QUATERNION DEFINITIONS, NOTATION, AND IDENTITIES

Abstract

Standard notation, definitions, and identities for quaternion algebra and direction cosine matrices are given. In particular, quaternion product operators and operator matrices are defined. The quaternion product operators are necessary to avoid ambiguity in quaterion algebra. The cross product operator and operator matrix are also defined.

1 Introduction

The classical definition of the quaterion product is the algebraic product of two hypercomplex numbers. The product is computed by ordinary multiplication and by using the product rules for the basis elements $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, and $\hat{\boldsymbol{k}}$ of the hypercomplex numbers. The order of the quaternion product is the reverse of the corresponding product of direction cosine matrics (DCM). In analytical work, it is convenient to write the quaternion as a column matrix. An operator then becomes necessary to define the quaternion product of the 4×1 vectors. The quaternion product can then be written in either order with a corresponding product operator symbol.

It is not uncommon to find ambiguous and incorrect notation in the literature. One purpose of this note is to provide a reference to standard notation. The $\circledast$ operator, which was first introduced by the author in an earlier paper, may not be considered standard but has gained acceptance. A second purpose of this note is to provide definitions and identities that are useful in attitude analysis.

Standard notation, definitions, quaternion properties and identities, and direction cosine matrix (DCM) identities are provided in this note. The identities are stated without proof. The studious reader can easily prove the identities or find them in standard texts and references such as [1] and [2], albeit with notational differences. We will denote a 3×3 identity matrix by $\mathbf{I}$ and all other identity matrices by $\boldsymbol{I}$. A vector $\boldsymbol{v}$ in a particular frame a is denoted by $\boldsymbol{v}^a$. A transformation matrix is a DCM that can be used to transform a vector from a frame a to a frame b, which we will denote by $\mathbf{T}_a^b$, so we have $\boldsymbol{v}^b = \mathbf{T}_a^b \boldsymbol{v}^a$. Quaternions may be adorned in the same manner as transformation matrices.

2 Cross Product Operator

The cross-product operator $\times$ is such that for vectors $\boldsymbol{u} = [u_x,\, u_y,\, u_z]^T$ and $\boldsymbol{v} = [v_x,\, v_y,\, v_z]^T$, we have

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{bmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{bmatrix} . \qquad (1)$$

The cross-product operator matrix $[\boldsymbol{u}\times]$ is defined such that

$$[\boldsymbol{u}\times]\boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{v} \qquad (2)$$

where

$$[\boldsymbol{u}\times] = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} . \tag{3}$$

The cross-product matrix matrix is often written as $\boldsymbol{u}^\times$ such that $\boldsymbol{u}^\times \equiv [\boldsymbol{u}\times]$ and $\boldsymbol{u}^\times \boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{v}$. The notation $\llbracket \boldsymbol{u} \rrbracket$, where $\llbracket \boldsymbol{u} \rrbracket \equiv -[\boldsymbol{u}\times]$, is apparently attributable to and unique to the work of Shuster; it first appeared in [3] and no mention of its origin is given in [2]. In his earlier work Shuster used $[\boldsymbol{u}] \equiv -[\boldsymbol{u}\times]$. Some authors use the clumsier notation $S = \mathrm{skew}(\boldsymbol{u}) \equiv [\boldsymbol{u}\times]$.

Since the cross-product matrix is skew-symmetric, we have $[\boldsymbol{u}\times]^T = -[\boldsymbol{u}\times]$ and $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$. A useful identity is $[\boldsymbol{u}\times]^2 = |\boldsymbol{u}|^2 \mathbf{I} - \boldsymbol{u}\boldsymbol{u}^T$. A reference frame transformation may be applied to the cross-product matrix so that

$$[\mathbf{T}_a^b \boldsymbol{v}^a \times] = \mathbf{T}_a^b [\boldsymbol{v}^a \times] \mathbf{T}_b^a \, . \tag{4}$$

3 Quaternion Operators

In the definitions that follow, let $\boldsymbol{p}$, $\boldsymbol{q}$, and $\boldsymbol{r}$ be quaternions represented as 4×1 column matrices and let $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{q}}$, and $\tilde{\boldsymbol{r}}$ be the corresponding hypercomplex quaternions. Also let $\mathbf{T}_{\boldsymbol{p}}$, $\mathbf{T}_{\boldsymbol{q}}$, and $\mathbf{T}_{\boldsymbol{r}}$ be the corresponding orthonormal direction cosine matrices. The subscript here is a shorthand notation that simply indicates that $\mathbf{T}$ corresponds to a particular quaternion, otherwise we would have to write, for example, $\mathbf{T}(\boldsymbol{q})$.

Quaternion operator notation is necessary because while the hypercomplex product $\tilde{\boldsymbol{p}}\tilde{\boldsymbol{q}}$ is well-defined, it is technically incorrect and sometimes causes confusion to write the matrix product $\boldsymbol{p}\boldsymbol{q}$. Some authors write the matrix product as $\boldsymbol{p} * \boldsymbol{q}$. The use of operator notation removes these difficulties and provides much greater flexibility in performing quaternion algebra. The quaternion operators and operator matrices to be introduced are analagous to the cross-product operator $\times$ and operator matrix $[\boldsymbol{u}\times]$. The quaternion $\boldsymbol{q} = [q_x,\, q_y,\, q_z,\, q_s]^T$ comprises the vector $[q_x,\, q_y,\, q_z]^T$ and the scalar q_s. The norm (2-norm) of a quaternion is given by $|\boldsymbol{q}| = (q_x^2 + q_y^2 + q_z^2 + q_s)^{1/2}$. It is usually assumed in attitude work that quaternions have unit norm, but for generality we will not make that assumption here. (Ultimately, however, an implicit or explicit normalization occurs when computing a rotation vector, Gibbs or Rodriguez vector, DCM, or any other attitude representation.) The conjugate $\boldsymbol{q}^*$ of a quaternion $\boldsymbol{q} = [q_x,\, q_y,\, q_z,\, q_s]^T$ is $\boldsymbol{q}^* = [-q_x,\, -q_y,\, -q_z,\, q_s]^T$. The inverse of $\boldsymbol{q}$ is $\boldsymbol{q}^{-1} = \boldsymbol{q}^*/|\boldsymbol{q}|^2$ so that $\boldsymbol{q} \otimes \boldsymbol{q}^{-1} = \boldsymbol{q}^{-1} \otimes \boldsymbol{q} = [0, 0, 0, 1]^T$. For a unit quaternion, $\boldsymbol{q}^* = \boldsymbol{q}^{-1}$ so that $\boldsymbol{q}^* \otimes \boldsymbol{q} = \boldsymbol{q} \otimes \boldsymbol{q}^* = [0, 0, 0, 1]^T$ and similarly for the $\circledast$ operator, where $\otimes$ and $\circledast$ are quaternion multiplication operators defined below. The $[\boldsymbol{q}\otimes]$ and $[\boldsymbol{q}\circledast]$ operator matrices are also introduced. The $\otimes$ operator was introduced in [4]. The operator matrices and the $\circledast$ operator (in this context) are this author's invention.

4 The $\otimes$ Operator

The operator $\otimes$ and operator matrix $[\boldsymbol{q}\otimes]$ are defined by

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{p} \otimes \boldsymbol{q} \\ &= [\boldsymbol{p}\otimes]\boldsymbol{q} \\ &= \begin{bmatrix} p_s & p_z & -p_y & p_x \\ -p_z & p_s & p_x & p_y \\ p_y & -p_x & p_s & p_z \\ -p_x & -p_y & -p_z & p_s \end{bmatrix} \begin{bmatrix} q_x \\ q_y \\ q_z \\ q_s \end{bmatrix} \end{aligned} \tag{5}$$

(Shuster [2] defines the less obvious notation $\{\boldsymbol{q}\}_L \equiv [\boldsymbol{q}\otimes]$.) For unit-norm quaternions, the quaternion product corresponds to the DCM product

$$\boldsymbol{r} = \boldsymbol{p} \otimes \boldsymbol{q} \Longleftrightarrow \mathbf{T}_{\boldsymbol{r}} = \mathbf{T}_{\boldsymbol{p}}\mathbf{T}_{\boldsymbol{q}}$$

where the order of corresponding terms is the same, whereas in comparison with the hypercomplex product

$$\boldsymbol{p} \otimes \boldsymbol{q} \Longleftrightarrow \tilde{\boldsymbol{q}}\tilde{\boldsymbol{p}}$$

the order of corresponding terms is reversed. Define

$$\varphi = |\boldsymbol{\phi}| \qquad \bar{\boldsymbol{\phi}} = \begin{bmatrix} \boldsymbol{\phi} \\ 0 \end{bmatrix} \qquad \boldsymbol{q} = \begin{bmatrix} \frac{1}{2}\boldsymbol{\phi}\,\dfrac{\sin(\varphi/2)}{\varphi/2} \\ \cos(\varphi/2) \end{bmatrix} \tag{6}$$

Some useful identities follow:

$$(\boldsymbol{p} \otimes \boldsymbol{q})^* = \boldsymbol{q}^* \otimes \boldsymbol{p}^* \tag{7a}$$

$$[\boldsymbol{q}\otimes][\boldsymbol{q}^*\otimes] = [\boldsymbol{q}\otimes][\boldsymbol{q}\otimes]^T = |\boldsymbol{q}|^2\boldsymbol{I} \tag{7b}$$

$$[\boldsymbol{q}^{-1}\otimes] = [\boldsymbol{q}\otimes]^{-1} = [\boldsymbol{q}\otimes]^T/|\boldsymbol{q}|^2 \tag{7c}$$

$$[\boldsymbol{q}\otimes] = \exp\{\tfrac{1}{2}[\bar{\boldsymbol{\phi}}\otimes]\} = \cos(\varphi/2)\boldsymbol{I} + \tfrac{1}{2}[\bar{\boldsymbol{\phi}}\otimes]\,\frac{\sin(\varphi/2)}{\varphi/2} \tag{7d}$$

5 The $\circledast$ Operator

The $\circledast$ operator and operator matrix $[\boldsymbol{q}\circledast]$ are defined by

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{q} \circledast \boldsymbol{p} \\ &= [\boldsymbol{q}\circledast]\boldsymbol{p} \\ &= \begin{bmatrix} q_s & -q_z & q_y & q_x \\ q_z & q_s & -q_x & q_y \\ -q_y & q_x & q_s & q_z \\ -q_x & -q_y & -q_z & q_s \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ p_s \end{bmatrix} \end{aligned} \tag{8}$$

(Shuster [2] defines the less obvious notation $\{\boldsymbol{q}\}_R \equiv [\boldsymbol{q}\circledast]$.) For unit-norm quaternions, the quaternion product corresponds to the DCM product

$$\boldsymbol{r} = \boldsymbol{q} \circledast \boldsymbol{p} \Longleftrightarrow \mathbf{T}_{\boldsymbol{r}} = \mathbf{T}_{\boldsymbol{p}}\mathbf{T}_{\boldsymbol{q}} \tag{9}$$

where the order of corresponding terms is reversed, whereas in comparison with the hypercomplex product

$$\boldsymbol{q} \circledast \boldsymbol{p} \Longleftrightarrow \tilde{\boldsymbol{q}}\tilde{\boldsymbol{p}} \tag{10}$$

the order of corresponding terms is the same. Some useful identities are

$$(\boldsymbol{q} \circledast \boldsymbol{p})^* = \boldsymbol{p}^* \circledast \boldsymbol{q}^* \tag{11a}$$

$$[\boldsymbol{q}\circledast][\boldsymbol{q}^*\circledast] = [\boldsymbol{q}\circledast][\boldsymbol{q}\circledast]^T = |\boldsymbol{q}|^2 \boldsymbol{I} \tag{11b}$$

$$[\boldsymbol{q}^{-1}\circledast] = [\boldsymbol{q}\circledast]^{-1} = [\boldsymbol{q}\circledast]^T / |\boldsymbol{q}|^2 \tag{11c}$$

$$[\boldsymbol{q}\circledast] = \exp\{\tfrac{1}{2}[\bar{\boldsymbol{\phi}}\circledast]\} = \cos(\varphi/2)\boldsymbol{I} + \tfrac{1}{2}[\bar{\boldsymbol{\phi}}\circledast]\,\frac{\sin(\varphi/2)}{\varphi/2} \tag{11d}$$

$$[\boldsymbol{q}\otimes][\boldsymbol{q}\circledast]^{-1} = [\boldsymbol{q}\circledast]^{-1}[\boldsymbol{q}\otimes] = \begin{bmatrix} \mathbf{T}_{\boldsymbol{q}} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{11e}$$

$$\boldsymbol{p} \otimes \boldsymbol{q} = \boldsymbol{q} \circledast \boldsymbol{p} \tag{11f}$$

where φ, $\bar{\boldsymbol{\phi}}$, and $\boldsymbol{q}$ as a function of $\boldsymbol{\phi}$ were defined in equation (6).

The $[\boldsymbol{q}\circledast]$ matrix can be partitioned such that $[\boldsymbol{q}\circledast] = [\Xi \;\; \boldsymbol{q}]$, where Ξ is a 4×3 matrix. We will write $\Xi(\boldsymbol{p})$ for the partition of $[\boldsymbol{p}\circledast]$, but otherwise assume that Ξ without an argument depends on $\boldsymbol{q}$ in order to simplify the notation. Since $[\boldsymbol{q}\circledast]$ is an orthonormal matrix for $|\boldsymbol{q}| = 1$, we have that $\Xi\Xi^T + \boldsymbol{q}\boldsymbol{q}^T = |\boldsymbol{q}|^2\boldsymbol{I}$, $\Xi^T\Xi = |\boldsymbol{q}|^2\mathbf{I}$, and $\Xi^T\boldsymbol{q} = \mathbf{0}$.

6 Quaternion Transformation of a Vector

Let $\boldsymbol{v}^a$ be a vector in a frame a and $\boldsymbol{v}^b$ be a vector in a frame b. If a quaternion $\boldsymbol{q}$ represents the attitude of frame b with respect to frame a, the $\boldsymbol{v}^a$ is transformed to the frame b by

$$\bar{\boldsymbol{v}}^b = \boldsymbol{q} \otimes \bar{\boldsymbol{v}}^a \otimes \boldsymbol{q}^{-1} \tag{12a}$$

$$= \boldsymbol{q}^{-1} \circledast \bar{\boldsymbol{v}}^a \circledast \boldsymbol{q} \tag{12b}$$

where $\bar{\boldsymbol{v}}^a = \begin{bmatrix} \boldsymbol{v}^a \\ 0 \end{bmatrix}$ and $\bar{\boldsymbol{v}}^b = \begin{bmatrix} \boldsymbol{v}^b \\ 0 \end{bmatrix}$. The equivalence of equations (12a) and (12b) follows from equation (11f). From equation (11e), we have that equations (12a) and (12b) are the same as

$$\begin{bmatrix} \boldsymbol{v}^b \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{\boldsymbol{q}} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^a \\ 0 \end{bmatrix} . \tag{13}$$

7 Associativity and Commutativity

Quaternion operators of the same kind are associative, that is,

$$\boldsymbol{p} \otimes \boldsymbol{q} \otimes \boldsymbol{r} = (\boldsymbol{p} \otimes \boldsymbol{q}) \otimes \boldsymbol{r} = \boldsymbol{p} \otimes (\boldsymbol{q} \otimes \boldsymbol{r}) \tag{14a}$$

$$\boldsymbol{p} \circledast \boldsymbol{q} \circledast \boldsymbol{r} = (\boldsymbol{p} \circledast \boldsymbol{q}) \circledast \boldsymbol{r} = \boldsymbol{p} \circledast (\boldsymbol{q} \circledast \boldsymbol{r}) \tag{14b}$$

However, quaternion operators that are not the same kind are not associative, that is,

$$(\boldsymbol{p} \otimes \boldsymbol{q}) \circledast \boldsymbol{r} \neq \boldsymbol{p} \otimes (\boldsymbol{q} \circledast \boldsymbol{r}) \tag{15a}$$

$$(\boldsymbol{p} \circledast \boldsymbol{q}) \otimes \boldsymbol{r} \neq \boldsymbol{p} \circledast (\boldsymbol{q} \otimes \boldsymbol{r}) \tag{15b}$$

so the expressions $\boldsymbol{p} \otimes \boldsymbol{q} \circledast \boldsymbol{r}$ and $\boldsymbol{p} \circledast \boldsymbol{q} \otimes \boldsymbol{r}$ are ambiguous.

Equation (11f) indicates that quaternions are, in general, not commutative in their product. They are commutative only if their axes of rotation are in the same direction, which is also true of direction cosine matrices.

8 Parameterizations of the Attitude Matrix

Equation (11e) shows one way to compute the attitude matrix $\mathbf{A}(\boldsymbol{q}) = \mathbf{T}_{\boldsymbol{q}}$, namely

$$\begin{aligned} \mathbf{A}(\boldsymbol{q}) &= (s^2 - |\boldsymbol{r}|^2)\mathbf{I} - 2s[\boldsymbol{r}\times] + 2\boldsymbol{r}\boldsymbol{r}^T && \text{(16a)} \\ &= (s^2 + |\boldsymbol{r}|^2)\mathbf{I} - 2s[\boldsymbol{r}\times] + 2[\boldsymbol{r}\times]^2 && \text{(16b)} \end{aligned}$$

where the quaternion is partitioned into its vector part $\boldsymbol{r}$ and a scalar part s,

$$\boldsymbol{q} = \begin{bmatrix} \boldsymbol{r} \\ s \end{bmatrix} . \tag{17}$$

The attitude matrix can also be written in terms of a rotation vector $\boldsymbol{\phi}$ with $\varphi = |\boldsymbol{\phi}|$,

$$\begin{aligned} \mathbf{A}(\boldsymbol{\phi}) = \exp(-[\boldsymbol{\phi}\times]) &= (\cos\varphi)\mathbf{I} - \frac{\sin\varphi}{\varphi}[\boldsymbol{\phi}\times] + \frac{1-\cos\varphi}{\varphi^2}\boldsymbol{\phi}\boldsymbol{\phi}^T && \text{(18a)} \\ &= \mathbf{I} - \frac{\sin\varphi}{\varphi}[\boldsymbol{\phi}\times] + \frac{1-\cos\varphi}{\varphi^2}[\boldsymbol{\phi}\times]^2 && \text{(18b)} \end{aligned}$$

References

[1] WERTZ, J. R., ed., *Spacecraft Attitude Determination and Control*, Kluwer Academic Publishers, 1978, pp. 410–416, 510–512, 564–565, 755–762.

[2] SHUSTER, M. D., "A Survey of Attitude Representations", The Journal of the Astronautical Sciences, Vol. 41, No. 4, October–December 1993, pp. 439–517.

[3] SHUSTER, M. D., "A Simple Kalman Filter and Smoother for Spacecraft Attitude", The Journal of the Astronautical Sciences, Vol. 37, No. 1, January–March, 1989, pp. 89–106.

[4] LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D., "Kalman Filtering for Spacecraft Attitude Estimation", *AIAA J. Guidance, Control, Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429.