned} p_1(x) = 1 &- 16.221x + 141.983x^2 - 663.074x^3 \\ &+ 1661.218x^4 - 2099.526x^5 + 1047.439x^6 \end{aligned} \quad [17.7]$$

The fitted function is shown in Figure 17.4.

In order to verify this calibration, the statistic

$$\delta v = \frac{\Delta v_F - \Delta v_N}{[\sigma_F^2 - 2\rho_0(\sigma_F, \sigma_N)\sigma_F\sigma_N + \sigma_N^2]^{1/2}} \quad [17.8]$$

was computed for each star and analysed as a function of astrometric or photometric data. The random variable δv should have zero expectation and unit variance if the correlation coefficient is correctly calibrated and if the standard errors σ_F and σ_N are representative of the true variations of the random errors on the abscissae. A magnitude effect was however found, due to the fact that, for bright stars, the abscissa standard errors were only weakly correlated with magnitude. Developing $\mathrm{Var}(\delta v) - 1$ to first order in ρ allowed the required correction to ρ_0 to be found. This was fit with a cubic polynomial in magnitude.

Finally, there was another variation of the correlation coefficient with time, the correlation being maximum around half-way through the mission and smaller at the beginning and the end of the mission. This effect is probably related to the precision of the proper motions. Using the same method as described above, an additional correction was determined as a cubic polynomial of time.

In summary, the calibration of the correlation coefficient was finally expressed as a function of the standard errors, magnitude and time:

$$\begin{aligned} \rho(\sigma_F, \sigma_N, Hp, t) = \rho_0(\sigma_F, \sigma_N) &- 0.1205 \\ &- 0.02770Hp + 0.010990Hp^2 - 0.0006509Hp^3 \\ &- 0.0038t - 0.0314t^2 + 0.00415t^3 \end{aligned} \quad [17.9]$$

where ρ_0 is given by Equation 17.6 and t is the time in years from J1991.25; the calibrated correlation was further constrained to the interval $0.2 \le \rho \le 0.99$.

17.5. The Least-Squares Solutions

Together with the standard error of the FAST and NDAC abscissae, the calibrated correlation coefficients provided the necessary information about the covariance matrix **V** of the observations. The covariances between abscissae from different great circles were neglected. Grouping the observations by pairs, corresponding to the FAST and

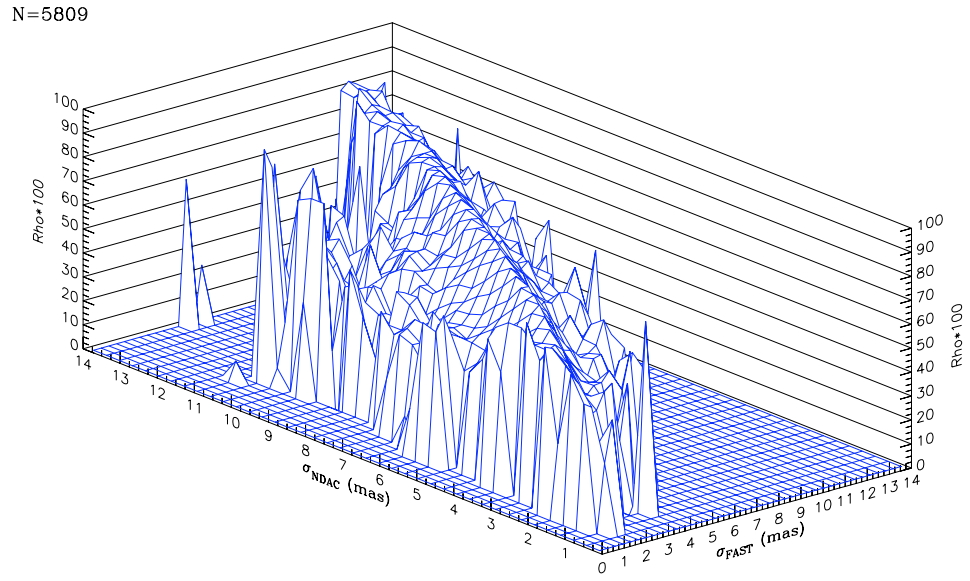

Figure 17.3. *Sample correlation coefficient ρ between the FAST and NDAC abscissae, calculated in bins of the standard errors* σ_F *and* σ_N.

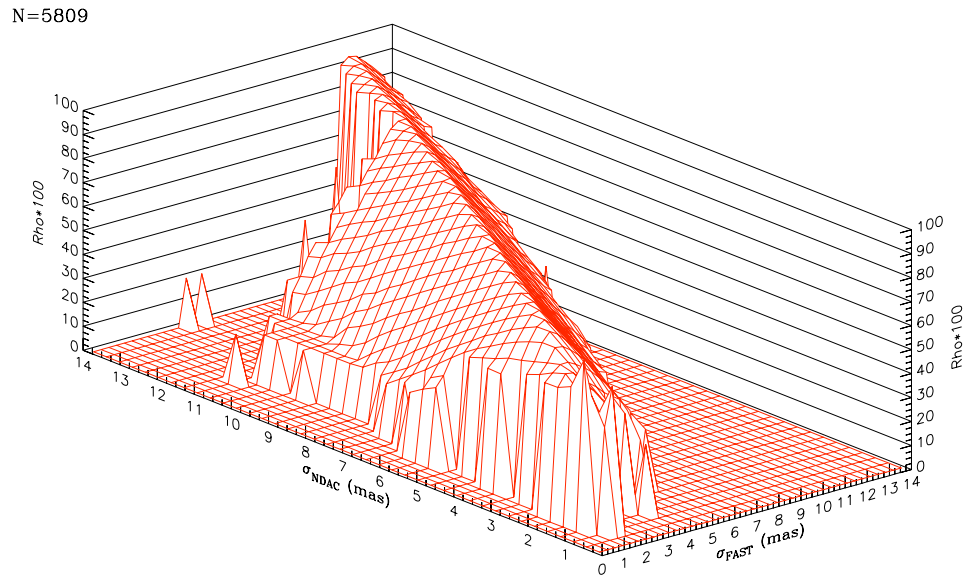

Figure 17.4. *Calibration function* ρ_0 *(Equation 17.6) fitted to the sample correlation as a function of the standard errors* σ_F *and* σ_N.

NDAC abscissae on each orbit, results in a block-diagonal structure of the covariance matrix:

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{V}_2 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad [17.10]$$

where the submatrix for the jth pair is given by:

$$\mathbf{V}_j = \begin{pmatrix} \sigma_{\mathrm{F}j}^2 & \rho_j \sigma_{\mathrm{F}j} \sigma_{\mathrm{N}j} \\ \rho_j \sigma_{\mathrm{F}j} \sigma_{\mathrm{N}j} & \sigma_{\mathrm{N}j}^2 \end{pmatrix} \qquad [17.11]$$

For orbits in which only one consortium provided an abscissa $\mathbf{V}_j$ reduces to the 1×1 matrix containing $\sigma_{\mathrm{F}j}^2$ or $\sigma_{\mathrm{N}j}^2$.

Given the partial derivatives of the abscissae with respect to the five astrometric parameters, $\partial v_j / \partial a_i$, the corrections Δa_i to the reference parameters a_i were found by minimising:

$$\chi^2 = (\Delta \mathbf{v} - \frac{\partial \mathbf{v}}{\partial \mathbf{a}'} \Delta \mathbf{a})' \mathbf{V}^{-1} (\Delta \mathbf{v} - \frac{\partial \mathbf{v}}{\partial \mathbf{a}'} \Delta \mathbf{a}) \qquad [17.12]$$

This differs from a standard weighted least-squares solution only in that the weight matrix $\mathbf{V}^{-1}$ is not diagonal, which requires a trivial modification of the normal equations.

The algorithm was also modified for robustness. Initially all observations were kept, but outliers were not unexpected, due for instance to veiling-glare effects. Robust methods using a non-Euclidean metric were tried, but abandoned as the variance of the estimated astrometric parameters increased too much for stars without outliers, which were in the majority. A conventional 3σ rejection was used instead: when the absolute value of a residual exceeded three times its formal error, the observation was excluded and a new solution computed. This process was iterated until the set of outliers remained constant. Pairs of observations were also rejected if their normalised difference (Equation 17.8) was greater than $3\sqrt{2}$.

For the single stars only, an average of 62 observations were used per star, with an outlier rate of 0.5 per cent. The median unit-weight error of the astrometric solutions was about 1.01, and this slight departure from unity explains why the median of the goodness-of-fit statistic F2 (Field H30) is about 0.2. The unit-weight error exhibited no significant variations with astrometric or photometric data.

17.6. Merged Solutions of Non-Single Stars

The standard five-parameter model did not always adequately represent the observations. More complex models were constructed (Volume 1, Section 2.3) and tested on all stars. Seven- and nine-parameter solutions (type 'G') were obtained the same way as in Equation 17.12, using the supplementary partial derivatives:

$$\frac{\partial v}{\partial a_{i+5}} = \tfrac{1}{2}(t^2 - 0.81) \frac{\partial v}{\partial a_i} \qquad [17.13]$$

$$\frac{\partial v}{\partial a_{i+7}} = \tfrac{1}{6}(t^2 - 1.69) \frac{\partial v}{\partial a_{i+3}} \qquad [17.14]$$

for $i = 1, 2$, where t is the time in years from J1991.25. The constants 0.81 yr^2 and 1.69 yr^2 were chosen to make the quadratic and cubic terms approximately orthogonal

to the position and proper motion terms. Orbital solutions (type 'O') with up to twelve parameters were similarly obtained, using a special routine to calculate the partial derivatives of the tangential coordinates ξ and η (Volume 1, Section 2.3.4) with respect to the seven orbital elements $\mathbf{o}$, from which:

$$\frac{\partial v}{\partial \mathbf{o}'} = \frac{\partial v}{\partial a_1}\frac{\partial \xi}{\partial \mathbf{o}'} + \frac{\partial v}{\partial a_2}\frac{\partial \eta}{\partial \mathbf{o}'} \qquad [17.15]$$

Several trials were made before defining the criteria for selection of a model for a given star. The adopted criteria are described in Volume 1, Section 2.3.1.

Although good results were obtained for almost all stars, there existed some stars where none of the above-mentioned models was adequate, thus leading to a high outlier rejection rate or a bad goodness-of-fit. This could be due for instance to unrecognised duplicity or orbital motion. Following a suggestion by R. Wielen, a stochastic model was assumed for these stars. Superposed on the uniform motion of the centre of mass, unmodelled photocentric displacements were assumed to behave in a stochastic manner, with a standard deviation given by the 'cosmic error' ϵ. The cosmic error was determined by adding a variance ϵ^2 to each of the individual abscissa variances, the constraint being that the unit-weight variance $u^2 = \chi^2/(n-5)$ of the five-parameter astrometric solution must be equal to unity. A normal rate of possible outliers (≤ 2 per star) was however still allowed. The cosmic error was added quadratically to the abscissa standard errors, and the modified standard errors $(\sigma_F^2 + \epsilon^2)^{1/2}$, $(\sigma_N^2 + \epsilon^2)^{1/2}$ were used to compute the correlation coefficient between the FAST and NDAC abscissae according to Equation 17.9. The calculated correlation coefficient therefore increased in the presence of a cosmic error.

For each star, the adopted procedure was the following iteration:

1. a normal astrometric solution, as described above, was performed. If $u^2 \neq 1$, the maximum allowed number of outliers (rejected abscissae) was set to 2;
2. if no cosmic error was indicated ($u^2 \leq 1$), the maximum number of outliers was decreased to 1, then 0. Evidently, the standard five-parameter model was adequate for the majority of the stars, in which case the cosmic error was set to $\epsilon = 0$;
3. otherwise, ϵ was computed from the non-rejected observations. Assuming Gaussian residuals, $(n-5)u^2$ follows a χ^2_{n-5} distribution, where $n = n_F + n_N$ is the total number of observations, and the standard error on ϵ is then approximately given by:

$$\sigma_\epsilon = \left[\frac{2}{n-5}\right]^{1/2} \left[2\epsilon\frac{\partial(u^2)}{\partial(\epsilon^2)}\right]^{-1} \qquad [17.16]$$

   In order to avoid unrealistically small standard errors on some stars, a lower limit of $\epsilon/\sqrt{2(n-5)}$ was introduced on σ_ϵ;
4. the covariance matrix $\mathbf{V}$ was computed for a new iteration.

This procedure was applied to the whole catalogue. A sample of about 94 000 single stars was first defined, i.e. stars not recognised as double or multiple, with a goodness-of-fit statistic F2 < 3, and with at most two rejected abscissae. From these stars, it was empirically found that the criterion $\epsilon > 5\sigma_\epsilon$ corresponded to a significance level similar

to the usual 3σ two-sided test on a Gaussian distribution. This criterion was therefore used to decide whether a stochastic solution was significant.

One of the possible causes for the presence of a cosmic error is unrecognised duplicity. Known double stars should therefore also produce significant cosmic error if given a stochastic solution instead of the proper solution with two or more components. In order to test this hypothesis, stochastic solutions were computed for all Hipparcos stars resolved as double and thus contained in Part C of the Double and Multiple Systems Annex. The cosmic errors were plotted against the effective magnitude difference D (Equation 13.40) of the doubles. This quantity, equal to the real magnitude difference ΔHp when the separation exceeds 0.32 arcsec but increasing for small separations, is a measure of the difficulty by which the secondary may be detected in the Hipparcos detector signal. A strong correlation was found between the cosmic error and the effective magnitude difference (Figure 17.5).

Other stars which received a stochastic solution may be astrometric binaries with a period of less than a few years, and the cosmic error could then be taken as an order-of-magnitude estimate of the semi-major axis of the orbit of the photocentre (a_0). Significant stochastic solutions were obtained for some of the objects finally retained as orbital astrometric binaries in the Hipparcos Catalogue. In Figure 17.6 the cosmic errors for these objects are compared with the semi-major axes as given in Field DO4 of the Double and Multiple Systems Annex. As expected, there is a rough proportionality between the two quantities, with $a_0 \simeq 2.4\epsilon$ at least for small a_0.

A grossly erroneous starting position for the astrometric adjustment usually resulted in a number of grid-step errors in the abscissa residuals, which were then more or less uniformly distributed between $-s/2$ and $+s/2$, where $s = 1.2074$ arcsec is the grid step. These cases show up among the stochastic solutions with a cosmic error of the order of $s/\sqrt{12} \simeq 300$ mas (Figure 17.5 contains a number of such solutions). A similar effect tends to occur for non-detections (very weak signal, e.g. due to erroneous pointing of the instantaneous field of view). After elimination of all systems solved by other models, a number of stars with large cosmic errors ($\epsilon > 100$ mas) still remained and had to be examined more closely. Some of them were indeed found to be affected by grid-step errors, and good five-parameter solutions could be obtained by changing the reference positions and the abscissae by multiples of the grid step. Remaining cases with a cosmic error exceeding 100 mas were however rejected as being most probably invalid astrometric solutions.

17.7. Comparison with a Weighted Mean

An alternative to the adopted merging procedure described above would be to use, for each astrometric parameter a of a given star, a weighted mean of the FAST and NDAC parameters:

$$a_{\mathrm{W}} = w a_{\mathrm{F}} + (1 - w) a_{\mathrm{N}} \qquad [17.17]$$

This method (with a constant $w = 0.5$) was used for the construction of the provisional catalogues H18 and H30, but a slightly better result would be obtained by optimising the weight w for each star, or even for each parameter. It is of interest to compare the precision of this simple method to the more elaborate merging adopted for the final catalogue.

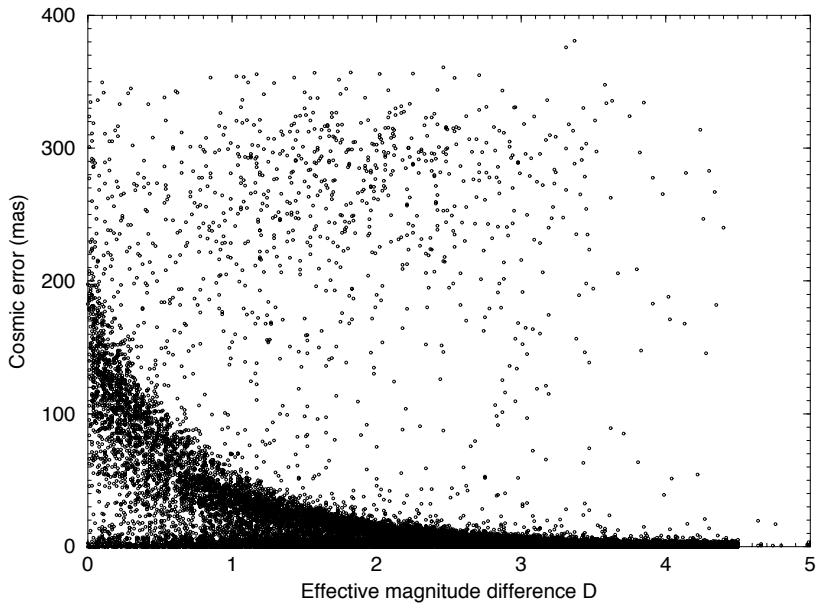

Figure 17.5. *Cosmic error ϵ (mas) versus the effective magnitude difference D (Equation 13.40) for systems retained as resolved double stars in the final catalogue (Part C of the Double and Multiple Systems Annex). Stochastic solutions with cosmic errors above some 200 mas are generally due to grid-step errors, which were not eliminated in this sample.*

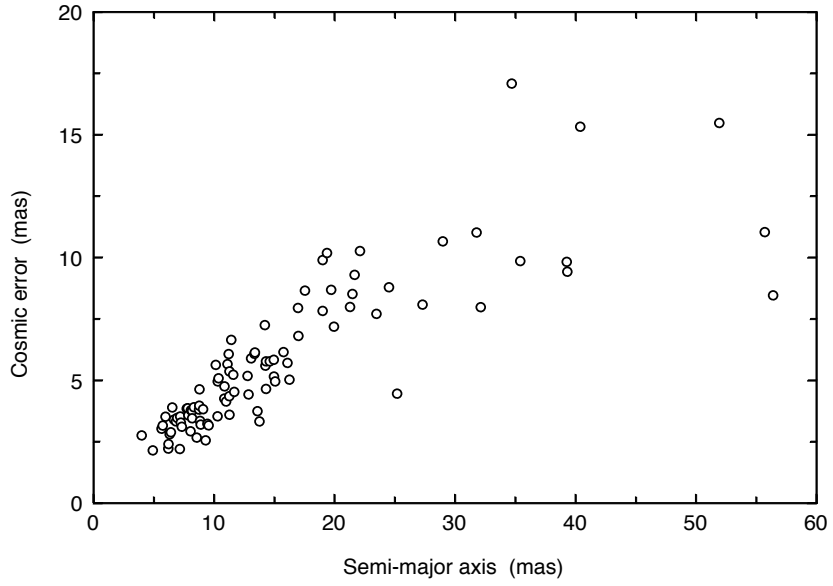

Figure 17.6. *Cosmic error (mas) versus the semi-major axis of the barycentric orbit of the photocentre. The 95 systems shown in this diagram were retained as orbital astrometric binaries in the final catalogue (Part O of the Double and Multiple Systems Annex), but stochastic solutions were also computed for this comparison.*

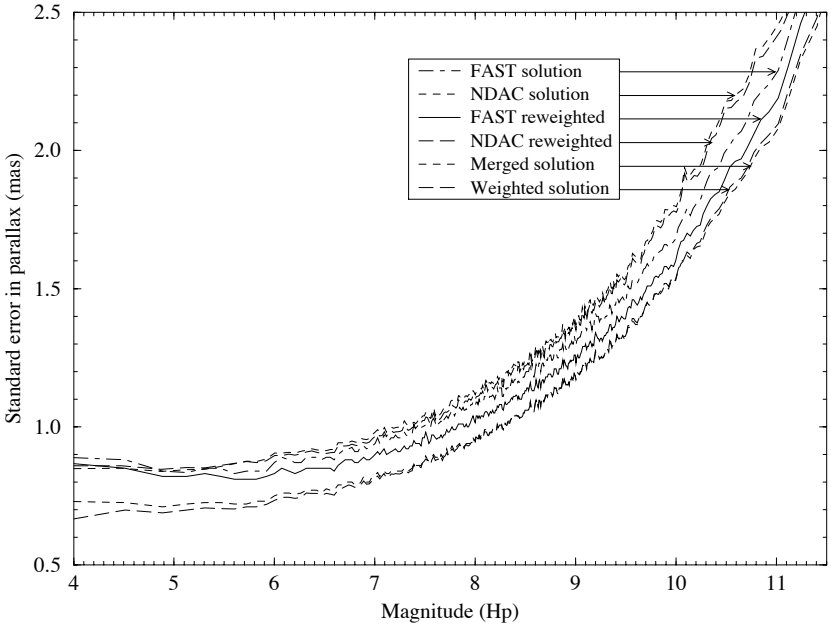

Figure 17.7. *Median precision of parallaxes of single stars versus magnitude Hp for various astrometric solutions: (a) initial FAST solution; (b) initial NDAC solution; (c) FAST solution with re-weighted abscissa standard errors; (d) NDAC solution with re-weighted abscissa standard errors; (e) merged solution adopted for the final Hipparcos Catalogue; (f) a weighted mean of the FAST and NDAC parallaxes.*

The value w which minimises the variance σ_W^2 of a_W is:

$$w = \frac{1 - \rho q}{1 - 2\rho q + q^2} \qquad [17.18]$$

where $q = \sigma_F/\sigma_N$ is the known ratio of the standard errors and ρ the (initially) unknown correlation coefficient between a_F and a_N. The variance of the weighted mean parameter is given by:

$$\sigma_W^2 = \frac{1 - \rho^2}{1 - 2\rho q + q^2}\,\sigma_F^2 \qquad [17.19]$$

The correlation coefficient may be estimated the same way as in Equation 17.8 under the assumption that ρ and q are constant for a given magnitude; the result is:

$$\rho \simeq \left\langle \frac{1 + q^2}{2q} \right\rangle [1 - \mathrm{Var}(\overline{\Delta a})] \qquad [17.20]$$

where $\overline{\Delta a} = (a_F - a_N)/(\sigma_F^2 + \sigma_N^2)^{1/2}$ are the normalised differences introduced in Equation 16.23. This procedure allowed ρ to be computed in bins of magnitude. Each of the two factors of the product in Equation 17.20 was then calibrated as a polynomial in the Hp magnitude. Introducing this $\rho(Hp)$ into Equations 17.17–17.19 gives the weighted astrometric parameter and associated standard error for a star of magnitude Hp.

Applying the above method to the parallaxes of single stars gives the median value of σ_W as drawn in Figure 17.7. In terms of the astrometric standard errors, the adopted merging procedure appears roughly equivalent to the present method using a simple weighted mean of the astrometric parameters. However, the more elaborate merging of the abscissa data allows the covariances of the merged astrometric parameters also to be computed in a rigorous manner; this is not possible in the simpler method.

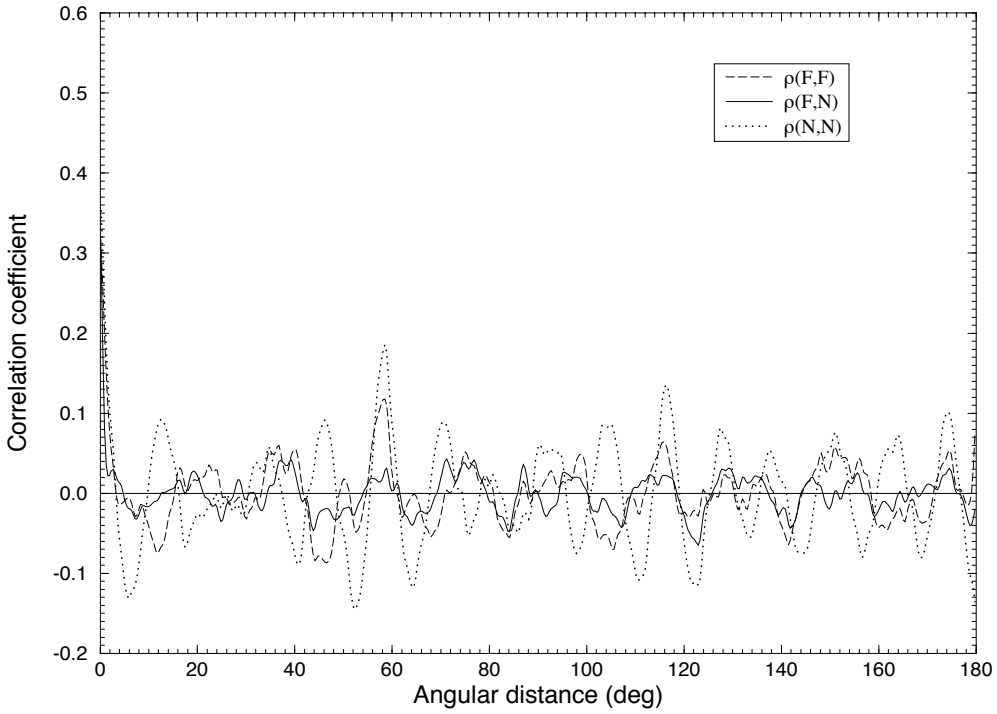

Figure 17.8. *The auto and cross-correlation functions $\rho_{\rm FF}(\theta)$, $\rho_{\rm FN}(\theta)$ and $\rho_{\rm NN}(\theta)$ of the abscissa residuals in orbit number 1001, plotted against the angular separation of the stars (θ). A bin size of $0\overset{\circ}{.}35$ in θ was used to compute the correlations, which were then smoothed with a Gaussian kernel of standard deviation $\simeq 0\overset{\circ}{.}43$.*

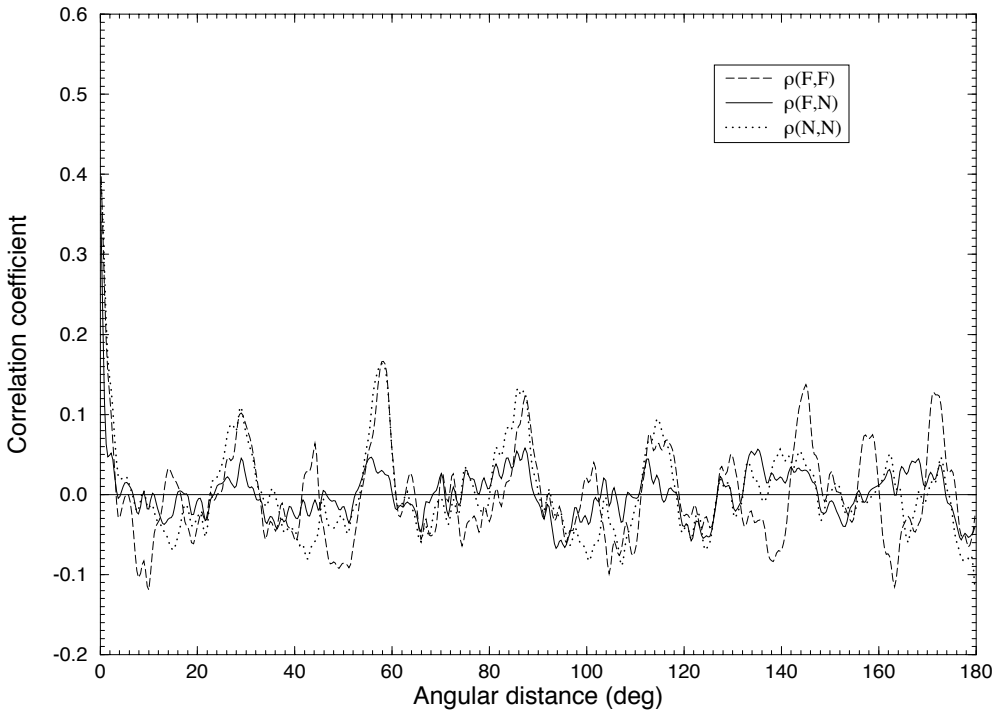

Figure 17.9. *Same as Figure 17.8 but for orbit 2000, illustrating the typical variation of the correlation functions from one orbit to another.*

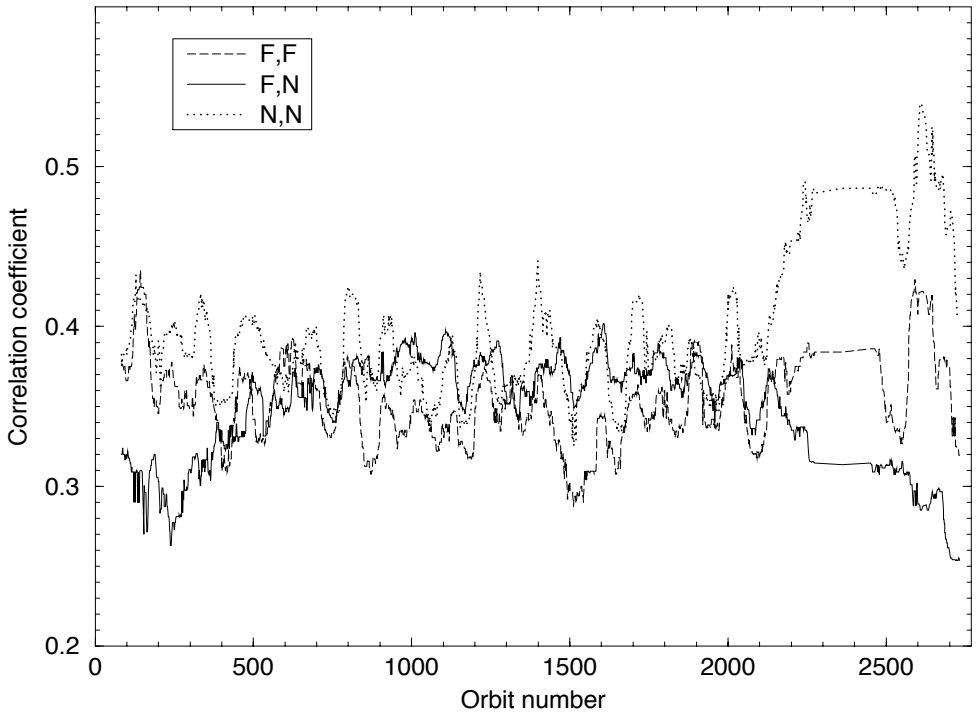

Figure 17.10. *The first bin of the auto and cross-correlation functions $\rho_{\rm FF}(\theta)$, $\rho_{\rm FN}(\theta)$ and $\rho_{\rm NN}(\theta)$, i.e. for separations $\theta < 0\overset{\circ}{.}35$, plotted against the orbit number. Each curve is a running median over 50 orbits ($\simeq$ 22 days).*

17.8. Correlations Between Different Stars on the Same Great Circle

Already in the early stages of the preparations of the Hipparcos data reductions it was recognised that the observation mode, and especially the great-circle reductions, would generate cyclic correlation patterns among the stars observed on the same great circle (Høyer *et al.* 1981). More detailed predictions were included in the pre-launch documentation (Perryman *et al.* 1989 Volume III, Chapter 23). A significant positive correlation (of a few tenths) was expected between the abscissae v_i and v_j of stars for which $|v_i - v_j|$ was less than a few degrees or close to a small multiple of the basic angle, modulo 2π.

The abscissa residuals from the merging allowed the sample correlation coefficients to be computed among pairs of the normalised residuals on the same great circle ($\Delta v_{{\rm F}i}/\sigma_{{\rm F}i}$, $\Delta v_{{\rm F}j}/\sigma_{{\rm F}j}$, $\Delta v_{{\rm N}i}/\sigma_{{\rm N}i}$ and $\Delta v_{{\rm N}j}/\sigma_{{\rm N}j}$ for $i \neq j$). This was done as functions of the angular separation θ between the stars (see Equation 16.25). Note that the actual angular separation of the stars was used, rather than the abscissa difference. Two autocorrelation functions [$\rho_{\rm FF}(\theta)$ and $\rho_{\rm NN}(\theta)$] and one cross-correlation function [$\rho_{\rm FN}(\theta)$] were calculated, using a bin size of $180^\circ/512 \simeq 0\overset{\circ}{.}35$. The results for two fairly typical great circles are shown in Figures 17.8–17.9.

Although most correlation functions show the expected peaks at multiples of the basic angle (0°, 58°, 116°, 174°, 128°, ...), the amplitude of the peaks varies considerably between different great circles, and also between the consortia results on the same

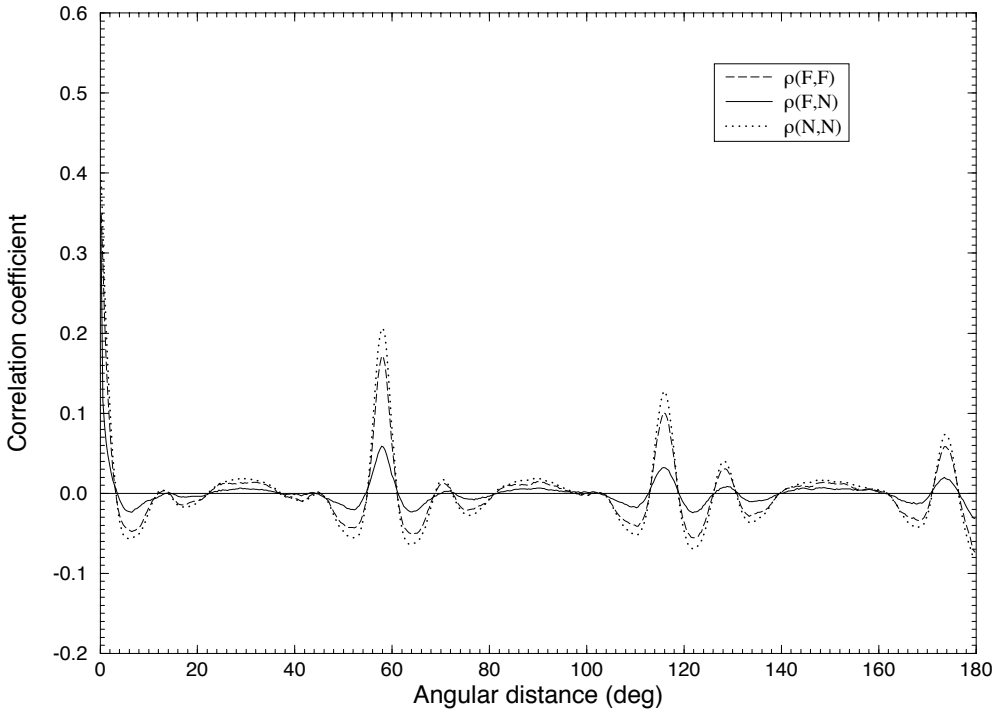

Figure 17.11. *The average auto and cross-correlation functions $\rho_{\rm FF}(\theta)$, $\rho_{\rm FN}(\theta)$ and $\rho_{\rm NN}(\theta)$ obtained by considering all 2341 available great circles.*

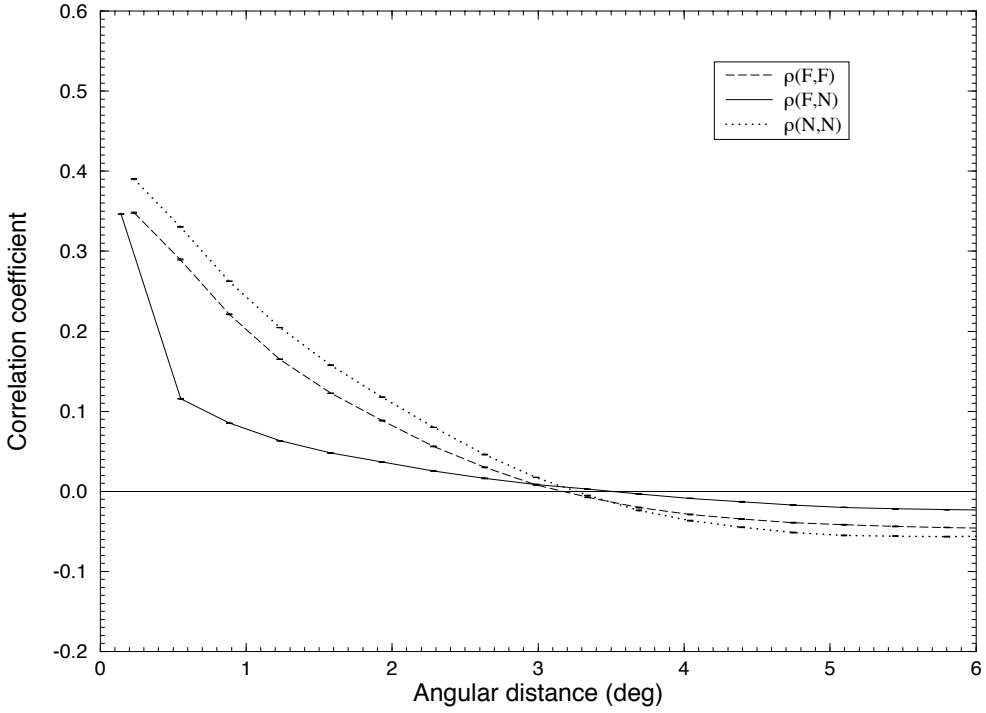

Figure 17.12. *Same as Figure 17.11 but on an enlarged scale showing the correlations for small angular separations.*

great circle. The variation of the correlations for small separations (the first bin, i.e. $\theta < 0\overset{\circ}{.}35$) is shown in Figure 17.10. The mean correlation functions for the whole mission are shown in Figures 17.11–17.12. The auto-correlations for the FAST data are slightly smaller than the NDAC ones, and the cross-correlation is much smaller than the autocorrelations.

It is interesting to see the influence of the abscissa correlations on the final astrometric parameters. As seen in Figures 16.36–16.37, the small-scale correlations ($\theta \lesssim 2°$) remain whereas the other correlation peaks at multiples of 58° are very strongly damped. The reason for this is that any given pair, separated by 58° on the sky, was very rarely observed on the same reference great circle. The peak at 174° is relatively less damped because stars at diametrically opposite points on the sphere were more likely to be observed on the same great circle.

The auto- and cross-correlation functions can be used, together with the Hipparcos Intermediate Astrometric Data (Volume 1, Section 2.8), to estimate the full covariance matrix of the abscissae for any group of stars, such as in a cluster. This may be useful for evaluating the correlations between the astrometric parameters of the different stars in a more rigorous manner than using the average correlation discussed in Section 16.6.

17.9. Conclusions

By using the FAST and NDAC star abscissae as a basis for the merging, rather than the astrometric solutions of each consortium, it was possible to find an optimum solution for all five (or more) astrometric parameters, including estimates of the standard errors and correlations of the merged astrometric parameters. The method required however very careful calibration of the standard errors of the abscissae and the correlation of abscissa errors between the consortia, for which special techniques were developed. The overall correlation was about 0.7, allowing a significant improvement in precision by the merging, as can be seen in Figure 17.7. The merging of the abscissae also made it relatively simple to choose between the several different models of star motion, from the standard five-parameter model to the twelve-parameter model for orbital binaries.

The final merging resulted in a catalogue called H37C, in which the positions and proper motions were still expressed in the provisional *ad hoc* reference frame defined by earlier mergings. The main part of the Hipparcos Catalogue was created from H37C by applying the rigid-body rotation derived from comparing H37C with the extragalactic reference frame, as described in Chapter 18. Detailed statistics of the Hipparcos Catalogue are given in Volume 1, Part 3.

An important by-product of the merging process was the generation of the complete set of FAST and NDAC abscissa residuals, expressed on the same reference frame as the Hipparcos Catalogue and including the calibrated standard errors and correlations; for future investigations this data set is made available on ASCII CD-ROM, in the form of the Hipparcos Intermediate Astrometric Data described in Volume 1, Section 2.8.

F. Arenou