Mathematical framework for the
AGIS System Orientation

L. Lindegren


Abstract. This note describes the mathematical framework for determining and correcting the system orientation of the source and attitude data resulting from the Astrometric Global Iterative Solution (AGIS).

1 Introduction

The astrometric measurement principle of Gaia only allows to determine the relative positions of sources, that is, without direct reference to any celestial coordinate system. More precisely, the only quantity that is directly observable with a Gaia-like instrument is the instantaneous angle between incident light rays in the proper frame of the observer, or \( \psi = \arccos(u_1' \cdot u_2) \) if \( u_1 \) and \( u_2 \) are the proper directions to a pair of sources. When such angular measurements are combined with an astrometric model in which the sources are constrained to move with uniform space velocity, the resulting reference frame (to which the positions and proper motions refer) has in practice six degrees of freedom, corresponding to a solid-body rotation with fixed inertial spin.\(^1\) This means that the geometrical problem of simultaneously determining the source and attitude parameters is, in reality, singular – it has a rank defect of 6.

While the solution of rank-deficient problems in general requires special attention to the singularities, for example by adding constraints to avoid numerical instability, no such complication arises here because of the way AGIS works. Basically, a solution is found by iterating between the source updating and the attitude updating.\(^2\) When the sources are updated, the reference frame is in reality set by the (then assumed) attitude; and when the attitude is updated, the frame is set by the (then assumed) source parameters. The end result is that AGIS converges to a solution with both the source and attitude parameters expressed in the same reference frame, which however is to some extent arbitrary and probably mainly determined by the initial source or attitude parameters, depending on the sequence of updating in AGIS.

The intention is however that the final source parameters (positions and proper motions) shall be expressed in a celestial reference frame that represents, as closely as possible, the

\(^{1}\)In principle all six degrees of freedom are fixed by the kinematic constraints and the use of external auxiliary data, although too weakly to be of any practical consequence. For example, an inertially rotating frame is in principle inconsistent with the assumption of uniform space motions, but the resulting constraint is extremely weak because of the small arcs covered even by high-proper motion stars in a few years. Also the frame orientation is in principle determined by the adopted orbit of Gaia, via the aberration and parallax effects, but again this constraint is too weak have any practical significance.

\(^{2}\)The calibration and global updating play no role here because they are to first order independent of the reference frame.
International Celestial Reference System (ICRS). For consistency, it is moreover necessary that the attitude parameters are expressed in exactly the same frame as the source parameters. It is the purpose of the System Orientation task to accomplish this. Briefly, the method is as follows:

1. Identify a subset \((S_1)\) of primary sources with \textit{a priori} known astrometric parameters. This subset will define the orientation of the final celestial reference frame. Typically it will contain the optical counterparts of a few thousand extragalactic objects with accurate VLBI positions in the ICRS.

2. Identify a larger subset \((S_2)\) of primary sources that can be assumed to define a kinematically non-rotating celestial frame. Typically it will contain some \(10^5\) to \(10^6\) quasars and point-like galactic nuclei mainly identified from ground-based surveys and photometric criteria. (Note that these sources in general have no accurately known prior positions in the ICRS, as they have not been observed by VLBI.)

3. Use the AGIS parameters for the sources in \(S_1\) (positions and proper motions) and \(S_2\) (proper motions) to make a robust, weighted least-squares solution of the rotation state of the AGIS sources. The solution estimates the nine parameters \(\omega_x, \omega_y, \omega_z, \epsilon_x, \epsilon_y, \epsilon_z, a_x, a_y, a_z\) defined in Sect. 2 and 5. They refer to some chosen frame reference epoch \(t_1\) on the TCB scale.

4. Apply the rotation correction \((\omega_x, \omega_y, \omega_z, \epsilon_x, \epsilon_y, \epsilon_z)\) to all the source parameters, both primary and secondary sources (Sect. 3).

5. Apply the rotation correction \((\omega_x, \omega_y, \omega_z, \epsilon_x, \epsilon_y, \epsilon_z)\) to the attitude parameters (Sect. 4).

The sub-tasks corresponding to items 1 and 2 above (identification of \(S_1\) and \(S_2\)) are not further discussed here. The other three sub-tasks (items 3, 4 and 5) are described in some detail in the following sections. An Appendix outlines concepts and notational conventions used in this and some other technical notes concerning vectors, reference frames, etc.

2 Relation between the provisional and final Gaia frames

This Section gives an exact model for the relation between the provisional reference frame, obtained through AGIS, and the final (extragalactic) reference frame in which the Gaia catalogue should be published. The model is exact in the sense that it makes no assumption about how close the two frames are.

2.1 Model description

Let the triad \(E = [x_E, y_E, z_E]\) represent the desired, kinematically non-rotating extragalactic reference frame in which we want to express the astrometric parameters of the sources.

---

3Basic formulae for linking an observed reference frame to the extragalactic frame were derived in [7]. Although that paper was concerned with linking a provisional Hipparcos catalogue to the ICRS, much of it applies also in the present case. In particular, the same sign convention is used here to describe the rotational relationship between the two frames. The provisional Hipparcos frame was denoted \(H\) in that paper; here we use \(G\) for the corresponding provisional Gaia frame.
Similarly, let the triad $G = [x_G, y_G, z_G]$ represent the reference frame in which AGIS provides the astrometric parameters and the attitude. As explained in Sect. 1, the astrometric model used for the primary sources in AGIS implies that $G$ rotates with constant angular velocity vector, which we denote $\Omega$. Thus,

$$\dot{G} = \Omega \times G$$

(1)

It is therefore necessary to consider $G$ as a function of time. Performing a Taylor expansion around the adopted frame reference epoch $t_1$, and using that $\dot{\Omega} = 0$, we have

$$G = G_1 + (t - t_1)\Omega \times G_1 + \frac{1}{2}((t - t_1)^2\Omega \times (\Omega \times G_1) + \ldots$$

(2)

where, for brevity, $G_1 = G(t_1)$ and $G$ (without subscript) means $G(t)$. The AGIS frame, as a function of time, is thus completely defined by $G_1$ and $\Omega$, which give the orientation and spin of the AGIS frame at the reference epoch. We must now specify a suitable parameterisation of these entities.

For consistency with [7] and existing software, we define the orientation and spin parameters in the sense of corrections to the AGIS frame. For example, the spin correction to be applied to $G$ is $\omega = -\Omega$. Furthermore, we choose to express this vector with respect to the $G$ frame, so that the spin correction parameters (the same in all $G$ frames!) are

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = G' \omega = -G' \Omega$$

(3)

As for the frame orientation correction, we know from Euler’s rotation theorem that there exists a unit vector $e$ and an angle $0 \leq \epsilon \leq \pi$ such that a rotation of $G_1$ by the angle $\epsilon$ around $e$ gives us $E$. We define the orientation correction parameters to be the components of the vector $e = ee$ in $E$:

$$\begin{bmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \end{bmatrix} = E \epsilon$$

(4)

Note that the components of $e$ are the same in the two frames $E$ and $G_1$ ($G_1 e = E' e$); thus they can be denoted $\epsilon_x, \epsilon_y, \epsilon_z$ without ambiguity.

In analogy with $e$ and $\epsilon$ we introduce the total spin $\omega = (\omega_x^2 + \omega_y^2 + \omega_z^2)^{1/2}$ and the unit vector $o = \omega^{-1}$ with components (in $G$) $o_x = \omega_x \omega^{-1}$, etc.

The relation between $G$ and $E$ is completely specified by the frame reference epoch $t_1$, the spin correction parameters $\omega_x, \omega_y, \omega_z$, and the orientation correction parameters (at epoch $t_1$) $\epsilon_x, \epsilon_y, \epsilon_z$.

### 2.2 Quaternion representation of the relation between the frames

The orientation of the AGIS frame $G$ with respect to $E$ can be represented by a unit quaternion. More precisely, starting with $G$ at some arbitrary $t$ we apply two successive rotations:

1. by the angle $(t - t_1)\omega$ about the unit vector $o$: this gives $G_1$;
2. by the angle $\epsilon$ about the unit vector $e$: this gives $E$. 
Using quaternion multiplication (see Appendix), the total rotation from $G$ to $E$ becomes

$$q_{G\rightarrow E} = \begin{pmatrix} o_x \\ o_y \\ o_z \end{pmatrix} \sin \frac{(t - t_1)\omega}{2}; \cos \frac{(t - t_1)\omega}{2} \begin{pmatrix} e_x \\ e_y \\ e_z \end{pmatrix} \sin \frac{\epsilon}{2}; \cos \frac{\epsilon}{2}$$  \hspace{1cm} (5)

In the following we need this quaternion either for the arbitrary time $t$ (when transforming attitude parameters; Sect. 4) or for the specific instant $t = t_0$, where $t_0$ is the source reference epoch (when transforming source parameters; Sect. 3).

### 2.3 General coordinate transformation between the frames

The algorithms described below involve the transformation of the arbitrary vector $v$ between the $G$ and $E$ frames. The coordinates of $v$ in the two frames are $G'v$ and $E'v$, respectively. The coordinate transformation from $G$ to $E$ is achieved through pre-multiplication by the orthogonal $3 \times 3$ matrix $E'G$, 

$$E'GG' = E'v$$  \hspace{1cm} (6)

since $GG'$ is the unit tensor. The transformation matrix $E'G$ can be calculated from the previously obtained quaternion $q_{G\rightarrow E}$. With $q_x, q_y, q_z, q_w$ denoting the components of the quaternion, that is

$$q_{G\rightarrow E} = \begin{pmatrix} q_x \\ q_y \\ q_z \\ q_w \end{pmatrix}$$  \hspace{1cm} (7)

we have

$$E'G = \begin{bmatrix} q_w^2 + q_x^2 - q_y^2 - q_z^2 & 2(q_x q_y + q_z q_w) & 2(q_x q_z - q_y q_w) \\ 2(q_x q_y - q_z q_w) & q_z^2 + q_y^2 - q_x^2 & 2(q_y q_z + q_x q_w) \\ 2(q_x q_z + q_y q_w) & 2(q_y q_z - q_x q_w) & q_x^2 - q_y^2 - q_z^2 \end{bmatrix}$$  \hspace{1cm} (8)

An alternative, and perhaps better method of transformation is based on quaternion algebra. If, corresponding to the coordinates of $v$ in the two frames, we introduce the quaternions

$$v_E = (E'v; 0) \hspace{1cm} \text{and} \hspace{1cm} v_G = (G'v; 0)$$  \hspace{1cm} (9)

we have

$$v_E = q_{G\rightarrow E}^{-1} v_G q_{G\rightarrow E}$$  \hspace{1cm} (10)

It is readily verified that this expression is equivalent to the matrix multiplication in Eqs. (6)–(8).

### 3 Transformation of the astrometric parameters

This Section gives an exact recipe for transforming the astrometric parameters between the frames.

The barycentric unit vector towards a (primary) source is denoted $r$ and is in general a function of $t$. Actually, only the first-order Taylor expansion of $r$ around the source reference epoch $t_0$ is of interest here:

$$r(t) = r_0 + (t - t_0) \mu + \cdots$$  \hspace{1cm} (11)
where $\mathbf{r}_0 = \mathbf{r}(t_0)$ and $\mathbf{\mu} = \dot{\mathbf{r}}(t_0)$ are the barycentric position and proper motion vectors at epoch $t_0$. (The source reference epoch $t_0$ may be different from the frame reference epoch $t_1$, and could even be different for different sources.) $\mathbf{r}_0$ and $\mathbf{\mu}$ must here be interpreted as physical vectors that are independent of the various coordinate frames.

In the extragalactic frame $\mathcal{E}$, the instantaneous spherical coordinates of the vector $\mathbf{r}(t)$ follow from

$$
\mathcal{E}' \mathbf{r}(t) = \begin{bmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{bmatrix} \tag{12}
$$

Clearly $\alpha$ and $\delta$ are functions of time because of the proper motion. Taking the time derivative and using that $\dot{\mathcal{E}} = 0$, we have

$$
\mathcal{E}' \dot{\mathbf{r}}(t) = \begin{bmatrix} -\sin \delta \sin \alpha \\ \cos \delta \cos \alpha \\ 0 \end{bmatrix} \dot{\alpha} + \begin{bmatrix} -\sin \delta \cos \alpha \\ -\sin \delta \sin \alpha \\ \cos \delta \end{bmatrix} \dot{\delta} \tag{13}
$$

The relevant astrometric parameters in the $\mathcal{E}$ frame at epoch $t_0$ are now defined as

$$
\begin{align*}
\alpha_{0\mathcal{E}} &\equiv \alpha_{\mathcal{E}}(t_0) \\
\delta_{0\mathcal{E}} &\equiv \delta_{\mathcal{E}}(t_0) \\
\mu_{\alpha^*\mathcal{E}} &\equiv \dot{\alpha}_{\mathcal{E}}(t_0) \cos \delta_{\mathcal{E}}(t_0) \\
\mu_{\delta\mathcal{E}} &\equiv \dot{\delta}_{\mathcal{E}}(t_0)
\end{align*} \tag{14}
$$

Introducing the normal matrix at $(\alpha_{0\mathcal{E}}, \delta_{0\mathcal{E}})$,

$$
\begin{bmatrix} p_{0\mathcal{E}} & q_{0\mathcal{E}} & r_{0\mathcal{E}} \end{bmatrix} = \begin{bmatrix} -\sin \alpha_{0\mathcal{E}} & \sin \delta_{0\mathcal{E}} \cos \alpha_{0\mathcal{E}} & \cos \delta_{0\mathcal{E}} \cos \alpha_{0\mathcal{E}} \\ \cos \alpha_{0\mathcal{E}} & -\sin \delta_{0\mathcal{E}} \sin \alpha_{0\mathcal{E}} & \cos \delta_{0\mathcal{E}} \sin \alpha_{0\mathcal{E}} \\ \sin \delta_{0\mathcal{E}} & \cos \delta_{0\mathcal{E}} & 0 \end{bmatrix} \tag{15}
$$

we thus have from (12)–(14)

$$
\mathcal{E}' \mathbf{r}_0 = \mathbf{r}_{0\mathcal{E}} \tag{16}
$$

and

$$
\mathcal{E}' \mathbf{\mu} = p_{0\mathcal{E}} \mu_{\alpha^*\mathcal{E}} + q_{0\mathcal{E}} \mu_{\delta\mathcal{E}} \tag{17}
$$

The corresponding astrometric parameters of the same source in the (rotating) $\mathcal{G}$ frame are denoted $\alpha_{0\mathcal{G}}$, $\delta_{0\mathcal{G}}$, $\mu_{\alpha^*\mathcal{G}}$, and $\mu_{\delta\mathcal{G}}$. They are the spherical coordinates of $\mathbf{r}$ in $\mathcal{G}$ at the source reference epoch $t_0$, and the time derivatives of these spherical coordinates at the same epoch. Introducing the normal matrix at $(\alpha_{0\mathcal{G}}, \delta_{0\mathcal{G}})$,

$$
\begin{bmatrix} p_{0\mathcal{G}} & q_{0\mathcal{G}} & r_{0\mathcal{G}} \end{bmatrix} = \begin{bmatrix} -\sin \alpha_{0\mathcal{G}} & \sin \delta_{0\mathcal{G}} \cos \alpha_{0\mathcal{G}} & \cos \delta_{0\mathcal{G}} \cos \alpha_{0\mathcal{G}} \\ \cos \alpha_{0\mathcal{G}} & -\sin \delta_{0\mathcal{G}} \sin \alpha_{0\mathcal{G}} & \cos \delta_{0\mathcal{G}} \sin \alpha_{0\mathcal{G}} \\ \sin \delta_{0\mathcal{G}} & \cos \delta_{0\mathcal{G}} & 0 \end{bmatrix} \tag{18}
$$

we have

$$
(G' \mathbf{r})_{t=t_0} = G'_0 \mathbf{r}_0 = \mathbf{r}_{0\mathcal{G}} \tag{19}
$$

where $G_0 = G(t_0)$, and

$$
\frac{d}{dt} (G' \mathbf{r})_{t=t_0} = p_{0\mathcal{G}} \mu_{\alpha^*\mathcal{G}} + q_{0\mathcal{G}} \mu_{\delta\mathcal{G}} \tag{20}
$$

\footnote{The remaining two astrometric parameters, the parallax $\pi$ and radial proper motion $\mu_r$, are independent of the reference frame and therefore not considered here.}
Combining Eqs. (16) and (19) we obtain

\[ r_{0E} = E' G_0 r_{0G} \]  

which provides the required transformation \((\alpha_{0G}, \delta_{0G}) \rightarrow (\alpha_{0E}, \delta_{0E})\). After this we may also compute \(p_{0E}\) and \(q_{0E}\) from (15).

In order to transform the proper motion components, we first use (1) to find

\[ \frac{d}{dt} (G' r) = \dot{G}' r + G' \dot{r} = (\Omega \times G') r + G' \dot{r} = G' (\dot{r} - \Omega \times r) \]  

and consequently

\[ \frac{d}{dt} (G' r) \bigg|_{t=t_0} = G'_0 (\mu + \omega \times r_0) \]  

where we substituted \(\Omega = -\omega\). By means of (20) we now find

\[ \mu = G_0 (p_{0G} \mu_{*G} + q_{0G} \mu_{SG}) - \omega \times r_0 = G_0 (p_{0G} \mu_{*G} + q_{0G} \mu_{SG} - (G'_0 \omega) \times r_{0G}) \]  

whereupon (17) gives \(^5\)

\[ p_{0E} \mu_{*E} + q_{0E} \mu_{SE} = E' G_0 (p_{0G} \mu_{*G} + q_{0G} \mu_{SG} - (G'_0 \omega) \times r_{0G}) \]  

The proper motion components \((\mu_{*E}, \mu_{SE})\) are found by taking the scalar product of (25) with respectively \(p_{0E}\) and \(q_{0E}\), since \(p_{0E} q_{0E} = 0\) and \(p_{0E} p_{0E} = q_{0E} q_{0E} = 1\).

Both (21) and (25) involve the coordinate transformation from \(G_0\) to \(E\) through a pre-multiplication with the orthogonal matrix \(E' G_0\). As explained in Sect. 2.3 this transformation can alternatively be carried out by quaternion algebra according to (10).

### 4 Transformation of the attitude parameters

The attitude describes the orientation of the Scanning Reference System (SRS) \(Z = [x\ y\ z]\), as a function of time, with respect to the Centre-of-Masses Reference System (CoMRS). For the present purpose we may consider the latter to be a local variant of \(G\) (for the attitude determined in AGIS) or \(E\) (for the desired attitude), see [2]. The AGIS attitude is thus given by the quaternion \(q_{G \rightarrow Z}(t)\) specifying how to turn \(G(t)\) into \(Z(t)\). Similarly, the desired (extragalactic) attitude is given by the quaternion \(q_{E \rightarrow Z}(t)\) specifying how to turn \(E\) into \(Z(t)\). Their relation immediately follows from (5), viz.:

\[ q_{E \rightarrow Z}(t) = q_{E \rightarrow G}(t) \ q_{G \rightarrow Z}(t) = q_{G \rightarrow Z}(t)^{-1} \ q_{G \rightarrow E}(t) \ q_{G \rightarrow Z}(t) \]  

which gives a recipe for correcting the AGIS attitude at the arbitrary time \(t\) for the difference in frame orientation.

In practice the components of the attitude quaternion are represented by splines, which are linear combinations of B-splines. Thus we may write

\[ q_{G \rightarrow Z}(t) = \sum_n c_n^{(G)} B_n(t) \]  

\(^5\)Since \(G' \omega\) is independent of \(t\), we drop the subscript in \(G_0 \omega\), cf. Eq. (3).
where $c_n^{(G)}$ are quaternions whose components consist of the attitude parameters obtained in AGIS and $B_n(t)$ are the B-spline functions on the given knot sequence $\{\tau_n\}$ (cf. [6]). The superscript $(G)$ indicates that the quaternions refer to the G frame. It follows that the attitude in the E frame is

$$q_{E-Z}(t) = \sum_n q_{G-E}^{-1}(t) c_n^{(G)} B_n(t) \quad (28)$$

This equation is strictly correct, but not very useful since the coefficients of the B-splines, $q_{G-E}^{-1}(t) c_n^{(G)}$, are now themselves functions of $t$. However, we are here helped by a property of B-splines, namely that their support is limited to $M$ consecutive knot intervals: $B_n(t) = 0$ except for $\tau_n < t < \tau_n+M$, where $M$ is the order of the spline [6]. Thus, if the variation of $q_{G-E}^{-1}(t)$ is small within that interval, we may just as well replace it with $q_{G-E}^{-1}(\bar{t}_n)$, where $\bar{t}_n$ is some intermediate time, for example

$$\bar{t}_n = \frac{1}{2}(\tau_n + \tau_n+M) \quad (29)$$

Then

$$q_{E-Z}(t) \simeq \sum_n c_n^{(E)} B_n(t) \quad (30)$$

where

$$c_n^{(E)} = q_{G-E}(\bar{t}_n) c_n^{(G)} \quad (31)$$

are the transformed attitude parameters.

The error of the approximation in (30) is at most $\frac{1}{2}(\tau_n+M - \tau_n)\omega \sim 1$ µas if $\omega = 1$ arc-sec yr$^{-1}$, $\Delta t = 15$ s and $M = 4$ (cubic spline). Since, realistically, $\omega \ll 1$ arcsec yr$^{-1}$ the error is in practice completely negligible. Thus (31) provides a very simple and accurate recipe for correcting the AGIS attitude.

5 Determination of the frame orientation/rotation parameters

The determination of $E'\epsilon = [\epsilon_x \epsilon_y \epsilon_z]'$ and $G'\omega = [\omega_x \omega_y \omega_z]'$, together with the acceleration parameters $a_x, a_y, a_z$, is done by a least-squares estimation of the astrometric parameters of subsets $S_1$ and $S_2$ in the two frames. The least-squares estimator should take into account the non-linear relationship between the source parameters in the two frames (as detailed in Sect. 3), allow consistent treatment of the different epochs involved, consider the heteroscedasticity of the data, and make allowance for possible outliers among the data. In the following we presume the availability of a general (non-linear, robust) least-squares method that solves the following problem:

Given a method to compute, for each ‘observation’ $i$,

- the residual $R_i(\theta)$,
- the associated standard error $\sigma_i$,
- the partial derivatives $\partial R_i/\partial \theta_j$,

where $\theta$ is an $n$-dimensional parameter array,

find the best estimate $\hat{\theta}$, and the associated covariance matrix $C$.

It is assumed that the number of observations $m$ is (much) larger than $n$, that the errors are uncorrelated among the observations, and that the problem is non-singular.
With such a method in mind, the parameter array is \( \theta = [x, y, z, \omega_x, \omega_y, \omega_z, a_x, a_y, a_z] \)' and it remains to define what the ‘observations’ are, and how the residuals and their partial derivatives can be computed for arbitrary \( \theta \).

Up until now, we have not introduced any approximations based on the assumption that \( \epsilon \) and \( \omega \) are small quantities. In the following, we will continue to use rigorous formulae for computing the residuals (\( R_i \)), but adopt the small-angle approximation for computing the partial derivatives (\( \partial R_i / \partial \theta_j \)). Thus quantities of second order in \( \epsilon \), \( \omega \), \( a \) and \( \mu \) (and their cross terms) are ignored in the derivatives.

### 5.1 Proper motion of an extragalactic source

Let us first consider the processing of the proper motion data from Subset \( S_2 \) (Sect. 1). In the \( E \) frame, the proper motions of these sources are statistically zero except for the apparent streaming-like motion caused by the cosmological acceleration \( g \) of the solar system barycentre [1]. To first order, the apparent proper motion of a source at position \( r \) (unit vector) is

\[
\mu = (U - rr')g c^{-1}
\]  
(32)

where \( c \) is the speed of light and \( U - rr' \) is the tensor effecting a projection onto the plane of the sky, normal to \( r \). The expected value of \( |g| \) due to the galactic potential is \((220 \text{ km s}^{-1})^2/(8 \text{ kpc}) \approx 2 \times 10^{-10} \text{ m s}^{-2} \) resulting in a maximum apparent proper motion of \( |g|c^{-1} \approx 4.3 \mu \text{as yr}^{-1} \). From the size of the effect it is clear that higher-order terms in the relativistic aberration formula need not be considered. We introduce \( a = gc^{-1} \), which is then expressed in the same unit as \( \omega \) (e.g., rad yr\(^{-1} \)). The parameters \( a_x, a_y, a_z \) should be interpreted as the components of \( a \) in \( E \), or \( E' a = [a_x a_y a_z]' \).

Taking the \( r \) in (32) to be the barycentric coordinate direction \( r_0 \) at the source reference epoch, the expected proper motion components in the \( E \) frame are

\[
\begin{align*}
\mu_{\alpha E} &= p'_{0E}(E'a) \\
\mu_{\delta E} &= q'_{0E}(E'a)
\end{align*}
\]

(33)

These values are to be compared with the ‘observed’ values from AGIS, obtained after rigorous transformation to the \( E \) frame according to (25). The residuals are

\[
\begin{align*}
R_{\mu \alpha} &= p'_{0E}E'G_0 \left( p_{0G} \mu_{\alpha \gamma} + q_{0G} \mu_{\delta \gamma} - (G' \omega \times r_{0G}) \right) - p'_{0E}(E'a) \\
R_{\mu \delta} &= q'_{0E}E'G_0 \left( p_{0G} \mu_{\alpha \gamma} + q_{0G} \mu_{\delta \gamma} - (G' \omega \times r_{0G}) \right) - q'_{0E}(E'a)
\end{align*}
\]

(34)

In the small-angle approximation the partial derivatives are found to be

\[
\begin{align*}
\frac{\partial R_{\mu \alpha}}{\partial (E' \epsilon)} &= 0, & \frac{\partial R_{\mu \alpha}}{\partial (G' \omega)} &= -q_{0E}, & \frac{\partial R_{\mu \alpha}}{\partial (E'a)} &= -p_{0E} \\
\frac{\partial R_{\mu \delta}}{\partial (E' \epsilon)} &= 0, & \frac{\partial R_{\mu \delta}}{\partial (G' \omega)} &= p_{0E}, & \frac{\partial R_{\mu \delta}}{\partial (E'a)} &= -q_{0E}
\end{align*}
\]

(35)

### 5.2 Position of a source known in the ICRS

For a source in Subset \( S_1 \) we have a position measured (by VLBI) directly in the \( E \) frame, but referring to some different epoch of observation, say \( t_2 \). Again, the sources are assumed to have no proper motion in \( E \), except for the acceleration effect as in (32). Thus
we need to take into account the epoch difference between the VLBI observations \((t_2)\) and the Gaia observations \((t_0)\). Let \(r_{2E}^{(ICRS)}\) be the array of direction cosines calculated from the ICRS data (referring to \(t_2\)). This should be compared with the AGIS position transformed to the \(E\) frame according to (21), after correcting for the apparent proper motion in the intervening time interval. Resolving the difference in its components along \(\alpha\) and \(\delta\) we obtain the position residuals

\[
\begin{align*}
R_\alpha &= p'_{0E} \left( E'G_0 r_{0G} - r_{2E}^{(ICRS)} \right) - p'_0E(\alpha)(t_0 - t_2) \\
R_\delta &= q'_{0E} \left( E'G_0 r_{0G} - r_{2E}^{(ICRS)} \right) - q'_0E(\alpha)(t_0 - t_2)
\end{align*}
\]

with partial derivatives

\[
\begin{align*}
\frac{\partial R_\alpha}{\partial (E'\epsilon)} &= -q_{0E}, & \frac{\partial R_\alpha}{\partial (G'\omega)} &= -q_{0E}(t_0 - t_1), & \frac{\partial R_\alpha}{\partial (E'a)} &= -p_{0E}(t_0 - t_2) \\
\frac{\partial R_\delta}{\partial (E'\epsilon)} &= p_{0E}, & \frac{\partial R_\delta}{\partial (G'\omega)} &= p_{0E}(t_0 - t_1), & \frac{\partial R_\delta}{\partial (E'a)} &= -q_{0E}(t_0 - t_2)
\end{align*}
\]

In (37) we have made use of the small-angle approximation

\[
E'G_0 \simeq \begin{bmatrix} 1 & +\epsilon_z & -\epsilon_y \\ -\epsilon_z & 1 & +\epsilon_x \\ +\epsilon_y & -\epsilon_x & 1 \end{bmatrix} + \begin{bmatrix} 0 & +\omega_z & -\omega_y \\ -\omega_z & 0 & +\omega_x \\ +\omega_y & -\omega_x & 1 \end{bmatrix} (t_0 - t_1)
\]

The proper motion data for the sources in \(S_1\) are treated as in Sect. 5.1.
Appendix: Vectors, triads, matrices and quaternions

In the following we work exclusively in Euclidean space. This is possible when considering entities (directions, angles, orientation, rotation) in the local rest frame of Gaia (the Centre-of-Mass Reference System, CoMRS) or the co-rotating frame (the Scanning Reference System, SRS), since all relativistic effects have been taken care of when transforming to the CoMRS.

A.1. Vectors

We think of vectors as physical entities that have direction, magnitude (length), and sometimes an origin, but which exist independently of any reference frames or coordinate systems. Examples are positions (displacements), velocities, accelerations, forces, angular velocities, angular momenta, and directions (usually represented by vectors of unit length). We use (lower-case) boldface italics to denote vectors, e.g. \( \mathbf{r}, \mathbf{v} \). Operations such as vector addition and substraction, multiplication with a scalar, scalar (inner) product and vector (cross) product have their usual meaning independent of coordinate systems.

Actually, such vectors can be regarded as a special case of a tensor (of rank 1); in this context we can also think of a scalar (e.g. \( s = r'v' \)) as a tensor of rank 0, and the outer product of two vectors (e.g. \( \mathbf{T} = \mathbf{r} \mathbf{v}' \)) as a tensor of rank 2. All of them represent physical entities independent of coordinate systems.

A.2. Triads and coordinates

A coordinate frame can be represented by an orthogonal, right-handed triad of unit vectors, e.g. \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) (with the same starting point) satisfying

\[
\begin{aligned}
i'i = j'j = k'k &= 1, \\
i'j' &= j'k' = k'i = 0, \\
i \times j &= k, \\
j \times k &= i, \\
k \times i &= j
\end{aligned}
\]  

(39)

Following [8], upper-case sans-serif letters are used to denote triads, as in

\[
\mathbf{K} = [i \ j \ k]
\]  

(40)

Formally, the triad is therefore a (1,3)-matrix of vectors. The coordinates of the arbitrary vector \( \mathbf{r} \) in frame \( \mathbf{K} \) are the elements of the (3,1)-matrix

\[
\begin{bmatrix}
\mathbf{r}_i \\
\mathbf{r}_j \\
\mathbf{r}_k
\end{bmatrix} = \begin{bmatrix}
i'r \\
j'r \\
k'r
\end{bmatrix} = \begin{bmatrix}
i' \\
j' \\
k'
\end{bmatrix} \mathbf{r} = [i \ j \ k]' \mathbf{r} = \mathbf{K}' \mathbf{r}
\]  

(41)

where the prime (') is used both for the scalar product of vectors and the transpose of a matrix. Conversely, the vector can be written in terms of its coordinates as

\[
\mathbf{r} = i\mathbf{r}_i + j\mathbf{r}_j + k\mathbf{r}_k = [i \ j \ k] \begin{bmatrix}
\mathbf{r}_i \\
\mathbf{r}_j \\
\mathbf{r}_k
\end{bmatrix} = \mathbf{K} \begin{bmatrix}
\mathbf{r}_i \\
\mathbf{r}_j \\
\mathbf{r}_k
\end{bmatrix}
\]  

(42)

The notations introduced above allow a simple and unambiguous way to distinguish between a vector (\( \mathbf{r} \)) and its coordinate representation in a given frame (\( \mathbf{K}' \mathbf{r} \)). Sometimes it is however convenient to have a separate (one-character) symbol for the coordinate representation, which is a (3,1)-matrix. The boldface roman (upright) font may be used for this purpose, as in \( \mathbf{r} = \mathbf{K}' \mathbf{r} \).
More generally, we use here boldface roman letters for matrices, with upper-case letters for two-dimensional matrices and lower-case letters for one-dimensional matrices or arrays. Within the context of Euclidean space there is a clear correspondence between vectors/tensors and their coordinate representations in one- and two-dimensional matrices. This system of notation is further explored in Table 1 at the end of this document.\(^6\)

**A.3. Rotations**

We use the term *rotation* in the usual geometrical sense, namely as a rigid-body movement in three-dimensional space that keeps one point fixed. (The fixed point is here always considered to be the origin of the coordinate frame.)

A perpetual source of confusion when dealing with rotations is to specify unambiguously what is actually rotating. Consider a rigid body that is being rotated about a fixed point at the origin, and the previously defined vectors \(i, j, k\), and \(r\). We may distinguish two cases:

1. The coordinate axes \(i, j, k\) are fixed in space, while the vector \(r\) is attached to the rotated body.

2. The coordinate axes \(i, j, k\) are attached to the rotated body, while \(r\) is a vector fixed in space.

In both cases the rotation causes a change in the coordinates \(r_i, r_j, r_j\). However, a given physical rotation of the body will have different numerical representations in the two cases. Alternatively, we could say that a given change in the coordinates of \(r\) corresponds to different physical rotations in the two cases.

Linguistically, we may distinguish between the two cases by using the name *vector rotation* in the first case, and *frame rotation* in the second.

To further illustrate the point of confusion, consider the definition of elementary rotation matrices. A rotation by the angle \(\psi\) about the 3rd spatial axis (sometimes referred to as the \(z\) axis) may be expressed by the matrix

\[
\mathbf{R}_3(\psi) = \begin{bmatrix}
\cos \psi & \sin \psi & 0 \\
-\sin \psi & \cos \psi & 0 \\
0 & 0 & 1
\end{bmatrix}
\]  

(43)

When this matrix is applied to \(r_i, r_j, r_k\) we get a new set of coordinates:

\[
\begin{bmatrix}
\bar{r}_i \\
\bar{r}_j \\
\bar{r}_k
\end{bmatrix} = \mathbf{R}_3(\psi) \begin{bmatrix}
r_i \\
r_j \\
r_k
\end{bmatrix}
\]  

(44)

But what has actually been rotated? The answer – in this case – is that it is the coordinate frame that has been rotated by the angle \(\psi\), if we insist that the rotation is defined in the positive sense about the 3rd axis. Equivalently, we may of course say that it is the vector \(r\) that has been rotated, but then it would have to be by the angle \(-\psi\). The confusion is

\(^6\)Although this system of notation has some advantages, it would not be meaningful to try to enforce it rigorously, due to its restricted context and because other applications have different needs. Moreover, its implementation in LaTeX is not without difficulty.
further increased by some authors using the opposite convention, i.e., defining the rotation matrix to be the transpose of (43), see for example [9]. To be precise, we should therefore say that (43) is a recipe for frame rotation by the angle \( \psi \) about the 3rd axis.\(^7\)

Loosely speaking, much of the confusion derives from the notational difficulty to separate a vector (as a physical entity) from its coordinates (i.e., the numerical representation of the same vector in a particular coordinate frame). This problem can be overcome, at the expense of some notational overhead, by consistently using the conventions introduced above for the notation of vectors, frames (triads), and coordinates.

The usefulness of this notation becomes apparent when another coordinate frame is introduced. Let \( Z = [x \ y \ z] \) be the triad of orthogonal unit vectors defining the new frame. The coordinates of \( r \) in the new frame are

\[
\begin{bmatrix}
  r_x \\
  r_y \\
  r_z
\end{bmatrix} = Z'r
\]

If we know the coordinates of \( r \) in the original frame \( K \), we can use (44) to obtain

\[
Z'r = Z'K'K'r = (Z'K)(K'r)
\]

where parentheses have been introduced to help the interpretation of the different factors. Equation (46) is in fact just a short-hand notation for the standard matrix formulation of the coordinate transformation:

\[
\begin{bmatrix}
  r_x \\
  r_y \\
  r_z
\end{bmatrix} = A
\begin{bmatrix}
  r_i \\
  r_j \\
  r_k
\end{bmatrix}
\]

where

\[
A = Z'K = [x \ y \ z]'[i \ j \ k] = \begin{bmatrix}
  x' \\
  y' \\
  z'
\end{bmatrix}[i \ j \ k] = \begin{bmatrix}
  x'\ i & x'\ j & x'\ k \\
  y'\ i & y'\ j & y'\ k \\
  z'\ i & z'\ j & z'\ k
\end{bmatrix}
\]

is the (3,3) transformation matrix. The inverse transformation is

\[
K'r = K'ZZ'r = (K'Z)(Z'r)
\]

with transformation matrix \( K'Z = (ZK)' = A' \). (\( A \) is an orthogonal matrix; therefore \( A^{-1} = A' \).) The relation between the two frames can also be written

\[
K = ZA \quad \text{or} \quad Z = KA'
\]

Note that Eqs. (46)–(50) express what we previously called frame rotation, i.e., the transformation of one frame into another (as in Eq. 50), or between the coordinate representations of a fixed vector \( r \) in the two frames (as in Eqs. 46 and 49).

We may of course also use an orthogonal matrix like \( A \) to express a vector rotation, as in

\[
K'r = AK'r
\]

In this case we stay in the same frame (K), and rotate the vector (from \( r \) to \( \tilde{r} \)). Numerically, (47) and (51) involve exactly the same calculations, but the results have quite different physical interpretations.

\(^7\)An implicit assumption is that coordinate triplets are represented by column matrices, as in (44).
A.4. Quaternions

Any quaternion \( q \) is represented by four real numbers which we denote \( q_x, q_y, q_z \) and \( q_w \). The first three numbers may be referred to as the vector part of the quaternion, while \( q_w \) is the scalar part. An alternative notation is

\[
q = \begin{pmatrix}
q_x \\
q_y \\
q_z \\
q_w
\end{pmatrix}
\]

(52)

Note that we deliberately avoid writing \( q \) as a \((4,1)\)-matrix (or 4-dimensional vector), since that would make it necessary to agree on the order of the vector and scalar parts – and this might lead to confusion as different conventions exist in the literature.\(^8\)

The quaternion conjugate is obtained by reversing the sign of the vector part,

\[
q^* = \begin{pmatrix}
-q_x \\
-q_y \\
-q_z \\
q_w
\end{pmatrix}
\]

(53)

The product of two quaternions \( c = ab \) is defined in terms of its components as

\[
\begin{align*}
    c_x &= a_x b_w + a_y b_z - a_z b_y + a_w b_x \\
    c_y &= -a_x b_z + a_y b_w + a_z b_x + a_w b_y \\
    c_z &= a_x b_y - a_y b_x + a_z b_w + a_w b_z \\
    c_w &= -a_x b_x - a_y b_y - a_z b_z + a_w b_w
\end{align*}
\]

(54)

We have

\[
q^* q = qq^* = \begin{pmatrix}
0 \\
0 \\
0 \\
q_x^2 + q_y^2 + q_z^2 + q_w^2
\end{pmatrix}
\]

(55)

which is purely scalar; we may therefore define the absolute value (or length) of the quaternion as

\[
|q| = \sqrt{q^* q} = \sqrt{qq^*} = \sqrt{q_x^2 + q_y^2 + q_z^2 + q_w^2}
\]

(56)

For a non-zero quaternion the inverse may be computed as

\[
q^{-1} = q^* / (q^* q)
\]

(57)

Unit quaternions, having \(|q| = 1\), are useful for representing rotations. In a given reference frame \( Z = [x \ y \ z] \), rotation about the unit vector \( u \) by the angle \( \phi \) is represented by the unit quaternion

\[
q = \left( Z' u \sin \frac{\phi}{2}; \cos \frac{\phi}{2} \right)
\]

(58)

In the spirit of this Appendix it would be nice if we could distinguish quaternions from vectors and matrices by some special font, but unfortunately we have already exhausted the standard options in LaTeX. Given the choice between roman and italic boldface, we choose to use the same as for matrices, namely roman boldface, since a quaternion is after all a collection of numbers, like a matrix.
In the given reference frame \((Z)\), the arbitrary vector \(v\) may be represented as a quaternion with zero scalar part,

\[
v = (Z'v; 0)
\]  

(59)

Vector rotation, corresponding to the \(q\) in (58), is then accomplished by two quaternion multiplications:

\[
(Z'\tilde{v}; 0) = qvq^{-1}
\]  

(60)

Here, \(\tilde{v}\) is the vector obtained by rotating \(v\) the angle \(\phi\) about the unit vector \(u\).

Frame rotation is similarly accomplished by two quaternion multiplications. Let the \(q\) in (58) be the rotation that brings the original frame \(Z\) into coincidence with the new frame \(K\). (Note that \(K'u = Z'u\), so there is no ambiguity as to the coordinates of \(u\) to be used in the quaternion.) Then for the arbitrary fixed vector \(v\),

\[
(K'v; 0) = q^{-1}vq
\]  

(61)

Note the difference between (60) and (61), which parallels the difference between (51) and (46) in the previous section.

References


Acronyms

The following table has been generated from the on-line Gaia acronym list:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGIS</td>
<td>Astrometric Global Iterative Solution</td>
</tr>
<tr>
<td>CoMRS</td>
<td>Centre of Mass Reference System</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>GIS</td>
<td>(Astrometric) Global Iterative Solution</td>
</tr>
<tr>
<td>ICRS</td>
<td>International Celestial Reference System</td>
</tr>
<tr>
<td>SIM</td>
<td>Space Interferometry Mission</td>
</tr>
<tr>
<td>SRS</td>
<td>Scanning Reference System</td>
</tr>
<tr>
<td>TCB</td>
<td>Barycentric Coordinate Time</td>
</tr>
<tr>
<td>VLBI</td>
<td>Very Long Baseline Interferometry</td>
</tr>
</tbody>
</table>
Table 1: The tensor–matrix isomorphism and corresponding notations.

<table>
<thead>
<tr>
<th>Tensor (name, notation)</th>
<th>Matrix (name, notation)</th>
<th>Tensor–matrix relations using $Z = [x \ y \ z]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>scalar (tensor of rank 0) $s$</td>
<td>(1,1)-matrix $s = s$</td>
<td>$s = s$</td>
</tr>
<tr>
<td>vector (tensor of rank 1) $v$</td>
<td>(3,1)-matrix $v = \begin{bmatrix} v_x \ v_y \ v_z \end{bmatrix}$</td>
<td>$v = Z'v$ $v = Zv$</td>
</tr>
<tr>
<td>covector (tensor of rank 1) $v'$</td>
<td>(1,3)-matrix $v' = [v_x \ v_y \ v_z]$</td>
<td>$v' = v'Z$ $v' = v'Z'$</td>
</tr>
<tr>
<td>tensor (of rank 2) $T$</td>
<td>(3,3)-matrix $T = \begin{bmatrix} T_{xx} &amp; T_{xy} &amp; T_{xz} \ T_{yx} &amp; T_{yy} &amp; T_{yz} \ T_{zx} &amp; T_{zy} &amp; T_{zz} \end{bmatrix}$</td>
<td>$T = Z'TZ$ $T = ZTZ'$</td>
</tr>
<tr>
<td>special case: unit tensor (of rank 2) $U$</td>
<td>identity matrix $U = \begin{bmatrix} 1 &amp; 0 &amp; 0 \ 0 &amp; 1 &amp; 0 \ 0 &amp; 0 &amp; 1 \end{bmatrix}$</td>
<td>$U = Z'Z$ $U = ZZ'$</td>
</tr>
<tr>
<td>vector addition (parallelotram) $u \pm v$</td>
<td>matrix addition $u \pm v = \begin{bmatrix} u_x \pm v_x \ u_y \pm v_y \ u_z \pm v_z \end{bmatrix}$</td>
<td>$u \pm v = (Z'u) \pm (Z'v) = Z'(u \pm v)$</td>
</tr>
<tr>
<td>vector–scalar multiplication $vs$</td>
<td>matrix multiplication (3,1)(1,1)=(3,1) $vs$</td>
<td>$vs = Z'vs$</td>
</tr>
<tr>
<td>vector length $</td>
<td>v</td>
<td>$</td>
</tr>
<tr>
<td>scalar product $u'v = \frac{u \cdot v}{</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>cross product $u \times v = \hat{n}</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>outer product $uv'$</td>
<td>matrix multiplication (3,1)(3,1)=(3,3) $uv' = \begin{bmatrix} u_xv_x &amp; u_xv_y &amp; u_xv_z \ u_yv_x &amp; u_yv_y &amp; u_yv_z \ u_zv_x &amp; u_zv_y &amp; u_zv_z \end{bmatrix}$</td>
<td>$uv' = Z'uv'Z$</td>
</tr>
<tr>
<td>tensor–vector multiplication $Tv$</td>
<td>matrix multiplication (3,3)(1,1)=(3,1) $Tv = \begin{bmatrix} T_{xx}v_x + T_{xy}v_y + T_{xz}v_z \ T_{yx}v_x + T_{yy}v_y + T_{yz}v_z \ T_{zx}v_x + T_{zy}v_y + T_{zz}v_z \end{bmatrix}$</td>
<td>$Tv = Z'TZZ'v = Z'Tv$</td>
</tr>
<tr>
<td>special case: $Uv = v$</td>
<td>matrix multiplication (3,3)(3,1)=(3,1) $Uv = v$</td>
<td>$Uv = (Z'Z)(Z'u) = Z'(Z'Z)v = Z'Uv$</td>
</tr>
</tbody>
</table>

* $\hat{n}$ is a unit vector normal to $u$ and $v$, such that $(u, v, \hat{n})$ is right-handed.