Efficient Attitude Representations and Formulations

Jozef C. van der Ha

Department of Aeronautics and Astronautics, Kyushu University, Fukuoka, Japan

The paper presents a summary of efficient attitude representations and formulations with relevance to actual spacecraft applications. The presentation follows an approach that makes generous use of Euler’s Theorem. The mathematical results are supported by geometrical interpretations. The principal motivation for the present paper originates from educational objectives.

Key Words: Attitude Dynamics, Attitude Representations

1. Introduction

The modeling of a space vehicle’s attitude motion is an important aspect of its design and development activities. The mission performance usually depends on the proper execution of the attitude determination and control functions, for instance by providing the required pointing orientation for the instruments.

In order to model the evolution of the satellite’s attitude motion under the prevailing environmental and control torques, a mathematical representation of the attitude orientation as a function of time must be adopted. This is not a straightforward task because of the many available options, in particular for three-axis-stabilized satellites. The commonly used attitude representations can be established using coordinate transformations as shown in the reference book by [Wertz] and in the survey paper of [Shuster].

2. Vector Rotations

First, we provide a general model for describing arbitrary vector rotations in space, which forms the basis for the attitude representations presented below.

2.1 Euler’s Theorem

The starting point for describing vector and attitude rotations is Euler’s theorem [Hughes], p. 10:

The general displacement of a rigid body with one fixed point is a rotation about an axis through that point

This fixed rotation axis is designated as the Euler axis which is designated by the unit-vector \( \mathbf{e} \). The rotation angle about the Euler axis is represented by the Euler angle \( \Phi \) as shown in Figure 1.

We analyze the effect of a general rotation angle \( \Phi \) about the Euler axis \( \mathbf{e} \) on an arbitrary initial vector \( \mathbf{a} \). The result can best be visualized by the elementary geometry illustrated in Figure 1. The vector \( \mathbf{b} \) resulting from the rotation about the axis \( \mathbf{e} \) can be broken up as the sum of three vectors. After starting with the vector \( \mathbf{a} \) itself, we add the vector of length \( \rho \) \((1 - \cos \Phi)\), with radius \( \rho = |\mathbf{e} \times \mathbf{a}| = |\mathbf{a}| \sin \alpha \), in the direction of the vector \( \mathbf{e} \times (\mathbf{e} \times \mathbf{a}) \). Finally, we add the third vector of length \( \rho \sin \Phi \) along the vector \( \mathbf{e} \times \mathbf{a} \). The complete expression for the vector \( \mathbf{b} \) is thus:

\[
\mathbf{b} = \mathbf{a} + (1 - \cos \Phi) \mathbf{e} \times (\mathbf{e} \times \mathbf{a}) + \sin \Phi (\mathbf{e} \times \mathbf{a}) \tag{1}
\]

A more formal derivation of this result can be found in [Shuster], pp. 449-451, in particular eq. (101b).
2.2 Infinitesimal Vector Rotation

In order to obtain a deeper understanding of the nature of vector rotations we study the effect of an infinitesimal rotation $\Delta \Phi$ about the Euler axis on an arbitrary vector $\mathbf{r}$ as shown in Figure 2:

$$\Delta \mathbf{r} = (\mathbf{e} \times \mathbf{r}) \Delta \Phi \quad (2)$$

**Fig. 2 – Infinitesimal Vector Rotation**

It is convenient to employ the skew-symmetric matrix $[\mathbf{e} \times]$, which is defined in analogy with the common vector cross-product:

$$\mathbf{e} \times \mathbf{r} = [\mathbf{e} \times] \mathbf{r}, \quad \text{with} \quad [\mathbf{e} \times] = \begin{bmatrix}
0 & -e_3 & e_2 \\
e_3 & 0 & -e_1 \\
-e_2 & e_1 & 0
\end{bmatrix} \quad (3)$$

The entries $e_j \ (j = 1, 2, 3)$ refer to the components of the (Euler axis) unit-vector $\mathbf{e}$ in an arbitrary inertial reference frame.

When performing the limit for $\Delta \Phi \to 0$, eq. (2) takes the form of a linear vector differential equation with constant coefficients:

$$\mathbf{r}'(\Phi) = \lim_{\Delta \Phi \to 0} \left\{ \frac{\Delta \mathbf{r}}{\Delta \Phi} \right\} = [\mathbf{e} \times] \mathbf{r} \quad (4)$$

2.3 General Large Angle Rotation

The exact solution of eq. (4) for an arbitrary initial condition $\mathbf{r}_0 = \mathbf{r} \ (\Phi = 0) = \mathbf{a}$ can be obtained from linear system theory, see for instance [Rugh]¹, p. 66:

$$\mathbf{r}(\Phi) = \exp[\Phi [\mathbf{e} \times]] \mathbf{a} \quad (5)$$

The infinite series expansion of the exponential function allows us to express $\mathbf{r} \ (\Phi)$ in the form:

$$\mathbf{r}(\Phi) = \sum_{k=0}^{\infty} \frac{1}{k!} (\Phi [\mathbf{e} \times])^k \mathbf{a}$$

The second power of the matrix can be written in explicit form by using the definition of $[\mathbf{e} \times]$ in eq. (3):

$$[\mathbf{e} \times]^2 = \begin{bmatrix}
e_1^2 -1 & e_1 e_2 & e_1 e_3 \\
e_2 e_1 & e_2^2 -1 & e_2 e_3 \\
e_3 e_1 & e_3 e_2 & e_3^2 -1
\end{bmatrix} = [\mathbf{e} \times][\mathbf{e} \times] \quad (7)$$

Here, $[\mathbf{e} \times]^T$ denotes the vector outer-product and $[\mathbf{I}]$ is the identity matrix. It can easily be confirmed that the product of the matrices $[\mathbf{e} \times]$ and $[\mathbf{e} \times]^T$ vanishes. Therefore, eq. (7) produces $[\mathbf{e} \times]^3 = -[\mathbf{e} \times]$ and the higher powers of $[\mathbf{e} \times]$ can be reduced to:

$$[\mathbf{e} \times]^{2j} = (-1)^{j+1} [\mathbf{e} \times]^j \quad j = 1, 2, \ldots \quad (8)$$

After substituting these results in the series expansion of eq. (6), we obtain:

$$\mathbf{r}(\Phi) = \mathbf{a} + \{ \Phi - \Phi^3 / 3! + \Phi^5 / 5! + \ldots \} [\mathbf{e} \times] \mathbf{a} +$$
$$+ \{ \Phi^2 / 2! - \Phi^4 / 4! + \Phi^6 / 6! + \ldots \} [\mathbf{e} \times]^2 \mathbf{a} =$$

$$= ([\mathbf{I}] + \sin \Phi [\mathbf{e} \times] + (1 - \cos \Phi) [\mathbf{e} \times]^2) \mathbf{a} \quad (9)$$

or:

$$\mathbf{r}(\Phi) = [\mathbf{R}(\Phi)] \mathbf{a}, \quad \text{with} :$$

$$[\mathbf{R}(\Phi)] = [\mathbf{I}] + \sin \Phi [\mathbf{e} \times] + (1 - \cos \Phi) [\mathbf{e} \times]^2 \quad (10)$$

The expressions in eq. (1) and eq. (10) are obviously equivalent. Since both $[\mathbf{e} \times] \mathbf{e}$ and $[\mathbf{e} \times]^2 \mathbf{e}$ vanish, Eq. (10) leads to $[\mathbf{R}] \mathbf{e} = \mathbf{e}$, which implies that $\mathbf{e}$ is an eigen-vector of the matrix $[\mathbf{R}]$ with eigen-value $\lambda = 1$. This is self-evident because the vector $\mathbf{e}$ is invariant to a rotation about itself.

In general, the rotation over the positive angle $\Phi$ about the fixed Euler axis $\mathbf{e}$ maps the vector $\mathbf{a} = \mathbf{r}_0$ into the vector $\mathbf{b} = \mathbf{r} \ (\Phi) = [\mathbf{R}] \mathbf{a}$. Similarly, it can readily be shown that the inverse mapping is given by $\mathbf{a} = \mathbf{r} \ (-\Phi) = [\mathbf{R}]^T \mathbf{b}$, with:

$$[\mathbf{R}]^T = [\mathbf{I}] - \sin \Phi [\mathbf{e} \times] + (1 - \cos \Phi) [\mathbf{e} \times]^2 \quad (11)$$

Finally, it is a straightforward exercise to show that $[\mathbf{R}] \ [\mathbf{R}]^T = [\mathbf{I}]$ and that the determinant of $[\mathbf{R}] = 1$. Therefore, $[\mathbf{R}]$ is a proper orthogonal matrix.
3. Reference Frame Transformations

Now we consider the general transformation from one reference frame to another. This provides the mapping of a vector with components in one reference frame into the same vector (but with other components) in a different frame. The understanding of these transformations is important for describing the space vehicle’s attitude motion. After attaching a reference frame to the satellite (which is assumed rigid here), we describe the satellite’s attitude motion through the evolution of the satellite-fixed frame.

3.1 Direction-Cosine Matrix

We consider the fundamental reference frame with axes (X₁, X₂, X₃) and the satellite-fixed body frame with axes (x₁, x₂, x₃) as shown in Figure 3. Usually, but not always, the (X₁, X₂, X₃) frame is an inertial reference frame. The unit-vectors along the axes of these two coordinate frames are designated by Xᵢ (i = 1, 2, 3) and xⱼ (j = 1, 2, 3), respectively.

![Fig. 3 – Satellite Frame in Reference Frame](image)

The projections of the unit-vectors xⱼ (j = 1, 2, 3) along the satellite body axes onto the inertial reference axes with unit-vectors Xᵢ (k = 1, 2, 3) are:

\[
x_j = \sum_{k=1}^{3} (x_j \cdot X_k) X_k = \sum_{k=1}^{3} (\cos \theta_{jk}) X_k \quad (12)
\]

We introduce the Direction-Cosine Matrix (DCM) (also known as the attitude matrix) [A] as follows:

\[
[A] = \begin{bmatrix}
(x_1 \cdot X_1) & (x_2 \cdot X_1) & (x_3 \cdot X_1) \\
(x_1 \cdot X_2) & (x_2 \cdot X_2) & (x_3 \cdot X_2) \\
(x_1 \cdot X_3) & (x_2 \cdot X_3) & (x_3 \cdot X_3)
\end{bmatrix} \quad (13)
\]

When using the projections of the unit-vector in eq. (12) we can show that:

\[
\sum_{i=1}^{3} (x_j \cdot X_i) (X_i \cdot x_k) = (x_j \cdot x_k) = \delta_{jk} \quad (j, \ell = 1, 2, 3) \quad (14)
\]

where $\delta_{jk}$ is the Kronecker delta. Eq. (14) is useful for confirming that $[A] [A]^T$ is equal to the identity matrix. Also it follows by explicit calculation that the determinant of $[A]$ equals $x_i \cdot (x_i \times x_i) = 1$ for a right-handed reference frame. Thus, $[A]$ is a proper real orthogonal matrix, see [Wertz]¹, section 12.1.

3.2 Transformation of Arbitrary Vector

We consider the arbitrary vector $a^b = (a^b_1, a^b_2, a^b_3)^T$ with components in the body reference frame:

\[
a^b = a^b_1 x_1 + a^b_2 x_2 + a^b_3 x_3 = (a^b_1, a^b_2, a^b_3) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (15)
\]

We wish to express the body vector $a^b$ in terms of its components along the (X₁, X₂, X₃) reference axes. From eqs. (12) and (13) produce:

\[
a^b = (a^b_1, a^b_2, a^b_3) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (a^b_1, a^b_2, a^b_3) [A] \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad (16)
\]

The components of $a^b$ along the (X₁, X₂, X₃) axes are denoted by $a^i = (a^i_1, a^i_2, a^i_3)^T$. Eq. (16) allows us to write $a^i$ in the components of $a^b$ and vice versa:

\[
(a^i_1, a^i_2, a^i_3) = (a^b_1, a^b_2, a^b_3) [A] \quad \Rightarrow \quad \begin{bmatrix} a^i_1 \\ a^i_2 \\ a^i_3 \end{bmatrix} = [A]^T a^b \quad (17)
\]

The final equation maps the vector $a^i$ with its components in the fundamental reference frame into the vector $a^b$ with components in the body frame. In explicit terms we find from eq. (13):

\[
a^i_j = \sum_{k=1}^{3} (x_j \cdot X_k) a^b_k = (x_j \cdot a^b_k) \quad (j = 1, 2, 3) \quad (18)
\]

Thus, the structure of the transformation between $a^b$ and $a^i$ in eq. (18) is identical to the one between the body unit-vectors $x_j$ and the $X_k$ vectors in eq. (12).

3.3 Geometrical Interpretation

The transformation between reference frames can also be understood in terms of Euler’s Theorem, see [Schaub and Junkins]⁵, p. 87:

A reference frame can be brought from an arbitrary initial orientation to an arbitrary final orientation by means of a single rotation about a fixed rotation axis.
The Euler axis and angle are of course the Euler axis and angle introduced in Figure 1. The Euler axis \( \mathbf{e} \) maintains a fixed orientation in both the initial and the final frames as illustrated in Figure 4.

The general transformation between reference frames can be understood by employing the result for the vector rotation in eq. (10) and Figure 1. Each of the body unit-vectors \( \mathbf{x}_i \) \( (i = 1, 2, 3) \) originates from the mapping of the corresponding fundamental unit-vectors \( \mathbf{X}_i \) \( (i = 1, 2, 3) \) due to the rotation about the Euler axis \( \mathbf{e} \) over the angle \( \Phi \) (see Figure 4). Therefore, we have as in eq. (10):

\[
\mathbf{x}_i(\Phi) = [\mathbf{R}(\Phi)]\mathbf{X}_i \quad (i = 1, 2, 3)
\]

The matrix \([\mathbf{R}]\) is given by:

\[
[\mathbf{R}] = [\mathbf{I}] + S[\mathbf{e}_x \times] + (1-C)[\mathbf{e}_x \times]^2 =
= C[\mathbf{I}] + S[\mathbf{e}_x \times] + (1-C)[\mathbf{e}_x \times][\mathbf{e}_x \times]^T
\]

with abbreviations \( C = \cos \Phi \) and \( S = \sin \Phi \).

Eqs. (19) and (20) produce explicit results, in terms of the Euler axis and angle, for the projections of the body unit-vectors \( \mathbf{x}_i \) \( (j = 1, 2, 3) \) onto the fundamental reference axes. When substituting the explicit unit-vectors \( \mathbf{X}_i \) \( (i = 1, 2, 3) \) and the entries of the matrices \([\mathbf{e}_x \times]\) in eq. (3) and \([\mathbf{e}_x \times]^T\) in eq. 7, we find:

\[
\begin{align*}
\mathbf{x}_1 & = [\mathbf{R}(\Phi)] \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} C + e_1^2(1-C) \\ e_2e_1(1-C) + e_1S \\ e_3e_1(1-C) - e_2S \end{pmatrix} \\
\mathbf{x}_2 & = [\mathbf{R}(\Phi)] \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e_2e_1(1-C) - e_1S \\ e_3e_1(1-C) + e_2S \\ C + e_3^2(1-C) \end{pmatrix}
\end{align*}
\]

Finally, we recall that the components of \( \mathbf{x}_i \) \( (j = 1, 2, 3) \) along the fundamental axes \( \mathbf{X}_i \) \( (k = 1, 2, 3) \) are the same as \( (\mathbf{x}_i \cdot \mathbf{X}_k) \). This means that the column vectors in eqs. (21) are precisely the rows of the DCM rotation matrix \([\mathbf{A}]\) in eq. (13):

\[
[\mathbf{A}] =
\begin{bmatrix}
C + e_1^2(1-C) & e_2e_1(1-C) + e_1S & e_3e_1(1-C) - e_2S \\
e_2e_1(1-C) - e_1S & C + e_2^2(1-C) & e_3e_1(1-C) + e_2S \\
e_3e_1(1-C) + e_2S & e_3e_1(1-C) - e_2S & C + e_3^2(1-C)
\end{bmatrix}
\]

When comparing this result with eq. (20) we find:

\[
[\mathbf{A}] = [\mathbf{R}]^T = C[I] - S[\mathbf{e}_x \times] + (1-C)[\mathbf{e}_x \times]^T
\]

It can also be shown that \([\mathbf{A}]\mathbf{e} = \mathbf{e}\) so that the Euler axis \( \mathbf{e} \) is an eigenvector of the matrix \([\mathbf{A}]\) with the eigenvalue \( \lambda = 1 \). This confirms that the Euler axis is invariant to the mapping induced by \([\mathbf{A}]\).

The result in eq. (22) represents the most general form for an arbitrary transformation matrix in (three-dimensional) space. Conversely, when an arbitrary transformation matrix \([\mathbf{B}]\) is given, it is easy to calculate the Euler angle and Euler axis from the components \( b_{ij} \) \( (j = 1, 2, 3) \) of \([\mathbf{B}]\):

\[
\Phi = \arccos\left(\frac{\text{trace}[\mathbf{B}] - 1}{2}\right)
\]

\[
\mathbf{e} = \left(b_{23} - b_{32}, b_{31} - b_{13}, b_{12} - b_{21}\right) / (2\sin \Phi)
\]

where \( \text{trace}[\mathbf{B}] \) is the sum of the diagonal terms \( b_{ii} \) \( (i = 1, 2, 3) \). There are two feasible solutions due to the fact that the rotation \( (\mathbf{e}, \Phi) \) is identical to \(- (\mathbf{e}, -\Phi)\).

4. Useful Attitude Representations

From the general results given above, different attitude representations can be established, see the exhaustive summary of [Shuster]. We present a few representations that are commonly used in practice.

4.1 Quaternion Representation

Perhaps the most popular attitude representation is the quaternion with the Euler symmetric parameters as its components, see [Wertz], p. 414. They are defined by the Euler axis and Euler angle as follows:

\[
q_j = e_j \sin(\Phi/2), \quad j = 1, 2, 3; \quad q_4 = \cos(\Phi/2)
\]

These four parameters are not independent because \( q_4^2 = 1 \). The terms in the direction-cosine matrix \([\mathbf{A}]\) of eq. (22) can be expressed in \( \mathbf{q} = (q_1, ..., q_4)^T \):
The quaternion offers a few significant advantages compared to the direction-cosine matrix. Only four parameters are needed and two individual rotations can be combined very easily ([Wertz]$^1$, p. 415-416).

Another efficient representation is the Gibbs vector with its components defined by the Euler parameters: $g_j = q_j/q_4$ ($j = 1, 2, 3$). The Gibbs vector consists of only three parameters, which is the minimum. The associated DCM matrix is given in [Wertz]$^1$, p. 416.

### 4.2 Tait-Bryan Angles

Finally, we present the Tait-Bryan (TB) attitude angles, which is one set of the classical Euler angles, see [Wertz]$^1$, pp. 417-420, and p. 764. The Tait-Bryan angles are defined as the 1-2-3 sequence of Euler angles, i.e. successive rotations about the $x$, $y$, $z$ axes over the angles $\phi$, $\theta$, $\psi$, respectively (Fig. 5).

![Fig. 5 - Tait-Bryan Angles (1-2-3 Sequence)](image)

The general transformation matrix $[A(\phi, \theta, \psi)]$ in terms of the TB angles is given in [Wertz]$^1$, p. 764. The use of the TB angles is especially attractive when the satellite frame is close to an adopted reference frame, e.g. Earth-pointing satellite missions.

In these applications, both the Euler angle and the TB angles are small quantities, written as $\Delta \Phi$ and $(\Delta \phi, \Delta \theta, \Delta \psi)^T$, respectively. Therefore, we may neglect terms of second and higher order in the TB angles and the general transformation matrix in [Wertz]$^1$, p. 764, can be simplified into the form:

$$[A(q)] = \begin{bmatrix}
q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_2q_4 + q_3q_1) & 2(q_3q_2 - q_4q_1) \\
2(q_2q_4 - q_3q_1) & q_1^2 - q_2^2 - q_4^2 + q_3^2 & 2(q_3q_4 + q_1q_2) \\
2(q_3q_2 + q_4q_1) & 2(q_3q_4 - q_1q_2) & q_1^2 - q_3^2 - q_4^2 + q_2^2
\end{bmatrix}
$$

or:

$$[A(q)] = (q_1^2 - q_2^2 - q_3^2 + q_4^2)[I] - 2q_4[q\times] + 2qq^T \quad (26.1)$$

In this special case, the general result of eq. (23) can be reduced to only first-order terms in $\Delta \Phi$:

$$[A] = [R]^T = [I] - [\Delta \Phi e] = \begin{bmatrix}
1 & \Delta \phi & -\Delta \theta \\
-\Delta \psi & 1 & \Delta \phi \\
-\Delta \theta & -\Delta \phi & 1
\end{bmatrix} \quad (27)$$

It is advantageous to define the vector $\Delta \Phi = \Delta \Phi e$. When comparing eq. (28.2) with eq. (27) we find:

$$\Delta \Phi = \sqrt{\Delta \phi^2 + (\Delta \theta)^2 + (\Delta \psi)^2} \quad (29.1)$$

$$e = (\Delta \phi, \Delta \theta, \Delta \psi)^T / \Delta \Phi \quad (29.2)$$

The same results may also be obtained from eqs. (24). That approach, however, would require second-order expansions (in the TB angles $\Delta \phi$, $\Delta \theta$, and $\Delta \psi$) in the diagonal terms of the matrix $[A]$ in eq. (27) in order to be able to calculate the Euler angle $\Delta \Phi$.

### 5. Conclusions

The paper provides a review of rotations of vector and reference frames based on the application of Euler’s Theorem. The resulting mathematical expressions are interpreted by straightforward geometrical arguments. The analysis aims at the selection of suitable attitude representations for practical space applications. In particular, the Tait-Bryan angles are proposed as an efficient representation for three-axis applications when the attitude is kept close to a reference attitude.

### 6. References