# Attitude parameterization for GAIA

L. Lindegren (1 July 2000)

SAG–LL–30

ABSTRACT. The GAIA attaitude may be described by four continuous functions of time, $q_1(t)$, $q_2(t)$, $q_3(t)$, $q_4(t)$, which form a quaternion of approximately unit length. The quaternion formalism is elegant and computationally efficient, reducing the use of trigonometric functions to a minimum. As examples of the use of quaternions, I discuss the nominal scanning law, the inertial rotation of the instrument, and the determination of the instantaneous attitude from direction measurements.

## 1 Introduction

The nominal scanning law (NSL) proposed for GAIA is descibed in SAG–LL–014, 014A and 026. The real attitude of GAIA should follow the NSL at least to within a few arcmin. For Hipparcos, the attitude was described differentially relative the NSL by means of three Euler angles. This allowed to use the minimum possible number of attitude parameters (i.e., 3) while avoiding singularities, and allowed the real-time control to be formulated linearly in the (small) Euler angles. However, for the accurate *a posteriori* attitude determination it lead to rather complicated relations e.g. between the object directions in the reference (celestial) frame and instrument frame, and in the accurate physical modelling of the satellite motion. For GAIA, it will be advantageous to adopt an attitude parameterization which directly relates the reference and instrument directions, and I propose that quaternions are ideal for this.

## 2 Definition of attitude

Let the triad $\mathbf{N} = [\mathbf{l}\ \mathbf{m}\ \mathbf{n}]$ represent the celestrial reference frame (ICRF). That is, $\mathbf{l}$ is a unit vector towards $(\alpha, \delta) = (0, 0)$, $\mathbf{n}$ a unit vector towards $\delta = 90^\circ$, and $\mathbf{m} = \mathbf{n} \times \mathbf{l}$ to complete the right-handed triad. Similarly, let $\mathbf{Z} = [\mathbf{x}\ \mathbf{y}\ \mathbf{z}]$ represent the instrument system, with $\mathbf{x}$ bisecting the two astrometric viewing directions, $\mathbf{z}$ along the nominal spin axis, and $\mathbf{y} = \mathbf{z} \times \mathbf{x}$. At a give time $t$, let $\mathbf{u}$ be the proper direction to a celestial object (i.e., the 'observable' satellitocentric direction including gravitational deflection and aberration). Given astrometric data and the barycentric position and velocity of the satellite, the components of $\mathbf{u}$ in the reference frame are known, i.e. the $3 \times 1$ matrix $\mathbf{N}'\mathbf{u} = [u_l\ u_m\ u_n]'$. By observing the object, we measure (at least in principle and for a fully calibrated instrument) its components in the instrument frame, i.e. the $3 \times 1$ matrix

$\mathbf{Z}'\mathbf{u} = [u_x\ u_y\ u_z]'$. The attitude at time $t$ is the mapping from $\mathbf{N}'\mathbf{u}$ to $\mathbf{Z}'\mathbf{u}$ for the arbitrary direction $\mathbf{u}$. The mapping is most directly specified by the $3 \times 3$ *attitude matrix* $\mathbf{A}$:

$$\begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = \mathbf{A} \begin{bmatrix} u_l \\ u_m \\ u_n \end{bmatrix} . \tag{1}$$

Equivalently, we may define $\mathbf{A} = \mathbf{Z}'\mathbf{N}$. The attitude matrix is a real, proper orthogonal matrix ($\mathbf{A}\mathbf{A}' = \mathbf{I}$, $|\mathbf{A}| = +1$). The inverse relation is therefore

$$\begin{bmatrix} u_l \\ u_m \\ u_n \end{bmatrix} = \mathbf{A}' \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} . \tag{2}$$

## 3 Attitude parameterization

The attitude of GAIA could be represented by the continuous function $\mathbf{A}(t)$, or rather by the nine continuous functions $A_{ij}(t)$, $i, j = 1, 2, 3$. However, this is not very convenient, because of the strong redundancy implied by the orthogonality (the nine functions would have to satisfy six orthogonality constraints). Alternatively, the attitude could be described by just three continuous functions, e.g. the Euler angles $\phi(t)$, $\theta(t)$, $\psi(t)$ with respect to $\mathbf{N}$, defined in some suitable sequence (there are 12 different ways to do this). This eliminates the redundancy problem but inevitably leads to singularity problems for certain values of $\theta$. A third possibility, which (almost) eliminates both problems, is to use the quaternion representation $\mathbf{q}(t)$, also known as the Euler symmetric parameters. This formalism has been used extensively for spacecraft control and attitude determination, and is also well-known in celestial mechanics. However, it is relatively unknown to astronomers in general, and a brief summary is therefore provided below, based on J.R. Wertz [ed.], *Spacecraft attitude determination and control*, ASSL Vol. 73, D. Reidel, 1978. (A slightly different formalism is described by Jefferys 1987, AJ 93, 755.)

## 4 Quaternion representation

Quaternions can be regarded as four-vectors, for which special rules of multiplication apply. Specifically, let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} , \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{3}$$

be two quaternions; then the product $\mathbf{c} = \mathbf{ab}$ is a quaternion with elements

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} b_4 & b_3 & -b_2 & b_1 \\ -b_3 & b_4 & b_1 & b_2 \\ b_2 & -b_1 & b_4 & b_3 \\ -b_1 & -b_2 & -b_3 & b_4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} . \tag{4}$$

Note that the order of multiplication is important, and that it is different in Eq. (4) than in $\mathbf{c} = \mathbf{ab}$. The length of a quaternion is defined as $|\mathbf{a}| = (a_1^2 + a_2^2 + a_3^2 + a_4^2)^{1/2}$.

The instantaneous attitude may be represented by the quaternion

$$\mathbf{q} = \begin{bmatrix} e_1 \sin(\Phi/2) \\ e_2 \sin(\Phi/2) \\ e_3 \sin(\Phi/2) \\ \cos(\Phi/2) \end{bmatrix} , \tag{5}$$

where $\mathbf{e}$ is a unit vector and $\Phi$ an angle. The physical interpretation is that a frame, originally aligned with $\mathbf{N}$, becomes aligned with $\mathbf{Z}$ after a rotation by the angle $\Phi$ about the axis $\mathbf{e}$. $e_1$, $e_2$ and $e_3$ are the components of $\mathbf{e}$ in the original frame. Since $\mathbf{e}$ is a unit vector, it follows that $|\mathbf{q}| = 1$. Such a unimodular quaternion is also called a *spinor*.

Because of the constraint $|\mathbf{q}| = 1$ there is one degree of redundancy in the quaternion representation of the attitude. A simple way to deal with this is suggested below. On the other hand, the quaternion representation has several advantages, e.g.:

- There are no singularities. The four functions $q_i(t)$ are perfectly well-behaved continuous functions of time.
- Successive rotations are represented by multiplications (from left to right) according to the rule (4). This involves fewer arithmetic operations than the corresponding multiplication of rotation matrices.
- The attitude matrix $\mathbf{A}$ is easily computed from $\mathbf{q}$, without trigonometric functions [see Eq. (7)].
- There are simple relations between the time derivative $\dot{\mathbf{q}}$ and the intertial rotation vector [see Eqs. (8)–(9)]; again these involve no trigonometric functions.

There is a sign ambiguity in the definition of $\mathbf{q}$, in that a rotation by $-\Phi$ about $-\mathbf{e}$ is equivalent to a rotation by $\Phi$ about $\mathbf{e}$. To avoid discontinuities in $q_i(t)$ it is necessary to choose the correct sign. This happens automatically if the attitude quaternion remains close to the NSL quaternion defined in the next section.

The attitude $\mathbf{q}$ at a given instant can be estimated from a set of measurements of known celestial directions in the instrument frame. In this process it is easy to incorporate the constraint $|\mathbf{q}| = 1$ by standard methods (e.g. using a Lagrangian multiplier; see Sect. 8). In practice, the functions $q_i(t)$ need to be represented by suitable basis functions (polynomials, trigonometric functions or splines) and it is then the coefficients of these basis functions that have to be estimated. In such a formulation it is difficult guarantee that $|\mathbf{q}(t)| = 1$ for every $t$. The proposed solution is that (small) deviations from this condition are permitted, by formulating all transformations to take into account that $\mathbf{q}$ is not necessarily of unit length.

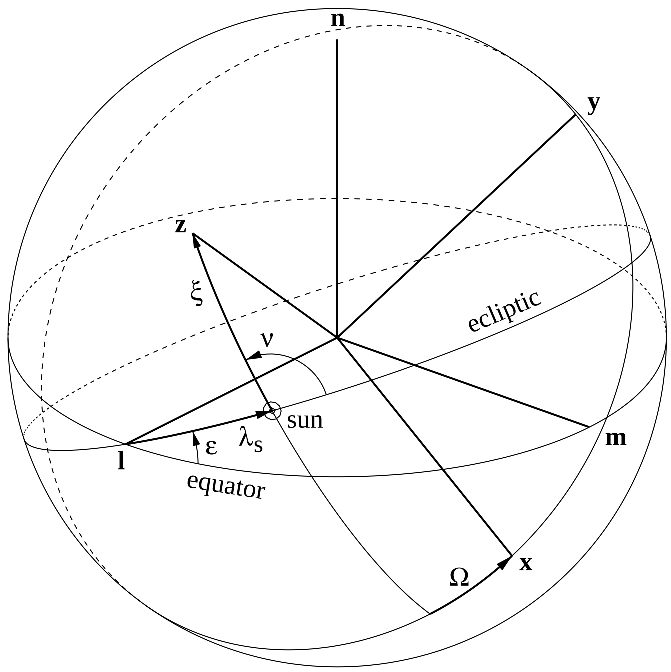


FIGURE 1. Definition of angles $\epsilon$, $\lambda_s$, $\nu$, $\xi$ and $\Omega$ in the nominal scanning law.

# 5 The nominal scanning law

The nominal scanning law for GAIA will be similar to that of Hipparcos, only with different constants. It is described by two constant angles ($\epsilon$ = obliquity of equator, and $\xi$ = revolving angle), and three angles which increase continuously but non-uniformly with time: $\lambda_s(t)$ = nominal longitude of the Sun, $\nu(t)$ = revolving phase, and $\Omega(t)$ = spin phase. See Fig. 1 for the definition of these angles.

As can be seen from the figure, a frame initially aligned with **N** becomes aligned with the nominal **Z** after the following sequence of rotations: (1) by $\epsilon$ about the first axis; (2) by $\lambda_s$ about the third axis; (3) by $\nu - 90°$ about the first axis; (4) by $90° - \xi$ about the second axis; (5) by $\Omega$ about the third axis. In terms of quaternion multiplications, this gives:

$$\mathbf{q} = \begin{bmatrix} \sin\frac{1}{2}\epsilon \\ 0 \\ 0 \\ \cos\frac{1}{2}\epsilon \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \sin\frac{1}{2}\lambda_s \\ \cos\frac{1}{2}\lambda_s \end{bmatrix} \begin{bmatrix} \sin\frac{1}{2}(\nu - 90°) \\ 0 \\ 0 \\ \cos\frac{1}{2}(\nu - 90°) \end{bmatrix} \begin{bmatrix} 0 \\ \sin\frac{1}{2}(90° - \xi) \\ 0 \\ \cos\frac{1}{2}(90° - \xi) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \sin\frac{1}{2}\Omega \\ \cos\frac{1}{2}\Omega \end{bmatrix} . \quad (6)$$

The precise specification of the functions $\lambda_s(t)$, $\nu(t)$ and $\Omega(t)$ remains to be done (although part of the discussion was made in SAG–LL–014), but the general principles are easily stated:

1. $\lambda_s(t)$ should be a reasonably simple analytical function which approximates the

satellitocentric proper direction to the sun to within the tolerances set by accepted variations in the *true* solar angle (probably some arcmin). This may be more complicated than for Hipparcos, due to the tighter tolerances for GAIA and the L2 orbit.

2. Given $\lambda_s(t)$, the function $\nu(t)$ should then be chosen to give a (nearly) constant precession rate $|\dot{\mathbf{z}}|$. The required precession rate is set by the loop overlap condition discussed in SAG–LL–014.

3. Finally, for given $\lambda_s(t)$ and $\nu(t)$, the function $\Omega(t)$ is chosen to give constant inertial rotation rate about the instrument $\mathbf{z}$ axis, $\omega_z = 120$ arcsec s$^{-1}$.

## 6 The attitude matrix

In terms of the components of the quaternion, the attitude matrix is

$$\mathbf{A}(\mathbf{q}) = \frac{1}{s}\begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix}, \tag{7}$$

where $s = |\mathbf{q}(t)|^2$. The normalization factor allows for the possible non-unit length of $\mathbf{q}$.

## 7 The inertial rotation vector

Let $\boldsymbol{\omega}$ denote the inertial rotation vector of the satellite. Its components in the instrument system are $\mathbf{Z}'\boldsymbol{\omega} = [\omega_x\ \omega_y\ \omega_z]'$. In the quaternion formalism, the kinematic equations are:

$$\dot{\mathbf{q}} = \frac{1}{2}\begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix}\mathbf{q}\,. \tag{8}$$

This equation actually assumes that $|\mathbf{q}| = 1$. If $\mathbf{q}$ is not normalized, introduce $s(t) = |\mathbf{q}(t)|^2$. Then Eq. (8) is modified in that the diagonal elements of the matrix become equal to $\dot{s}/s$. If $\dot{s}$ is negligible, then Eq. (8) remains valid also for the non-normalized $\mathbf{q}$.

Equation (8) can be inverted to give $\boldsymbol{\omega}$ in the instrument frame:

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \frac{2}{s}\begin{bmatrix} q_4 & q_3 & -q_2 & -q_1 \\ -q_3 & q_4 & q_1 & -q_2 \\ q_2 & -q_1 & q_4 & -q_3 \end{bmatrix}\dot{\mathbf{q}}\,. \tag{9}$$

This is strictly valid also when $\dot{s} \neq 0$. The rotation vector in this frame is needed to compute the scan rates along and across scan at different points of the fields, and it is therefore essential that it can be computed accurately and simply.

The rotation vector in the celestial frame is obtained from Eq. (9) and the attitude matrix,

$$\mathbf{N}'\boldsymbol{\omega} = \begin{bmatrix} \omega_l \\ \omega_m \\ \omega_n \end{bmatrix} = \mathbf{A}' \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} . \tag{10}$$

Combining Eqs. (7) and (9) we obtain after some algebra

$$\begin{bmatrix} \omega_l \\ \omega_m \\ \omega_n \end{bmatrix} = \frac{2}{s} \begin{bmatrix} q_4 & -q_3 & q_2 & -q_1 \\ q_3 & q_4 & -q_1 & -q_2 \\ -q_2 & q_1 & q_4 & -q_3 \end{bmatrix} \dot{\mathbf{q}} . \tag{11}$$

The simplicity of Eqs. (9) and (11) is a good illustration of the compact elegance of the quaternion formalism.

## 8 Attitude determination

As part of the 'global iterative solution' we need to determine $\mathbf{q}(t)$ from the following data: (i) a set of known celestial directions (computed from a star catalogue); (ii) the CCD observations of these directions; and (iii) a calibration of the instrument parameters. As already mentioned, the four components of $\mathbf{q}(t)$ must be expressed as continuous functions, whose coefficients ('attitude parameters') are to be determined. In this section I consider briefly the more limited problem, viz. how to determine the instantaneous attitude $\mathbf{q}$ from a set of (quasi-simultaneous) measurements.

Each measurement is essentially an association of a certain time instant $t$ with certain angular coordinates $(\eta, \zeta)$ in the instrument frame (Fig. 2), relating to a certain direction $\mathbf{u}$. For instance, $t$ could be the observed transit time of a star (with known $\mathbf{N}'\mathbf{u}$) at a certain CCD column, whose central coordinates $(\eta, \zeta)$ are known from the instrument calibration. Let us see how this translates into observation equations for $\mathbf{q}$. Since the measurements are non-linear functions of $\mathbf{q}$, it is necessary to linearize. Let $\mathbf{q}_0$ be the current attitude estimate, which in the next iteration is improved to $\mathbf{q} = \mathbf{q}_0 + \Delta\mathbf{q}$. From $\mathbf{N}'\mathbf{u}$ and $\mathbf{A}(\mathbf{q}_0)$ we obtain the calculated direction $\mathbf{Z}'\mathbf{u}$ and hence the calculated $(\eta, \zeta)$ from the relation

$$\left. \begin{array}{rcl} u_x & = & \cos\zeta\cos\eta \\ u_y & = & \cos\zeta\sin\eta \\ u_z & = & \sin\zeta \end{array} \right\} \tag{12}$$

The differences $\Delta\eta\cos\zeta$, $\Delta\zeta$, taken in the sense observed *minus* calculated, can now be related to $\Delta\mathbf{q}$ by means of a differential rotation obtained in analogy with Eq. (9):

$$\begin{bmatrix} \Delta\eta\cos\zeta \\ \Delta\zeta \end{bmatrix} = \frac{2}{s}\mathbf{C} \begin{bmatrix} q_4 & q_3 & -q_2 & -q_1 \\ -q_3 & q_4 & q_1 & -q_2 \\ q_2 & -q_1 & q_4 & -q_3 \end{bmatrix} \Delta\mathbf{q} , \tag{13}$$

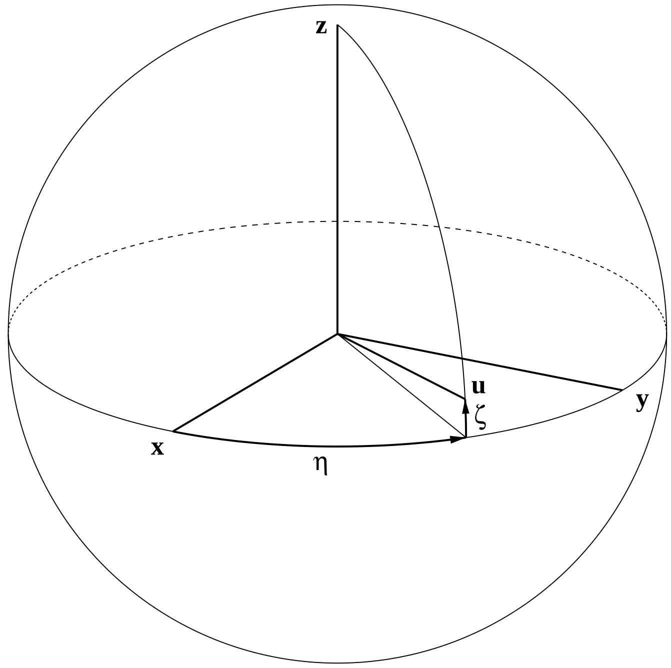


FIGURE 2. At a certain instant $t$ the celestial direction $\mathbf{u}$ is observed at spherical coordinates $(\eta, \zeta)$ in the instrument frame.

where $\mathbf{C}$ is fixed for a given CCD column:

$$\mathbf{C} = \begin{bmatrix} \sin\zeta\cos\eta & \sin\zeta\sin\eta & -\cos\zeta \\ -\sin\eta & \cos\eta & 0 \end{bmatrix} . \tag{14}$$

After normalization by the estimation errors $\sigma_\eta \cos\zeta$ and $\sigma_\zeta$, Eq. (13) gives two uncorrelated observation equations for each direction observed at the same time. Let $\mathbf{h} = \mathbf{B} \cdot \Delta\mathbf{q}$ be the full set of such equations. The required solution minimizes $|\mathbf{h} - \mathbf{B} \cdot \Delta\mathbf{q}|^2$ subject to the constraint $\mathbf{q}_0' \Delta\mathbf{q} = 0$ (to preserve the length of $\mathbf{q}$). Using the method of Lagrangian multiplier, the solution is obtained by the following steps. First, the normal+constraint equations

$$(\mathbf{B}'\mathbf{B}) \left[\ \mathbf{a} \quad \mathbf{b}\ \right] = \left[\ \mathbf{B}'\mathbf{h} \quad \mathbf{q}_0\ \right] \tag{15}$$

are solved to give the four-vectors $\mathbf{a}$ and $\mathbf{b}$. Then, the constrained solution is given by:

$$\Delta\mathbf{q} = \mathbf{a} - \mathbf{b}\lambda\,, \quad \text{where} \quad \lambda = \frac{\mathbf{q}'\mathbf{a}}{\mathbf{q}'\mathbf{b}}\,. \tag{16}$$

In the actual case the attitude is expressed in terms of basis functions, $\mathbf{q}(t) = \sum_i \mathbf{a}_k f_k(t)$, and $\mathbf{a}_k$ are then the attitude parameters to be determined. In terms of the observation equations (13), this is just a trivial modification. However, it is not obvious if and how normalization constraints should be introduced in the continuous case. Although many details such as this remain to be worked out, I think it is clear that the use of quaternions will have great advantages over alternative attitude parameterizations.