First Look Preprocessing

Hans-Heinrich Bernstein (s03@ix.urz.uni-heidelberg.de)
Sonja Hirte (hirte@ari.uni-heidelberg.de
Ulrich Bastian (bastian@ari.uni-heidelberg.de

November 30, 2005

Description of the Ring Solution

Documentcode: GAIA-ARI-BST-001, Issue 5

You can directly click on internal and external hyperlinks in the electronic version of this document.

Abstract

The task First Look Preprocessing (FLP) *provides a daily check of the functioning of GAIA on the basis of the scientific measurements of the satellite. Within this task, the Ring Solution is a special algorithm with some resemblances to the* Great-Circle Reduction of ESA's astrometry satellite HIPPARCOS. The observed stars are *connected by their astrometric, attitude and calibration parameters so they form a ring on the celestial sphere. Here, the special interest of the Ring Solution is the estimability and the accuracy of an appropriate set of parameters describing all quantities in question.*

Contents

1 Introduction to GAIA First Look Preprocessing

The scientific data analysis of the GAIA observations consists of a very complex and time consuming data reduction chain. Within a period of about five years approximately one to two hundred people will be active in different tasks. Probably, the success of GAIA becomes clear in the middle of the working time span only, unless special tasks clarify the functioning of the satellite at the beginning of the data analysis and constantly thrughout the mission. In order to accomplish this task different checks of the GAIA original measurements are foreseen. These checks are split in two levels: the Science Quick Look and the Detailed First Look (Biermann and Bastian 2005). The latter implies the necessity of a high precission auto-calibrating parameter adjustment which is the task **FLP**.

In parts, the GAIA FLP shows similarities to the Quick Look and the RGC Reduction tasks of HIPPARCOS (ESA-SP).

2 Details of the Ring Solution

In order to accomplish the GAIA FLP, several methods are conceivable. Regarding the observed stars of a measurement period of one day, all objects lie on a ring on the celestial sphere. The so-called Ring Solution combines the measurements of the observed stars within one day for a first high precission estimation of astrometric source, attitude and instrument calibration parameters. The main goal here is not the calculation of the most sophisticated values of these parameters, but the estimability and an acceptable error budget. This should show the proper functioning of GAIA. The Ring Solution accomplishes this analysis in a compact parameter adjustment without any iteration process. Nevertheless, the Ring Solution allows a star-by-star treatment of the measurements. The solution of the normal equations is a rigorous algorithm focused on the attitude and calibration parameters rather than on the astrometric source parameters. Here, the stars are used in a purely geometric way without any astrophysical meaning. The estimability of the parameter determination will be analysed during the calculation of the solution using a complete eigenvalue decomposition. So this method works almost independently from the number of used stars, which allows a comfortable treatment of the data.

2.1 The Observations

The location of a star in question in the focal plane at a certain time instant can be described by the corresponding field coordinates η and ζ. With the knowledge of the attitude and the instrument calibration parameters, one can derive the astrometric parameters of the observed stars. So, the field coordinates are the observations within the GAIA FLP task. As far as *a priori* values are available, η and ζ must be corrected concerning aberration, parallax and light bending effects, which is known as a full astrometric modelling, see (Lindegren 2001).

According to the observational principle of GAIA, the set of observations contains measurements of both fields of views of the instrument. Concerning observations of a star in question, we have double measurements, separated by the basic angle of GAIA and, with a spin rate of $60''/\mathrm{s}$, 1.65 hours apart of each other.

First of all, GAIA delivers one-dimensional information along scan. The star mapper CCDs and the AF1 CCDs (object confirmation) deliver observations across scan in addition with a reduced accuracy.

2.2 Quantities of the First Look Preprocessing

Normally, the positions of stars are described in equatorial or ecliptic coordinates (α, δ or λ, β). Since we limit our part of the celestial sphere to the mentioned ring, it is reasonable to use the spherical RGC coordinates v and r based on an arbitrary reference great circle within the ring (Bastian 2004). The timespan in question is about one day, so we neglect the proper motion, the parallax and the radial velocity as unknowns in the adjustment.

The attitude of GAIA will be described by quaternions at a dense grid of time instants and by B-splines between these time instants. In short, we summarize all these parameters in a part of the vector of unknowns denoted by $\mathbf{a} = (a_1, a_2, \ldots, a_k)$, consisting of B-spline coefficients. In a similar way we treat the calibration parameters $\mathbf{c} = (c_1, c_2, \ldots, c_l)$, not exactly specified at this stage of the work.

2.3 Modelling the Observations

In general, all unit vectors from the center to the celestial sphere's surface are connected by a more or less simple rotation matrix. So it is easy to get a mathematical relationship between our measurements and the three groups of unknowns. Observations and astrometric soure parameters form two of those unit vectors, while attitude and calibration parameters deliver a composed rotation matrix. Let **a** the vector of the attitude unknowns, **s** the vector of the astrometric source parameters of all sources and **c** the vector of the calibration parameters we get the following generalized observation equation using, see (Lindegren 2001):

$$\begin{pmatrix} \eta - \eta_\circ \\ \zeta - \zeta_\circ \end{pmatrix} = \begin{pmatrix} \partial\eta/\partial\mathbf{a} \\ \partial\zeta/\partial\mathbf{a} \end{pmatrix} \Delta\mathbf{a} + \begin{pmatrix} \partial\eta/\partial\mathbf{s} \\ \partial\zeta/\partial\mathbf{s} \end{pmatrix} \Delta\mathbf{s} + \begin{pmatrix} \partial\eta/\partial\mathbf{c} \\ \partial\zeta/\partial\mathbf{c} \end{pmatrix} \Delta\mathbf{c} + \begin{pmatrix} \epsilon_\eta \\ \epsilon_\zeta \end{pmatrix} \tag{1}$$

where, η, ζ are the observed field angles and $\eta_\circ, \zeta_\circ$ are the ones computed from *a priori* approximations of $\mathbf{a}, \mathbf{c}$ and $\mathbf{s}$, and e.g.

$$\begin{pmatrix} \partial\eta/\partial\mathbf{a} \\ \partial\zeta/\partial\mathbf{a} \end{pmatrix} = \begin{pmatrix} \partial\eta/\partial a_1 & \partial\eta/\partial a_2 & \ldots & \partial\eta/\partial a_k \\ \partial\zeta/\partial a_1 & \partial\zeta/\partial a_2 & \ldots & \partial\zeta/\partial a_k \end{pmatrix}.$$

ϵ_η and ϵ_ζ are the observational noise components. The vector **a** consists of several 10^3, **s** of several 10^7, and **c** of several 10^2 unknowns. In the following we denote these numbers of unknowns as n_a, n_s, and n_c. Their sum, the total number of unknowns, will be denoted u. **The handling of s is one of the central problems of GAIA FLP.**

2.4 Least Squares Approach and Direct Elimination of Unknowns

The character of the available information leads to the estimation of the unknowns using a least squares approach with constraints and directly observed parameters. We condense all (*a priori*) unknowns in a vector **x** and all observed minus computed values of $(\eta_i - \eta_{\circ_i})$ and $(\zeta_i - \zeta_{\circ_i})$ in a vector **l** and the directly observed parameters in the vector $\mathbf{l}_x$. One starts with the following two sets of observational equations and a set of constraints whereat the coefficient matrices **A**, **B** and **C** are containing the corresponding partial derivatives:

$$\mathbf{v} = \mathbf{A}\mathbf{x} - \mathbf{l}, \qquad \mathbf{v}_x = \mathbf{B}\mathbf{x} - \mathbf{l}_x \quad \text{and} \quad \mathbf{C}\mathbf{x} + \mathbf{w} = \mathbf{0}.$$

$\mathbf{v}$, $\mathbf{v}_x$ and $\mathbf{w}$ are the residual vectors of these two sets of observational equations and the set of constraints. Searching for an extremum with constraints, this least squares approach focusses on the minimisation of the following functional:

$$\mathcal{L} = \mathbf{v}^T \mathbf{Q}_{ll}^{-1} \mathbf{v} + \mathbf{v}_x^T \mathbf{Q}_{xx}^{-1} \mathbf{v}_x + 2\mathbf{k}^T (\mathbf{C}\mathbf{x} + \mathbf{w}) = Min$$

k is the vector of Lagrange multipliers. We need the following partial derivatives

$$
\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\mathbf{v}_l} &= 2\mathbf{Q}_{ll}^{-1}\mathbf{v}_l = 0, & \text{(a)}\\
\frac{\partial\mathcal{L}}{\partial\mathbf{v}_x} &= 2\mathbf{Q}_{xx}^{-1}\mathbf{v}_x = 0, & \text{(b)}\\
\frac{\partial\mathcal{L}}{\partial\mathbf{k}^T} &= 2(\mathbf{C}\mathbf{x}+\mathbf{w}) = 0, & \text{(c)}\\
\frac{\partial\mathcal{L}}{\partial\mathbf{x}} &= 2\mathbf{k}^T\mathbf{C} = 0. & \text{(d)}
\end{aligned}
$$

From (a) and (b) one gets:

$$
\begin{aligned}
\mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{A}\mathbf{x} &= \mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{l},\\
\mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{B}\mathbf{x} &= \mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{l}_x,
\end{aligned}
$$

summarized to

$$
(\mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{A}+\mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{B})\mathbf{x} = \mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{l}+\mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{l}_x.
$$

With (c) and (d) one gets the normal equations:

$$
\begin{pmatrix} \mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{A}+\mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{B} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{k} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^T\mathbf{Q}_{ll}^{-1}\mathbf{l}+\mathbf{B}^T\mathbf{Q}_{xx}^{-1}\mathbf{l}_x \\ -\mathbf{w} \end{pmatrix} \tag{2}
$$

In this least squares approach $\mathbf{Q}_{ll}$ is the covariance matrix of the measurements and $\mathbf{Q}_{xx}$ is the covariance matrix of the directly observed parameters. In general we expect non-vanishing correlations between the different measurements, but we estimate the correlation coefficients smaller than 0.1. So we treat $\mathbf{Q}_{ll}$ as a diagonal covariance matrix:

$$
\mathbf{Q}_{ll}^{-1} = \begin{pmatrix}
\frac{\sigma_\circ^2}{\sigma_{\eta_{1_1}}^2} & 0 & \dots & & & & & & & & & & \\
0 & \ddots & 0 & \dots & & & & & & & & & \\
\vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\eta_{i_1}}^2} & 0 & \dots & & & & & & & & \\
 & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\zeta_{1_1}}^2} & 0 & \dots & & & & & & & \\
 & & \vdots & 0 & \ddots & 0 & \dots & & & & & & \\
 & & & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\zeta_{i_1}}^2} & 0 & \dots & & & & & \\
 & & & & \vdots & 0 & \ddots & 0 & \dots & & & & \\
 & & & & & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\eta_{j_1}}^2} & 0 & \dots & & & \\
 & & & & & & \vdots & 0 & \ddots & 0 & \dots & & \\
 & & & & & & & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\eta_{j_n}}^2} & 0 & \dots & \\
 & & & & & & & & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\zeta_{j_1}}^2} & 0 & \dots \\
 & & & & & & & & & \vdots & 0 & \ddots & 0 \\
 & & & & & & & & & & \vdots & 0 & \frac{\sigma_\circ^2}{\sigma_{\zeta_{j_n}}^2}
\end{pmatrix}
$$

The rms of unit weight $\sigma_\circ^2$ can in principle be set to any value. This simple weighting model could be changed during the execution of the ring solution. In general, it is possible to take larger correlation coefficients into account, see (Höpcke 1980) after the execution of the estimation process in an additional correction term to the unknowns and the corresponding rms values. Our treatment of the measurements requires uncorrelated observations between different sources. The arrangement of the measurements can be done in two blocks (the η measurements and the ζ measurements) per source, which allows a simple method of inversion of the covariance matrix of the observations, see (Bernstein 1993), when the offdiagonal elements between η and ζ of the same star do not vanish. For each source j, $\mathbf{Q}_{jj}$ denotes the corresponding part of the covariance matrix $\mathbf{Q}_{ll}$ and $\mathbf{l}_j$ is the corresponding part of the $\mathbf{l}$ vector.

$\mathbf{Q}_{xx}$is considered also as a diagonal matrix and will be discussed later.

The arrangements of the unknowns in the corresponding vectors and matrices is as follows:

$$\mathbf{x} = \begin{pmatrix} v_1 \\ r_1 \\ \vdots \\ v_n \\ r_n \\ - \\ a_1 \\ \vdots \\ a_k \\ - \\ c_1 \\ \vdots \\ c_l \end{pmatrix} = \begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_a \\ \mathbf{x}_c \end{pmatrix}, \qquad \mathbf{l} = \begin{pmatrix} (\eta-\eta_0)_{1_1} \\ \vdots \\ (\eta-\eta_0)_{i_1} \\ (\zeta-\zeta_0)_{1_1} \\ \vdots \\ (\zeta-\zeta_0)_{i_1} \\ \vdots \\ (\eta-\eta_0)_{j_1} \\ \vdots \\ (\eta-\eta_0)_{j_n} \\ (\zeta-\zeta_0)_{j_1} \\ \vdots \\ (\zeta-\zeta_0)_{j_n} \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} \frac{\partial l_m}{\partial x_u}, & \cdots \\ \vdots & \ddots \end{pmatrix} \tag{3}$$

Due to the value of u, Equation (2) cannot be used directly to solve the FLP problem. It is noteworthy that the inversion of Equation (2) is an u^3-process. So, one needs a different strategy to get $\mathbf{x}$. Our strategy is to split up $\mathbf{x}$ into several groups and to estimate the unknowns of these groups in such a way as to accomplish Equation (2) directly. Our working strategy utilizes the technique of the direct elimination of unknowns using the Gaussian elimination process in connection with an adjustment of different groups of unknowns and profiting from the addition theorem of normal equation matrices. Those algorithms are well established in geodesy for the unification of several originally separate parts of a network, see (Wolf 1975) and (Großmann 1975). In the FLP the astrometric parameters are eliminated directly while the full set of all attitude and calibration parameters are common unknowns for all stars. It is remarkable that for each star the full set of all attitude and calibration parameters and eventually, the needed constraints must be taken into account, even those parameters, where the observations are not sensitive. This leads to an overwhelmingly large part of zero elements of the normal equation parts. However, after the direct elimination of the astrometric parameters this part changes into a non zero part, which is a consequence of the Gaussian algorithm. As mentioned, we split $\mathbf{x}$ in a source part, containing the astrometric parameters of all sources in question, and in a so-called satellite part containing the attitude and calibration parameters. Shown in Figure 1, we get the corresponding structure of the Jacobian matrix.

The corresponding observation equations for all stars n read as follows:

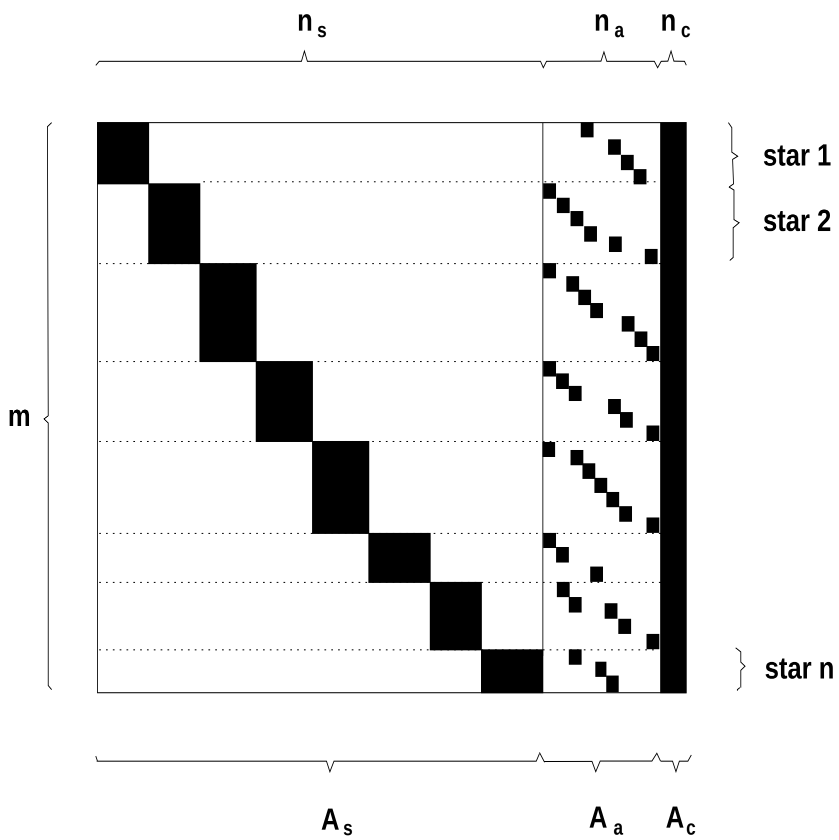

Figure 1: Structure of the Jacobian matrix. White areas contain zeros.

$$\begin{array}{lllllllll} \mathbf{v}_1 & = & \mathbf{A}_{s_1}\mathbf{x}_{s_1} & + & \mathbf{A}_a\mathbf{x}_a & + & \mathbf{A}_c\mathbf{x}_c & - & \mathbf{l}_1, \\ \mathbf{v}_2 & = & \mathbf{A}_{s_2}\mathbf{x}_{s_2} & + & \mathbf{A}_a\mathbf{x}_a & + & \mathbf{A}_c\mathbf{x}_c & - & \mathbf{l}_2, \\ \vdots & & & & \vdots & & & & \\ \mathbf{v}_n & = & \mathbf{A}_{s_n}\mathbf{x}_{s_n} & + & \mathbf{A}_a\mathbf{x}_a & + & \mathbf{A}_c\mathbf{x}_c & - & \mathbf{l}_n. \end{array} \tag{4}$$

So far, we have source, attitude and calibration parameters located along the same RGC without fixing a zero point and the basic RGC plane. In general, one can rotate the RGC plane in space. So we have three degrees of freedom within our adjustment problem. This leads to a rank deficient problem. In order to overcome this problem one has several possibilities:

First, one could introduce the approximate values of the source coordinates as directly observed unknowns with their corresponding accuracy. This would not disturb our strategy of the direct elimination of the astrometric parameters because this would not change the structure of the resulting normal equation matrix. The influence of those observational equations is like a 'weak constraint' which leads to an appropriate treatment of the approximate values with their rms errors. So every equation in (4) could get two more entries than the number of measurements of a star in question: one for υ and one for r:

$$\upsilon_{obs.} + v_\upsilon = x_\upsilon \qquad \text{and} \qquad r_{obs.} + v_r = x_r. \tag{5}$$

The design matrices $\mathbf{A}_{s_j}$ would be enlarged by two rows, the first one contains the number 1 for υ and the second row contains the number 1 for r all the other elements are zero. The covariance matrix of the observations would be enlarged by two diagonal elements ${m_0^*}^2/m_{\upsilon_j}^2$ and ${m_0^*}^2/m_{r_j}^2$, the off-diagonal elements are zero. ${m_0^*}^2$ is a free parameter and the subject of the fine tuning of the data analysis. The introduction of $\upsilon_{obs.}$ and $r_{obs.}$ with their corresponding accuracy is equivalent to the localization of the source network of a ring in question on the *a priori* known astrometric source parameters of reference. A disadvantage are the possible sytematic errors within the *a priori* used star catalogue which could lead to distortions in the attitude determination.

Secondly, one could stop those rotations during the utilization of the Moore-Penrose pseudoinverse. So one would get a mean location of the ring as a consequence of the minimal trace property of the Moore-Penrose pseudoinverse.

Thirdly, three rotation angles could be introduced as global parameters describing these global rotations. External information or additional constaints concerning these global rotations could determine the corresponding values.

At this stage of the work, we choose the second possibility, because the Moore-Penrose pseudoinverse will be used for other reasons within the parameter estimation process.

Furthermore, it is clear that one can move all stars along scan and one can change the CCD location parameters in the same way with the opposite sign, so we have a rotation about the RGC pole axis. In order to stop these rotations we suggest to introduce the following constraint:

$$\sum_{along\ scan} \mathbf{x}_c = 0. \tag{6}$$

In addition, one can tilt the basic RGC plane by a joint shift of the CCD location parameter across scan and one can change the attitude parameters correspondingly. In order to avoid those rotations, one needs two constraints in addition, one for each instantaneous field of view (IFOV):

$$\sum_{\substack{across\ scan\\ IFOV\ =\ 1}} \mathbf{x}_c = 0 \qquad \text{and} \qquad \sum_{\substack{across\ scan\\ IFOV\ =\ 2}} \mathbf{x}_c = 0. \tag{7}$$

Perhaps, a more sophisticated calibration model needs to distiguish short and long (time) scale calibration parameters and the number and form of these constraints will then change. At the present stage of the work we are dealing with large scale calibration parameters only.

Another group of constraints could be set up concerning the attitude parameters which consists of sixteen B-spline coefficients per time instant. The B-spline coefficients represent the four quaternion components $\mathbf{q} = (q_1, q_2, q_3, q_4)$ at the same time instants. It is well known that those quaternions are not independent of each other. In general we have the constraint

$$q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1 \,\big|_t. \tag{8}$$

Since these constraints do not contain any astrometric source parameters we could add these constraints directly to the reduced normal equation matrix, However, we ought to add those equations to each attitude parameter concerning to the same time instant. This means that the reduced normal equation matrix would be exhausted and we could not tackle such a matrix with our computer facilities. Much more convenient is the introduction of so-called 'weak constraints'. One can reformulate the quaternion constraint as an additional observation equation with a reasonable weight at each time instant. Such a procedure allows the separate treatment of these equations and one must add this resulting part of the normal equation matrix to the reduced (and accumulated) normal equation matrix of the attitude and calibration parameters. A reasonble weigth for these observational equations is the largest eigenvalue of the original reduced normal equation, which can be calculated easily.

The constraints concerning the calibration enter our adjustment problem using the Lagrange multipliers $\mathbf{k}_c$. With $\mathbf{w}_c$ containing the contradictions from Equations (6) and (7) before estimation of $\mathbf{x}_c$ we get the following constraints:

$$\mathbf{C}_c^T \mathbf{x}_c = -\mathbf{w}_c, \tag{9}$$

where $\mathbf{C}_c^T$ is the corresponding coefficient matrix projecting $\mathbf{x}_c$ along scan or across scan. Equation (8) contains no astrometric source parameters, so one can add these constraints after the elimination of υ_j and r_j.

The reformulation of equation (8) as an observation equation leads to

$$\mathbf{l}_x = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 2q_{1,t_1} & 2q_{2,t_1} & 2q_{3,t_1} & 2q_{4,t_1} & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 2q_{1,t_2} & 2q_{2,t_2} & 2q_{3,t_2} & 2q_{4,t_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 & 0 & 0 & \ddots \end{pmatrix}$$

$$\text{and} \quad \mathbf{B}^T \mathbf{Q}_{x,x}^{-1} \mathbf{B} = \mathcal{N}_{a,a}. \tag{10}$$

After some algebra one gets the normal equations containing the corresponding blocks:

$$\begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_a \\ \mathbf{x}_c \\ \mathbf{k}_c \end{pmatrix} = \begin{pmatrix} \mathbf{N}_{s,s} & \mathbf{N}_{s,a} & \mathbf{N}_{s,c} & \mathbf{0} \\ \mathbf{N}_{a,s} & \mathbf{N}_{a,a} + \mathcal{N}_{a,a} & \mathbf{N}_{a,c} & \mathbf{0} \\ \mathbf{N}_{c,s} & \mathbf{N}_{c,a} & \mathbf{N}_{c,c} & \mathbf{C}_c \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_c^T & \mathbf{0} \end{pmatrix}^{+} \cdot \begin{pmatrix} \begin{pmatrix} \mathbf{A}_{s_1}^T & \mathbf{A}_{s_2}^T & \cdots & \mathbf{A}_{s_n}^T & \mathbf{0} \\ \mathbf{A}_{a_1}^T & \mathbf{A}_{a_2}^T & \cdots & \mathbf{A}_{a_n}^T & \mathbf{B}^T \\ \mathbf{A}_{c_1}^T & \mathbf{A}_{c_2}^T & \cdots & \mathbf{A}_{c_n}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_{ll}^{-1}\mathbf{l} \\ \mathbf{Q}_{xx}^{-1}\mathbf{l}_x \end{pmatrix} \\ -\mathbf{w}_c \end{pmatrix}. \tag{11}$$

Since the normal equation matrix is symmetric, we have
$\mathbf{N}_{a,s} = \mathbf{N}_{s,a}^T$, $\mathbf{N}_{c,s} = \mathbf{N}_{s,c}^T$ and $\mathbf{N}_{c,s} = \mathbf{N}_{a,c}^T$.

Equation (11) deals with the Moore-Penrose pseudoinverse of the normal equation matrix, indicated by the superscript (+). In spite of the introduction of constraints like Equation (8), we expect a rank deficient problem since one does not know whether all reasonable and actually needed parameters of the full functional model are well represented in our data or not. This topic will be discussed in section 2.5 in more detail. The utilization of the Moore-Penrose pseudoinverse allows one to overcome those problems in general.

The matrix $\mathbf{N}_{s,s}$ is a very large block diagonal matrix, where each block is a 2×2 symmetric matrix and all offdiagonal blocks are zero matrices. So, one gets the structure of the full normal equation matrix as shown in Figure 2. This structure and the large number of blocks, perhaps 2.5×10^6, in connection with the main goal of this task as described in the introduction leads to the direct elimination techniques using the Gauss algorithm eliminate the astrometric source parameters from Equation (11). We expect to cope with the reduced normal equation matrix using all well known features of linear algebra.

Figure 3 shows the structure of the reduced normal equations.

The distances of the diagonal lines in Fig. 3 and the increasing size of the interuptions are consequences of the scanning law. The rough border of the diagonal lines, seen in Figure 4, as the fine structure of Fig. 3, shows the interactions between the scanning law and the source distribution remarkably. The last column in the pattern of Fig. 3 is due to the calibration parameters and the constraints (6) and (7).

This leads to the following working scheme illustrated on Figure 5:

1. starting a loop running over all objects in question,
2. setup of the normal equations for the current object with all needed constraints,
   (a) condition test and, if necessary, reduction of the number of unknowns or cancellation of the current object,

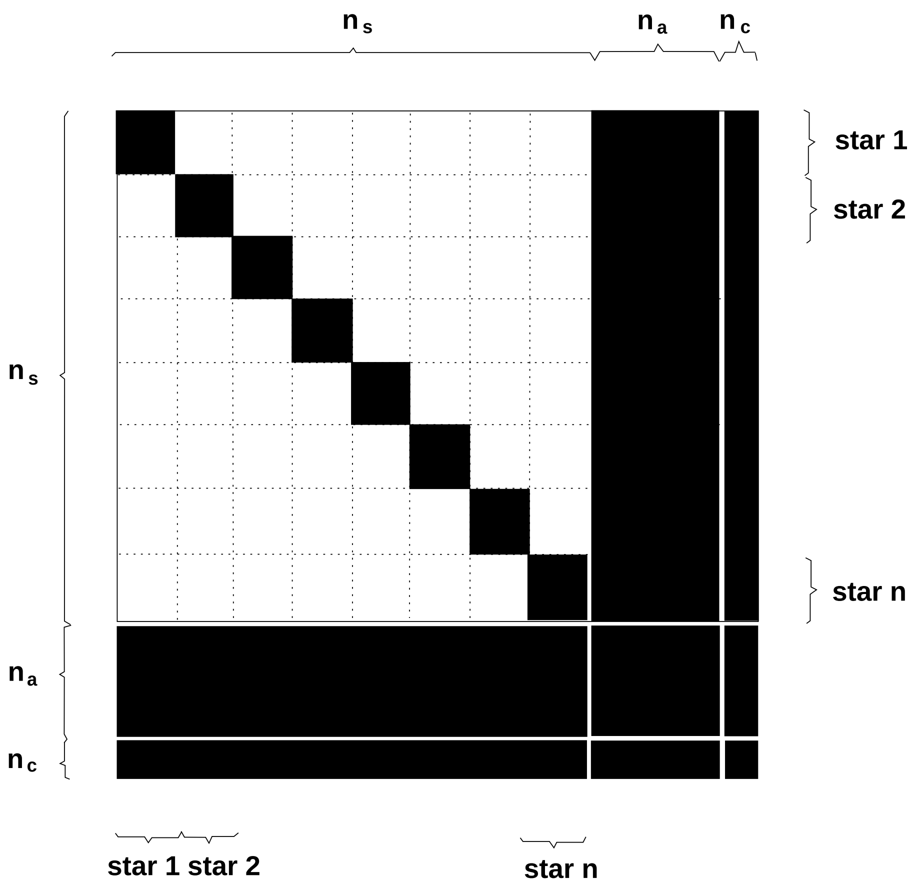

Figure 2: Structure of the original normal equation matrix. White areas contain zeros.

3. elimination of the astrometric parameters from these normal equations,
4. accumulation of the remaining parts of normal equations,
5. terminating this loop and solving the final part of the normal equation matrix, which is the part of **a** and **c**,
6. estimation of the astrometric parameters in a subsequent loop over all objects by back-substitution of the parts of the normal equations,
7. calculation of the corresponding error budget.

Using this scheme, equation (7) reduces to a new equation for all parameters, apart the source unknowns. For each star j, we get the corresponding normal equation part:

$$\left(\mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} \mathbf{A}_{s_j}\right)\mathbf{x}_s + \mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} \left(\mathbf{A}_{a_j}, \mathbf{A}_{c_j}\right) \begin{matrix} \mathbf{x}_a \\ \mathbf{x}_c \end{matrix} = \mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} \mathbf{l}_j. \tag{12}$$

$\mathbf{Q}_{jj}$ is the block diagonal part of $\mathbf{Q}_{ll}$ for a current source j. Direct elimination of the $\mathbf{x}_{s_j}$ from all equations (12) delivers a new normal equation part for each star:

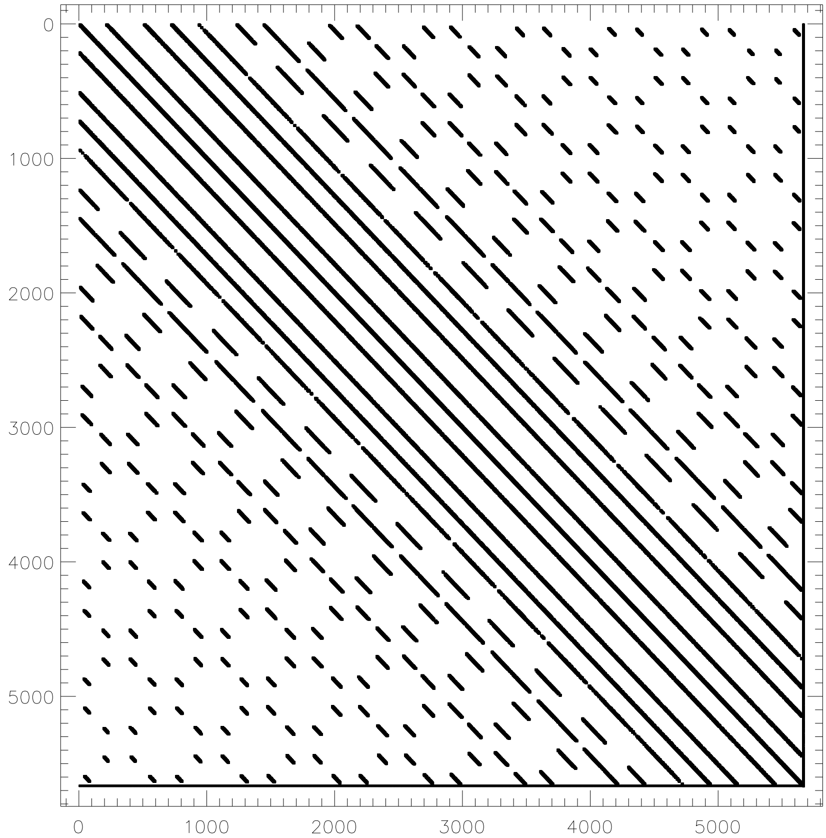

Figure 3: Structure of the reduced normal equation matrix based on dataset DSL. Black dots show the non-zero elements of the matrix.

$$\left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)^T\mathbf{Q}_{jj}^{-1}\left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)\Big|_{2\times red.}\begin{matrix}\mathbf{x}_a\\ \mathbf{x}_c\end{matrix} = \left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)^T\mathbf{Q}_{jj}^{-1}\mathbf{l}_j\Big|_{2\times red.}. \tag{13}$$

This is a new equation for each star with exactly the same unknowns in common. The accumulation delivers the final normal equation for these unknowns but still without the constraints:

$$\sum_{j=1}^{n}\left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)^T\mathbf{Q}_{jj}^{-1}\left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)\Big|_{2\times red.}\begin{matrix}\mathbf{x}_a\\ \mathbf{x}_c\end{matrix} = \sum_{j=1}^{n}\left(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\right)^T\mathbf{Q}_{jj}^{-1}\mathbf{l}_j\Big|_{2\times red.}. \tag{14}$$

According equation (2) the constraints change equation (14) to the border matrix structure:

$$\begin{pmatrix}\sum_{j=1}^{n}(\mathbf{A}_{a_j},\mathbf{A}_{c_j})^T\mathbf{Q}_{jj}^{-1}(\mathbf{A}_{a_j},\mathbf{A}_{c_j})\Big|_{2\times red.} & \mathbf{C}^T\\ \mathbf{C} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{x}_a\\ \mathbf{x}_c\\ \mathbf{k}\end{pmatrix} = \begin{pmatrix}\sum_{j=1}^{n}(\mathbf{A}_{a_j},\mathbf{A}_{c_j})^T\mathbf{Q}_{jj}^{-1}\mathbf{l}_j\Big|_{2\times red.}\\ -\mathbf{w}\end{pmatrix}. \tag{15}$$

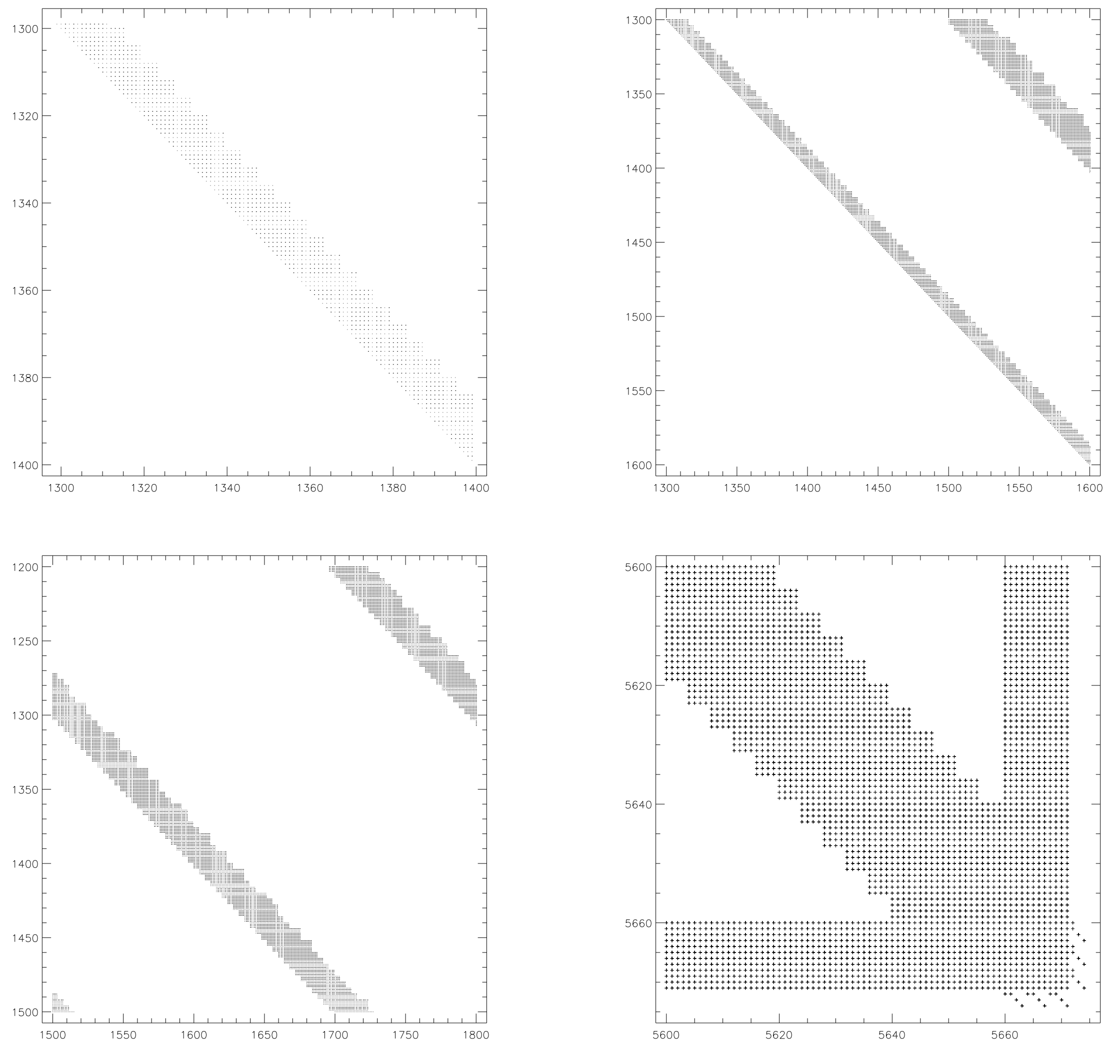

Figure 4: The four plots are magnified parts of the reduced normal equation matrix showing the fine structure of figure 3. Black dots and crosses mark the non-zero elements. The lower right panal shows the lower right corner of the matrix with the diagonal band of the attitude unknowns, the horizontal and vertical bands the calibration unknowns and the last three rows and columns the constraints.

The set up of equation (15) is the first *number crunching subtask* within the GAIA FLP. Equation (15) has approximately 5000 unknowns. In order to calculate the $\mathbf{x}_a$ and $\mathbf{x}_c$ from equation 15) we expect to tackle this equation directly, as described in the next chapter.

The astrometric parameter estimation for each star j is then given by

$$\mathbf{x}_{s_j} = \left(\mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} \mathbf{A}_{s_j}\right)^{-1} \left[\mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} \mathbf{l}_j - \mathbf{A}_{s_j}^T \mathbf{Q}_{jj}^{-1} (\mathbf{A}_{a_j}, \mathbf{A}_{c_j}) \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_c \end{pmatrix}\right], \tag{16}$$

obtained from (Wolf 1975).

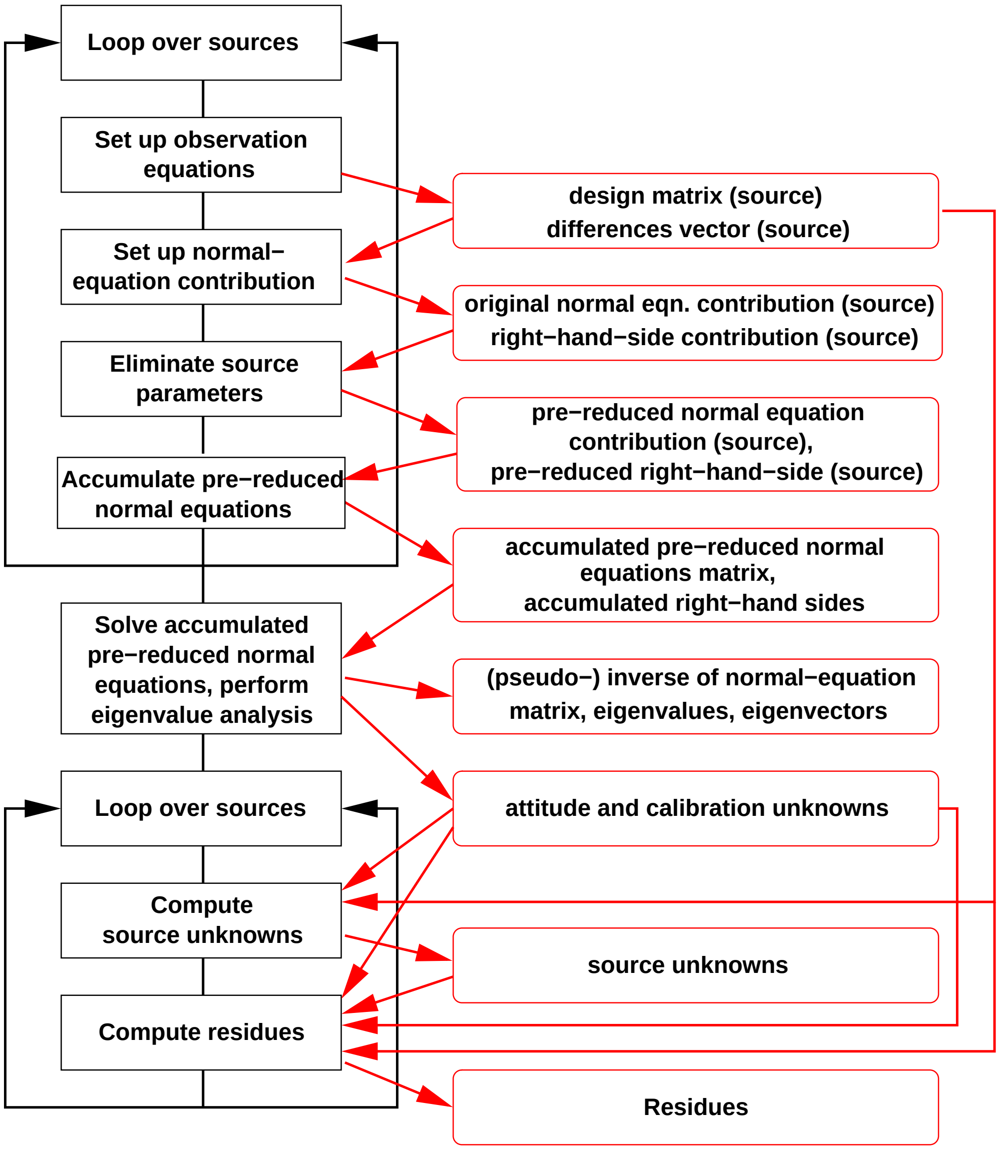

Figure 5: Illustration of the working scheme.

2.5 The Estimation and Estimability of the Attitude and Calibration Parameters

The inversion of the normal equation in (15) is a central problem in the GAIA FLP. In spite of the constraints (6) and (7) we consider equation (15) as a rank deficient problem. One can not be sure that the functional models of the attitude and the calibration will be full established in the approximately one day of measurements of the GAIA FLP. In addition, basic observations are done along scan, while the across scan information from the star mapper CCDs and the AF1 CCDs show much lower accuracy. This is the reason why we denoted the inversion in equation (11) as a Moore-Penrose pseudoinverse. In the case of a regular, positiv definite reduced normal equation matrix, the Moore-Penrose pseudoinverse equals a classical inverse by Gauß or Cholesky. The rank of a Moore-Penrose pseodoinverse is the same as the original reduced normal equation matrix. In practice we solve the complete eigenvalue problem with the modal matrix **M** containing the eigenvectors and the diagonal matrix Λ containing the eigenvalues:

$$\begin{pmatrix} \sum_{j=1}^{n} (\mathbf{A}_{a_j}, \mathbf{A}_{c_j})^T \mathbf{Q}_{jj}^{-1} (\mathbf{A}_{a_j}, \mathbf{A}_{c_j}) \Big|_{2\times red.} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{pmatrix} = \mathbf{M}\Lambda\mathbf{M}^T \tag{17}$$

The inverse is then:

$$\left(\mathbf{M}\Lambda\mathbf{M}^T\right)^+ = \mathbf{M}\Lambda^+\mathbf{M}^T. \tag{18}$$

Solving the complete eigenvalue problem allows the detection of over-stressing of the used functional model with respect to the short time span of data accumulation. A high-pass filter will be used to truncate *Lambda* for the rank partitioning. The needed treshold parameter is subject of some investigations and depends on the numerical condition of the reduced normal equation matrix. Using **M** and the zero subspace of *Lambda* we can detect whether certain parameters in the functional models are estimable or not. Under some circumsdances a reformulation of the functional model could be necessary. A rerun of the data analysis is then the consequence. This procedure shows some similarities to the well known truncated singular value decomposition in so-called *"schlecht gestellte Probleme"*, see (Louis 1989).

Figure 5 show the eigenvalue spectra for a given simulated data set. The spectral length of the eigenvalue spectra in Fig. 5 is of about 1×10^{15}. This indicates a numerical bad condition, and means, that the reduced normal equation is close to be positiv semidefinit with eigenvalues close to zero. This is an additional justification for the used Moore-Penrose pseodoinverse. The shape of the curve of the eigenvalue spectra shows four classes of eigenvalues: extreme high significant (a), high significant (b), significant (c) and not significant (d) region. The transitions from (a) to (b) and from (b) to (c) are well determined step functions. The transition (c) to (d) is a continuous slope. From such an eigenvalue classification results a special classification of the unknown parameters of the functional model, not investigated at this stage of the work.

2.6 The Error Budget

According the minimized functional $\mathcal{L}$ and using equation (4) one can calculate the rms of unit weight m_0:

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix}, \qquad m_0 = \sqrt{\frac{\mathbf{v}^T\mathbf{Q}_{ll}^{-1}\mathbf{v} + \mathbf{v}_x^T\mathbf{Q}_{xx}^{-1}\mathbf{v}_x}{m + m_x - u}}, \tag{19}$$

where m denotes the number of all observations, m_x denotes the number fictitious observations and u is the number of all unknowns including the astrometric source parameters. $\mathbf{Q}_{ll}$ ist the covariance matrix of the measurements with a diagonal or block diagonal structure. $\mathbf{Q}_{ll}^{-1}$ can be calculated easily, since this matrix is positive definite per definition and sparse. With the *a priori* known (or estimated) σ_0 one can use the

$$\chi^2 = \frac{m_\circ^2}{\sigma_\circ^2}(m + m_x - u) \tag{20}$$

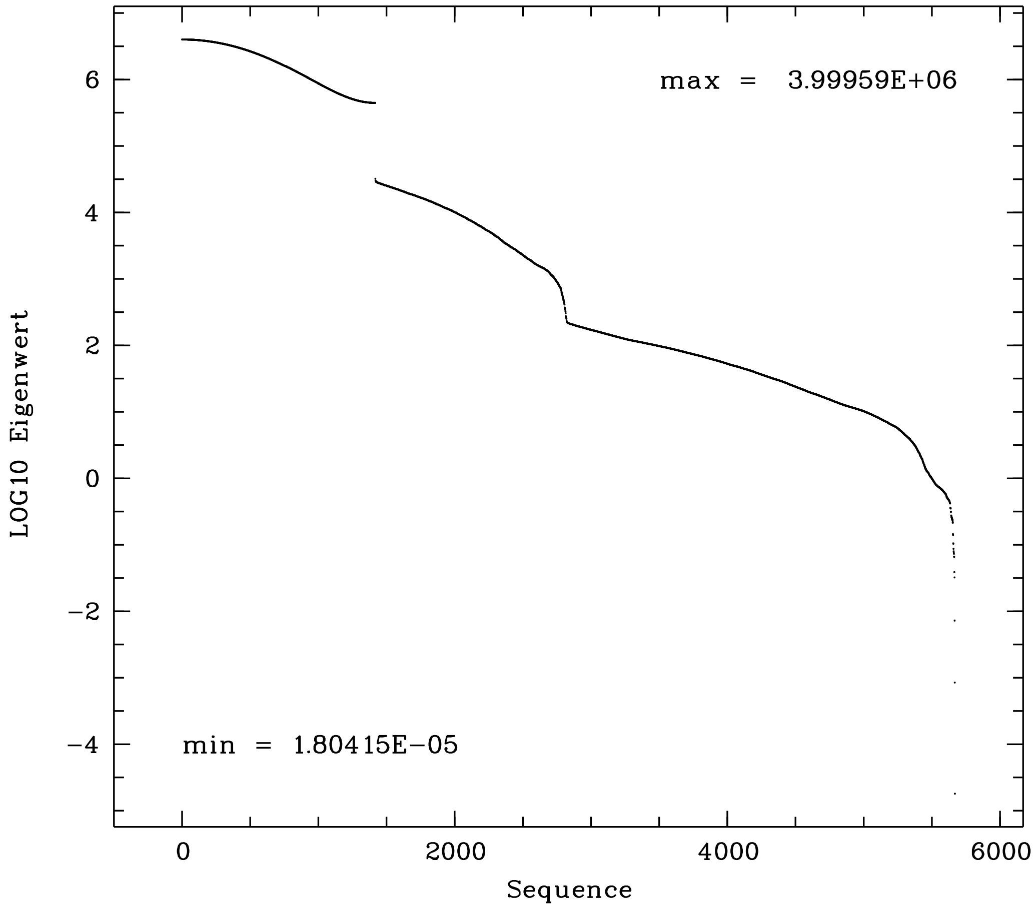

Figure 6: Eigenvalues of the reduced normal equation matrix with constraints based on dataset DSL.

for a lot of statistical tests. Especially the χ^2-test or the F-test allows one to check whether the satellite observations reached the expected accuracy or not: One of the first and main goals of this task. This is the first indication of the functioning of the satellite.

Denoting the covariance matrix of the unknowns in common by $\mathbf{G}_{x,x}$ which is the inverse of the reduced normal equation matrix in equation (15) the rms of an unknown x_n is given by

$$m_x^2 = m_\circ^2 \cdot g_n, \tag{21}$$

where g_n are the main diagonal elements of $\mathbf{G}_{x,x}$. Using (Wolf 1975) p.115 one can compute the block diagonal covariance matrix of the different sources j:

$$\mathbf{G}_{s_j,s_j} = \mathbf{G}_{j,j}^{-1} + \mathbf{G}_{j,j}^{-1}\mathbf{A}_{s_j}^T\mathbf{Q}_{jj}^{-1}\big(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\big)\mathbf{G}_{x,x}\big(\mathbf{A}_{a_j},\mathbf{A}_{c_j}\big)^T\mathbf{Q}_{jj}^{-1}\mathbf{A}_{s_j}\mathbf{G}_{j,j}^{-1}, \tag{22}$$

with $\mathbf{G}_{j,j}^{-1} = \big(\mathbf{A}_{s_j}^T\mathbf{Q}_{jj}^{-1}\mathbf{A}_{s_j}\big)^{-1}$. The rms errors of the source parameters are:

$$m_{s_j}^2 = m_\circ^2 \cdot g_{s_j}. \tag{23}$$

g_{s_j} are the main diagonal elements of $\mathbf{G}_{s_j,s_j}$. Equation 20 allows the computation of covariances of the unknowns for a star in question. The computation of covariances between unknowns of different stars

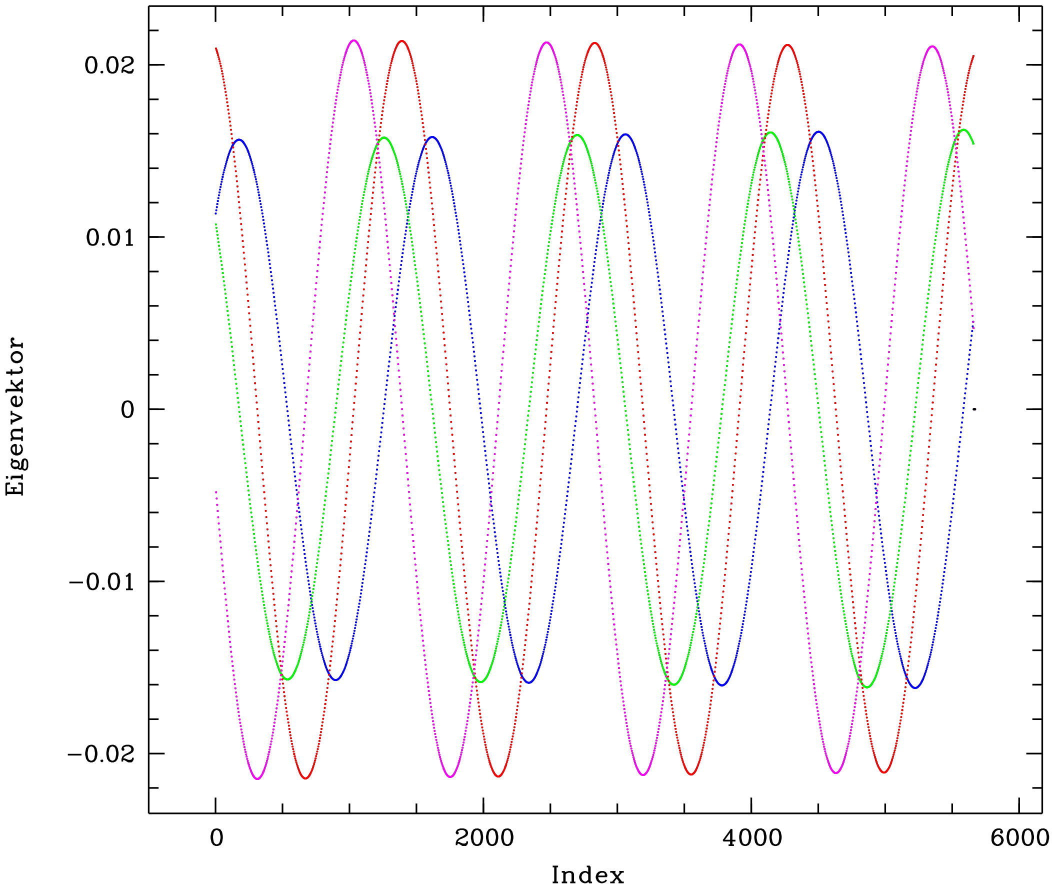

Figure 7: Eigenvectorcomponents of the first zero-eigenvalue of the reduced normal equation matrix with constraints based on dataset DSL.

is in practise not feasible in the Ring Solution. The calculation of the covariance matrix of the residuals $\mathbf{Q}_{vv}$ is feasible in blocks star by star:

$$\mathbf{Q}_{v_j v_j} = \mathbf{Q}_{jj} - \mathbf{A}_{s_j} \mathbf{G}_{j,j}^{-1} \mathbf{A}_{s_j}^T - \left(\mathbf{A}_{a_j}, \mathbf{A}_{c_j}\right) \mathbf{G}_{x,x} \left(\mathbf{A}_{a_j}, \mathbf{A}_{c_j}\right)^T . \tag{24}$$

2.7 Used Constants and Units

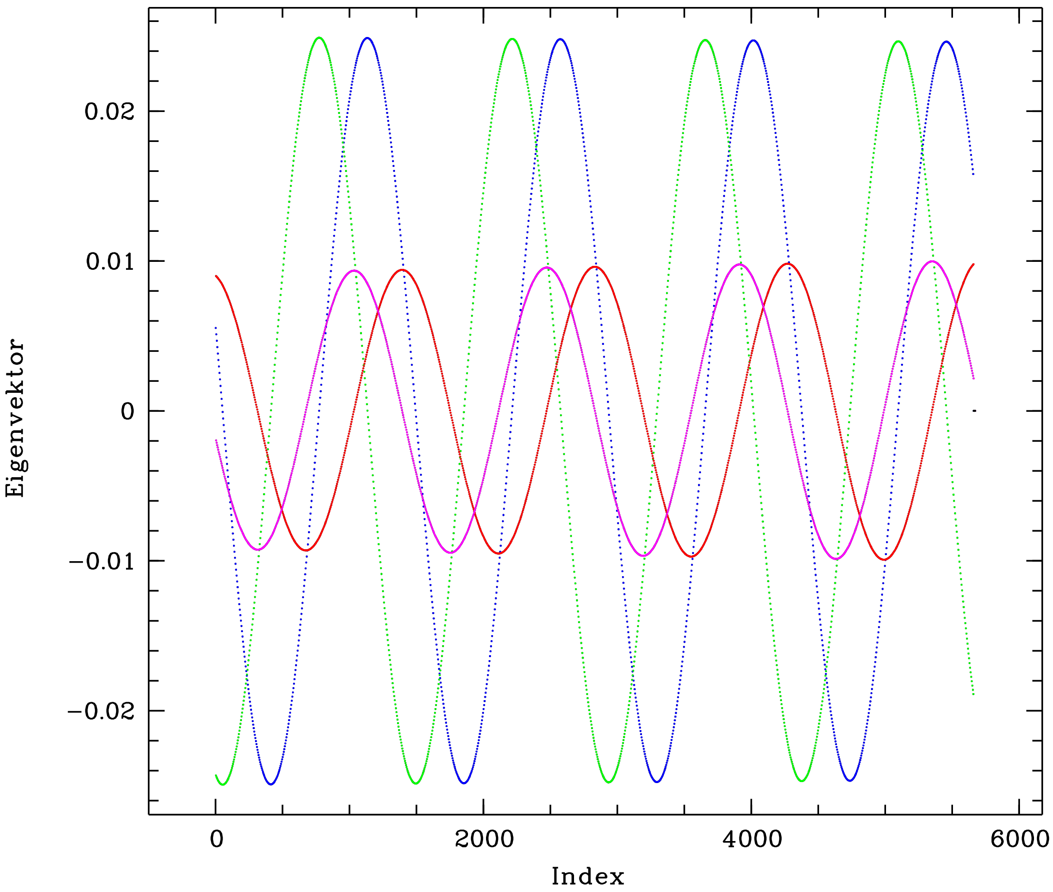

Figure 8: Eigenvectorcomponents of the second zero-eigenvalue of the reduced normal equation matrix with constraints based on dataset DSL.

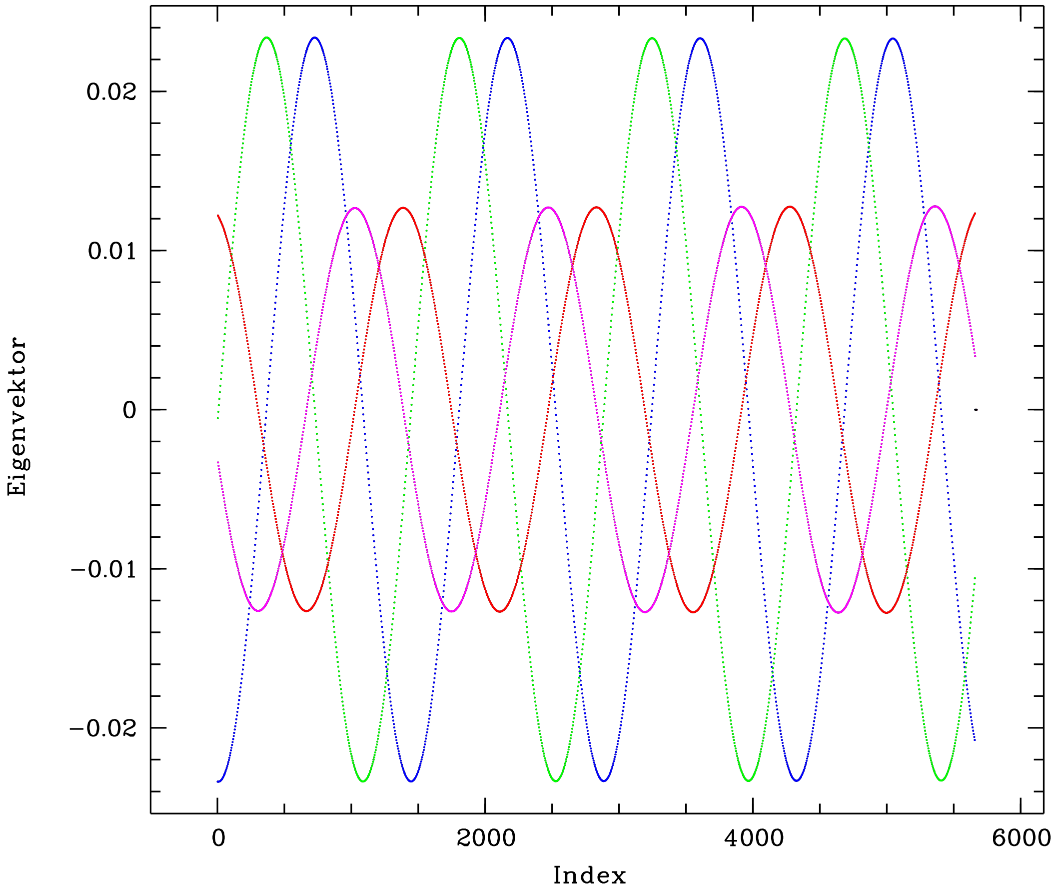

Figure 9: Eigenvectorcomponents of the third zero-eigenvalue of the reduced normal equation matrix with constraints based on dataset DSL.

3 Investigations of the Residuals

The investigations on the capability and functioning of the satellite includes the estimability of the parameters in question, the achievement of the expected error budget and a random and undisturbed behaviour of the residuals. Concerning the last demand on the measurements, we plan to estimate a series of appropriate covariance functions, some very special Fourier and/or Wavelet analysis, and we want to use the data to extract special effects by some special filterings. The residuals will be treated as prewhited time series of unevely spaced data. A reference for all those diagnostics could be obtained, when one executes the parameter adjustment and all computations concerning the residual investigations twice: first, a random number generator could be used to generate so-called quasi ideal observations which lead to a new vector of observed minus computed values $\mathbf{l}_\circ$. The application of this new vector to the full machinerie of the adjustment and error diagnostics delivers the referencee for the second step when one execute the same formulas with the real observations $\mathbf{l}$. A series of F-test statistics could be used to check whether the satellites measurements achieved the expected results or not.

For this purpose, one must summarize and reorder the residuals which are obtained star by star in section 2.7. concerning different criteria. For example, all residuals of all stars of magnitude m with $m_{min} \leq m \leq m_{max}$, where m_{min} and m_{max} must be defined appropriately, ordered by ascending time, form such a special class of data. Those classes can be formed by very different cryteria: magnitude differences can be used as well as colour differences. At this stage of the work, those classes are not established, however one needs such a classification during the residual analysis of the FLP Task. These classes of data are considered as the basis of the calculations of the subchapters **a**), **b**) and **c**).

3.1 Covariance Functions

In order to estimate covariances, one must bin the data of each class. Now, there is no experience concerning a reasonable time span for such a binning. But some simulations can help to obtain an appropriate binning. In general, such an empirical covariance function is given for evenly spaced data:

$$cov(\tau) = \frac{1}{N}\sum^{N}(v(t) - \widehat{v})(v(t+\tau) - \widehat{v}). \tag{25}$$

Applying equation (24), one gets a series of covariance functions where each will probably look like a series of Bell-shaped curves. Comparing these covariance curves, we expect a smooth crossing from one curve to each other. Any kind of discontinuty or leap is the reason for an alert about the functioning of the satellite.

3.2 Fourier and Wavelet Analysis

Here we summarize all residuals of all stars and reorder these data by ascending time instants. A subsequent Fourier analysis should not show any significant higher order of harmonics. Otherwise, one must introduce a deformation (warp) of the basic great circle plane, which is a contradiction to our assumptions in the data analysing method and a hint about a problem of the satellite.

In principle, disturbances of the satellite often enter as a cyclic but not strictly periodic effect in the observations. For example, a disturbance occure at different frequencies within certain time spans. In such a case a wavelet analysis is a better tool than a Fourier analysis. At this stage of the work no details about such an analysis is specified, but generally foreseen.

3.3 Filtering

In the subchapters **a**) and **b**) we tried to extract any kind of disturbances from the residuals, which are treated as time series. In contrast, one can benefit from a transformation of such time series. Moreover, we make an assumption of a special disturbance and we try to estimate the time location and the corresponding strength. Here we plan to accomplish several types of data filtering, like Wiener filter or Joseph-Bucy filter for example. During the measurement phase of the satellite, we hope to get **no** significant signal, which could give a hint for a failure of the satellite.

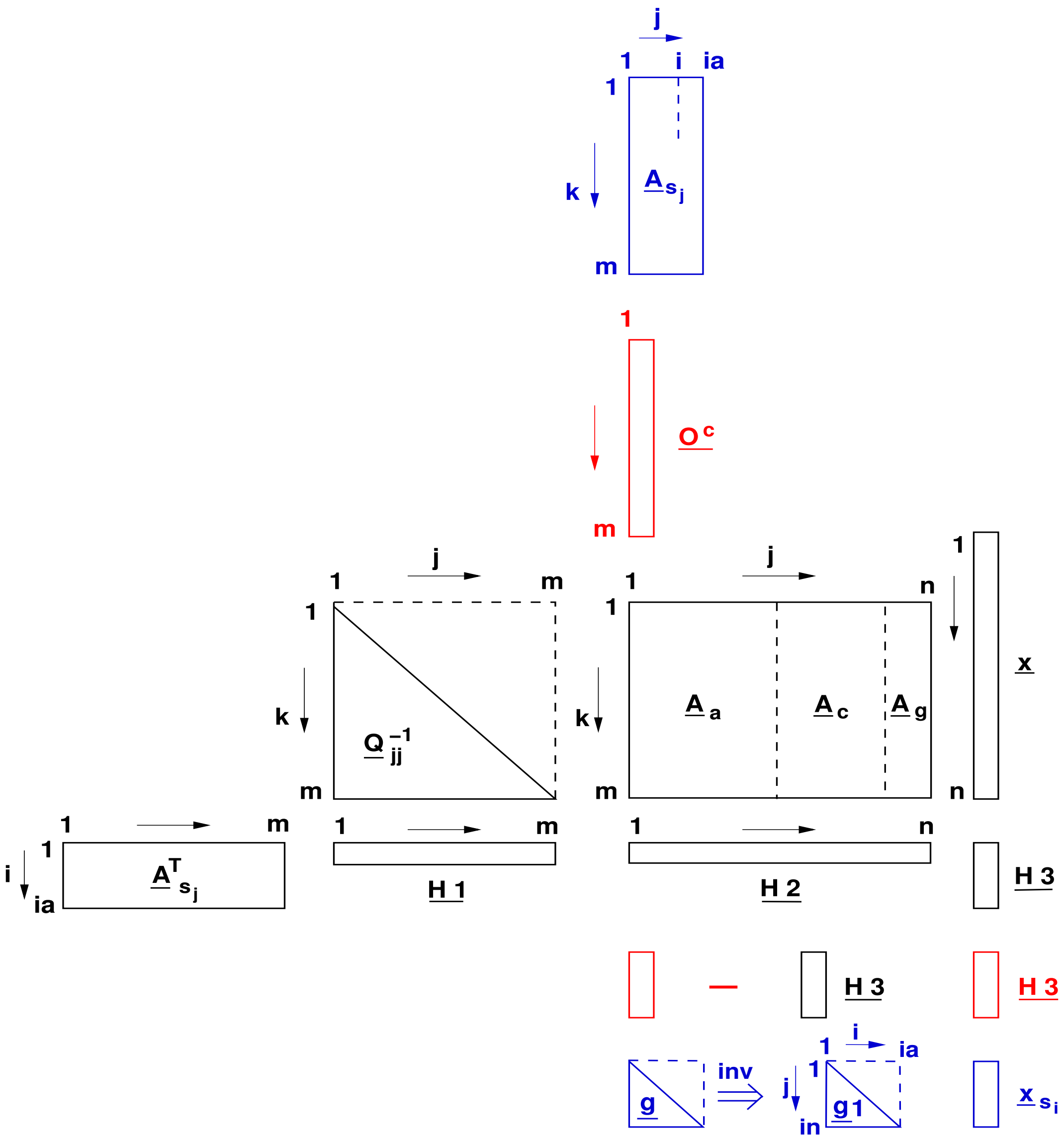

Figure 10: Programming Scheme of Equation 16

4 Software Prototype

4.1 Equipment

4.1.1 Hardware

large RAM necessary
tucana: DualPentiumIII 1000 MHz 4096 MB RAM
hydra: DualPentium 4 Xeon 2200 MHz 4096 MB RAM
ceres: Intel Pentium4-64 mit 3400 MHz and 2048 MB RAM

4.1.2 Software

External Software EISPACK http://www.netlib.org/eispack/
Random Number Generator
NumRec ??

GAIA-Software GAIA-LL37 (Lindegren 2001)

4.2 Input Data

4.2.1 Simulated GAIA-Data

Table 1: Data sets used

name	noise level	number of objects	number of measurements per co-ordinate
DSI	true	7716	94141
DSK	exp3	7716	94140
DSL	true	40091	488823
DSM	exp3	40091	488821

4.3 Software Implementation and Data Flow

4.3.1 Computing Time Simulation

In the first step a simulation of the computing time using different compilers and computers was provided (Hirte 2004). It could be shown that it is possible to deal with the size of the normal equations which is necessary for the first look.

4.3.2 Tests

Several tests were provided to check the quality and correctness of the developed programms.

Comparison of the observation equations with ODIS The ODIS was the first method developed and tested for the first look preprocessing. It used the same data sets as the ring solution. Therefore the proper directions, derivatives with respect to the unknowns, and the differences 'observed minus computed' could directly be compared with that from ODIS.

Structure of the normal matrix Another important test was to plot and interpret the structure of the accumulated pre-reduced normal-equation matrix. Figure 3 shows the structure of that matrix

Moore-Penrose test of the pseudo-inverse of the normal matrix blabla

4.3.3 Ring Solution

5 Results and Documentation

5.1 Validity Check with "true Data"

5.2 Noisy Attitude

5.3 Noisy Calibration

5.4 Realistic Experiment

6 File Utilization in the First Look Preprocessing

6.1 Contents of the Files

References

[1] dummy

[Bastian 2004] Bastian U., 2004, *Reference Systems, Conventions and Notations for GAIA.* GAIA-ARI-BAS-003

[Bastian et al. 2005] Bastian U., Hirte S., Jordan S., Lenhardt H., Bernstein H.-H., 2005, *First-Look Preprocessing.* GAIA-ARI-BAS-008, issue 2

[Bernstein 1993] Bernstein H.-H., 1993, *Global Astrometry of HIPPARCOS Double Stars within the FAST Consortium,* A&A Vol. 283, No. 1

[Biermann and Bastian 2005] Biermann M., Bastian U., 2005, *Task description Gaia First Look.* LSW-MB-001

[Großmann 1975] Großmann W., 1975, *Grundzüge der Ausgleichungsrechnung.* Springer Verlag, Berlin & New York

[Hirte 2004] Hirte S., 2004, *Simulation of the ring solution.* GAIA-ARI-HIR-002

[Höpcke 1980] Höpcke W., 1980, *Fehlerlehre und Ausgleichungsrechnung.* W. de Gruyter & Co., Berlin & New York

[Jordan et al. 2005] Jordan S., Lenhardt H., Bastian U., 2005, *Results from ODIS Experiments and Consequeences for GIS.* GAIA-ARI-SJ-003

[Lawson and Hanson 1974] Lawson C. L., Hanson R. J., 1974,*Solving Least Squares Problems.* Prentice-Hall, Englewood Cliffs

[Lindegren 2001] Lindegren L., 2001, *Global Iterative Solution – Distributed processing of the attitude updating.* SAG-LL-34(v.2)

[Louis 1989] Louis A. K., 1989, *Inverse und schlecht gestellte Probleme.* Teubner Verlag, Stuttgart

[Wolf 1975] Wolf H., 1975, *Ausgleichungsrechnung Formeln zur praktischen Anwendung.* Dümmlers Verlag, Bonn

[EISPACK] EISPACK http://www.netlib.org/eispack/

[ESA-SP] ESA SP 1200 Vol. 3

7 Appendices

List of Figures