Reconstitution of the celestial sphere — an alternative scheme

L. Lindegren 78-10-09

INTRODUCTION

The processing of the observations made during one scan (or group of scans) results in a set of relative positions of stars along the reference great circle, and the corresponding covariance matrix. The problem of reconstituting the celestial sphere consists in combining these data for all scans in order to obtain the astrometric parameters of the stars involved. This is accomplished in the second and third steps of the three-step procedure described e.g. in Chapter 8 of the report of the theoretical study.

I think, however, that this scheme is not sufficiently rigorous for a more detailed analytical and numerical investigation of the problem. Hence I propose here an alternative scheme which differs from the old one in the following respects:

1. Since instrumental parameters as well as the motion of the satellite are in principle solved for in the first step (i.e. the processing of one or a few scans), the only unknowns remaining are the astrometric parameters of the stars. Therefore it is very unsatisfactory that, in the old scheme, several additional unknowns were introduced, namely the zero-points of the scans, bj. These are in fact completely unnecessary and do not appear in the new scheme.

2. The abscissas of the stars along the reference great circle are reckoned from the abscissa of a particular star (let us call it the origin star), chosen for each scan. The solution for the zero-points of the scans in the old scheme (step 2) corresponds in the new scheme to solving the astrometric parameters of all origin stars; thus the number of origin stars should be kept as low as possible. The minimum number of origin stars can be estimated as follows. For a sun/spin-axis angle of 30° all scans will intersect the latitude circle at +60° twice. Half the circumference of this small circle is 90°. Since the width of the field of view is 0.9°, about 100 stars distributed along the +60° circle should in principle suffice. In practice, some 150 stars may be necessary. The number of unknowns to be solved in step 2 of the new scheme is then 750, which is considerably less than the number of scan zero-points to be solved for in the old scheme. So the new scheme is also computationally favourable in this respect.

3. The basic indetermination of the sphere, i.e. the six arbitrary constants describing the orientation and rotation of the sphere, is removed in the new scheme by the addition of six rigorous linear equations for the astrometric parameters. The proposed solution means that the ignorance is evenly spread over the sphere, producing a system of positions and proper motions which coincides in a least-squares sense with that of a given subset of stars (e.g. FK5).

4. An attempt was made to take into account the correlations between measurements of abscissas within the same scan (but neglecting as before the correlations between different scans). Unfortunately, this completely ruins the nice structure of the normal equations which makes a decomposition possible. In addition, strict decorrelation requires the full inverses of the covariance matrices from step 1, which are not obtained in the filtering algorithm. Part of
the correlations in step 1 is due to the use of an origin star (rather than
imposing, e.g., the the sum of abscissa corrections is zero); it is shown
how this can possibly be taken into account in an iterative way. An important
subject for a small-scale numerical study of the reconstitution problem is
possible effects of neglecting these correlations.

THE OBSERVATION EQUATION

The notations to be used here are essentially the same as the previously used,
but the definitions are repeated for convenience:

\( i = 1, 2, \ldots, I \) \hspace{1cm} \text{enumeration of stars}

\( j = 1, 2, \ldots, J \) \hspace{1cm} \text{enumeration of scans}

\( i_0(j) \) \hspace{1cm} \text{origin star (OS) number in scan } j

\( i_l(j), \ l = 1, 2, \ldots, L(j) \) \hspace{1cm} \text{the } l\text{th star in scan } j \text{ following OS}

\( L(j) + 1 \) \hspace{1cm} \text{number of stars in scan } j \text{ (including OS)}

\( a_i = (\Delta \lambda \cos \beta, \Delta \beta, \Delta \mu_\lambda \cos \beta, \Delta \mu_\beta, \Delta \omega)^* \) \hspace{1cm} \text{astrometric parameters for star } i

\( P_{j,i} = (-\sin \eta, -\cos \eta, \ldots) \) \hspace{1cm} \text{a } (1,5) \text{ row vector for star } i \text{ in scan } j

\( d_{j,i} \) \hspace{1cm} \text{abscissa correction for star } i \text{ in scan } j

\( e_{j,i} \) \hspace{1cm} \text{error in measured quantity } d_{j,i}

Note that by definition we have \( d_{j,i_0(j)} = e_{j,i_0(j)} = 0 \).

The observation equations are:

\[
P_{j,i_l(j)} a_i = P_{j,i_0(j)} a_i + d_{j,i_l(j)} + e_{j,i_l(j)}
\]

\( l = 1, 2, \ldots, L(j); \ j = 1, 2, \ldots, J. \) \hspace{1cm} (1)

THE NORMAL EQUATIONS

For writing the normal equations it is convenient to introduce the operators:

\( i^s \) \hspace{1cm} \text{sum over scans } (j) \text{ such that } i_0(j) = i

\( i^s \) \hspace{1cm} \text{sum over scans } (j) \text{ such that } i_l(j) = i \text{ for some } l > 0

\( i^{s,s'} \) \hspace{1cm} \text{sum over scans } (j) \text{ such that } i_0(j) = i \text{ and } i_l(j) = i' \text{ for some } l > 0

\( i^{s,s'} \) \hspace{1cm} \text{sum over scans } (j) \text{ such that } i_l(j) = i \text{ and } i_{l'}(j) = i' \text{ for some } l,l' > 0

The normal equation for star no. \( i \) then becomes:
\[
\begin{align*}
\left[ i_2 [L(j) P_j, i_d, i_j] + s_i' [P_j, i_d, i_j] \right] a_{ij} - \sum_i \left[ i_s' \cup i_s' [P_j, i_d, i_j] \right] a_{ij} &= \\
= S_i' [P_j, i_d, i_j] - s_i' [P_j, i_d, i_j] \sum_i \left[ S_i' \cup S_i' [P_j, i_d, i_j] \right] a_{ij},
\end{align*}
\]

The first term of the left member is the diagonal term \( i, i \); the second term contains the off-diagonal elements \( (i, i') \), \( i \neq i' \). The element \( (i, i') \) is non-zero if and only if either \( i \) or \( i' \) belongs to the set of origin stars.
Thus, if the stars are numbered in such a way that \( i = 1, 2, \ldots, I \) is an OS in some scan, while \( i = I + 1, I + 2, \ldots, I \) is never an OS, the system of normal equations will have the following structure:

\[
5I_0 \left\{ \begin{array}{ccc}
A & B & E \\
B^* & D & F \\
\end{array} \right\} \times \left\{ \begin{array}{c}
G \\
H \\
\end{array} \right\} = \left\{ \begin{array}{c}
E \\
F \\
\end{array} \right\}
\]

in which \( D \) is block-diagonal, whereas the other matrices are fairly dense. As in the old scheme, the solution is carried out in two steps:

\[
\begin{align*}
\text{step 2:} & \quad E = (A - B D^{-1} B^*)^{-1} (G - B D^{-1} F)
\end{align*}
\]

\[
\begin{align*}
\text{step 3:} & \quad F = D^{-1} (H - B E)
\end{align*}
\]

which involves the solution of a system of order \( 5I_0 \) (step 2) and one system of order 5 for each star (step 3).

REMOVAL OF THE BASIC INDETERMINATION

If the coordinate system is rotated a small angle \( \theta_x \) around the \( x \)-axis, \( \theta_y \) around the \( y \)-axis, and \( \theta_z \) around the \( z \)-axis, the coordinates \( (\lambda, \beta) \) of a star will change by \( (\delta \lambda, \delta \beta) \), where:

\[
\begin{align*}
\delta \lambda & = \begin{bmatrix}
\sin \beta \cos \lambda & \sin \beta \sin \lambda & -\cos \beta \\
-\sin \lambda & \cos \lambda & 0
\end{bmatrix} \times \begin{bmatrix}
\theta_x \\
\theta_y \\
\theta_z
\end{bmatrix}
\end{align*}
\]

From a set of observed differences \( (\delta \lambda, \delta \beta) \) we obtain a least-squares solution for \( (\theta_x, \theta_y, \theta_z) \) by forming observation equations according to (6). The three normal equations decorrelate if the stars are evenly distributed over the sky, resulting in the solution:
\( x: \Sigma (\sin^2 \beta \cos^2 \lambda + \sin^2 \lambda) = \Sigma (\sin \beta \cos \lambda)(\delta \lambda \cos \beta) - \Sigma (\sin \lambda)(\delta \beta) \)

\( y: \Sigma (\sin^2 \beta \sin^2 \lambda + \cos^2 \lambda) = \Sigma (\sin \beta \sin \lambda)(\delta \lambda \cos \beta) + \Sigma (\cos \lambda)(\delta \beta) \)

\( z: \Sigma (\cos^2 \beta) = \Sigma (-\cos \beta)(\delta \lambda \cos \beta) \)

A corresponding set of equations gives the rates of rotation, \( \dot{\theta}_x, \dot{\theta}_y, \dot{\theta}_z \), from the observed proper motion differences \((\delta \mu, \cos \beta, \delta \mu_\beta)\). To obtain a system of positions and proper motions which in a least-squares sense coincides with that of a given subset of stars, FK, we add to the normal equations the following six equations and corresponding Lagrangean multipliers:

\[
\begin{align*}
\Sigma_{i \in FK} (\sin \beta \cos \lambda a_{i1} - \sin \lambda a_{i2}) &= 0 \\
\Sigma_{i \in FK} (\sin \beta \sin \lambda a_{i1} + \cos \lambda a_{i2}) &= 0 \\
\Sigma_{i \in FK} (\cos \beta a_{i1}) &= 0 \\
\Sigma_{i \in FK} (\sin \beta \cos \lambda a_{i3} - \sin \lambda a_{i4}) &= 0 \\
\Sigma_{i \in FK} (\sin \beta \sin \lambda a_{i3} + \cos \lambda a_{i4}) &= 0 \\
\Sigma_{i \in FK} (\cos \beta a_{i3}) &= 0
\end{align*}
\]

in which \( a_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5})^* \). The structure of (8) is retained if (8) is placed on top of the normal equations; the dimension of \( \Lambda \) increases to \( 5I_0 + 6 \).

**DECORRELATION OF OBSERVATION EQUATIONS**

The observation equations (1) are however not uncorrelated and of equal weight, which is implicitly assumed in forming the normal equations. (Neglecting correlations leads to a least-squares solution which is less accurate than if correlations are allowed for, while the estimated variances are too optimistic.) The correlations within scan \( J \) are however known and contained in the covariance matrix \( C_J \) with elements \( C_{ij} = E(e_{ij} e_{ij}') \). A linear transformation by a matrix \( H_J \):

\[
\{d'_{j,i}\} = H_J \{d_{j,i}\}
\]

results in a set of "observed" quantities \( d'_{j,i} \) with covariance

\[
C'_{j} = H_J C_J H_J^* \]

which becomes the identity matrix of dimension \( L(J) \) if \( H_J \) is a square-root of \( C_J^{-1} \).
\[ H_j^* H_j = C_j^{-1} \]  

(11)

Therefore, we obtain formally an equivalent set of uncorrelated observation equations of equal weight (all variances of residuals = 1) by multiplying (1) by \( H_j \), or explicitly:

\[ L(j) \sum_{l=1}^{L} (H_j^*)_{i,l}(j) [P_{j,l}^* P_{j,l}(j)] a_{i,l}(j) - P_{j,l}^* a_{i,l}(j) = \]

\[ = \sum_{l=1}^{L} (H_j^*)_{i,l}(j) d_{j,l} a_{i,l}(j) \]  

(12)

The elements of \( H_j \) are specified by the star numbers rather than index \( l \).

Forming the normal equations, it is found that the elements of \( H_j \) recombine to form the elements of \( C_j^{-1} = Q_j \). For brevity, introduce:

\[ q_{j,i,l}^* = (Q_j^*)_{i,l} \]

(13)

\[ q_{j,i} = \sum_{l=1}^{L} q_{j,i,l}(j) \]

The normal equation for star \( i \) becomes:

\[ \left[ i^* S_q j, P_{j,i}^*, P_{j,i} \right] + S_i^* q_{j,i,l}^* P_{j,i}^*, P_{j,i} \]  

\[ + \sum_{i',i} \left[ i^* S_q j, P_{j,i}^*, P_{j,i} \right] + S_i^* q_{j,i,l}^* P_{j,i}^*, P_{j,i} \]  

\[ = S_i^* \left[ P_{j,i}^*, q_{j,i,l}, d_{j,i} \right] - i^* S_q j, P_{j,i}^*, q_{j,i,l}, d_{j,i} \]  

(14)

Apart from the obstacle that \( C_j^{-1} \) must be computed for each scan, there are now also non-zero off-diagonal elements \( (i, i') \) as soon as stars \( i \) and \( i' \) are both observed in the same scan. Apparently, it is impossible to reduce the system to a banded-bordered structure suitable for computation, because the coupling of pairs \( (i, i') \) does not form a reasonably simple chain, but embraces the whole sphere in a few links. The only way to take the correlations rigorously into account seems to be by iterations: A preliminary solution according to (3) will probably be close to the final solution, and it can be hoped that a single iteration is sufficient.

Part of the correlations within a scan is due to the very fact that the abscissas \( d_{j,i} \) are all referred to the abscissa of the same star \( i_0(j) \). Consider a very simplified model for this: Suppose that the abscissas of all \( L(j) + 1 \) stars are measured independently and relative to an external point of reference, and that the resulting errors \( r_{j,l} i_0(j), l = 0, 1, \ldots, L(j) \) are uncorrelated and have the same variance \( \sigma^2 \). From these we obtain the error of \( d_{j,i} \):

\[ \sigma_{j,i}^2 = r_{j,l}^2 - r_{j,i_0}(j) \quad l = 1, 2, \ldots, L(j) \]  

(15)
Obviously we have the error covariance matrix of dimension \((L(j), L(j))\):

\[
C_j = \sigma^2 \begin{bmatrix}
2 & 1 & 1 & \ldots & 1 \\
1 & 2 & 1 & \ldots & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1 & 1 & \ldots & 2
\end{bmatrix}
\]

with inverse:

\[
Q_j = (L(j)+1)^{-1} \sigma^{-2}
\begin{bmatrix}
L(j) & -1 & -1 & \ldots & -1 \\
-1 & L(j) & -1 & \ldots & -1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-1 & -1 & -1 & \ldots & L(j)
\end{bmatrix}
\]

from which:

\[
(L(j)+1) \sigma^2 q_{jj'} = \begin{cases} 
L(j) & \text{for } i' = i \\
-1 & \text{for } i' \neq i
\end{cases}
\]

\[
(L(j)+1) \sigma^2 q_{ji} = 1
\]

\[
(L(j)+1) \sigma^2 q_{ji} = L(j)
\]

Inserting this into (14), assuming that \(L(j) = L\) for all scans, and multiplying with \((L+1)\sigma^2\), we obtain the normal equation:

\[
L \left( i S^2 \left[ P^*_j, i \right. \left. P^*_j, i' \right] \right) a_i - \sum_{i' \neq i} \left( i S^2 \left[ i', i \right. \left. i', i' \right] \left[ P^*_j, i \right. \left. P^*_j, i' \right] \right) a_{i'} = \]

\[
= i S^2 \left[ P^*_j, i \right. \left. L d_j, i - \sum_{i' \neq i} d_j, i' \right] \}
\]

Note that this equation, in contrast to (2), makes no difference between origin stars and non-origin stars.

It is seen that (19) is obtained from (2) essentially by adding to the right member of (2) the terms:

\[
S^2 \left[ L(j) P^*_j, i \right. \left. \left( d_j, i - P^*_j, i \right) a_{i'} \right] - \sum_{i' \neq i} S^2 \left[ P^*_j, i \right. \left. (d_j, i' - P^*_j, i' a_{i'}) \right]
\]

This suggests an iterative procedure with the solution of (2) as the first approximation. Then the correction (20) is computed for each star and added to the right member of (2), which is then solved again. It appears that a small-scale numerical simulation of this (as well as a direct solution of (19) and a solution according to the "old scheme", using the zero-points \(b_j\)) is needed in order to see the effects of the various approximations.