

The Nature of the Quaternion

Malcolm D. Shuster¹

“I regard it as an inelegance, or imperfection, in quaternions, or rather in the state to which it has been hitherto unfolded, whenever it becomes or seems to become necessary to have recourse to x, y, z , etc.”

William Rowan Hamilton (quoted in a letter from Tait to Cayley)

“Quaternions came from Hamilton after his really good work had been done; and, though beautifully ingenious, have been an unmixed evil to those who have touched them in any way, including Clerk Maxwell.”

William Thomson, first baron Kelvin, 1892

“...quaternions appear to exude an air of nineteenth century decay, as a rather unsuccessful species in the struggle-for-life of mathematical ideas. Mathematicians, admittedly, still keep a warm place in their hearts for the remarkable algebraic properties of quaternions but, alas, such enthusiasm means little to the harder-headed physical scientist.”

Simon L. Altmann, 1986 [1]

Abstract

Some of the confusions concerning quaternions as they are employed in spacecraft attitude work are discussed. The order of quaternion multiplication is discussed in terms of its historical development and its consequences for the quaternion imaginaries. The different formulations for the quaternions are also contrasted. It is shown that the three Hamilton imaginaries cannot be interpreted as the basis of the vector space of physical vectors but only as constant numerical column vectors, the autorepresentation of a physical basis.

Introduction

The quaternion [2] is one of the most important representations of the attitude in spacecraft attitude estimation and control. Under the various guises of the Euler symmetric parameters, the Rodrigues symmetric parameters, the Euler-Rodrigues symmetric parameters, the Cayley-Klein parameters, and of course, the quaternion, quaternions have been in existence for nearly two-and-a-half centuries, longer, in fact, than the direction-cosine matrix. For a brief

historical discussion of the quaternion and other attitude representations with references, see reference [2], pp. 495–498. After such a long passage of time, the quaternion should be well understood and free of ambiguities. Surprisingly, the truth is different, and one of the most important inconsistencies has arisen during the past 30 years. The two most important confusions concern the order of quaternion multiplication and the nature of the quaternion “imaginaries,” both of which are the subject of this article.

Very little is derived in this article. A complete, detailed and consistent formulation of the quaternion as a 4×1 matrix appeared in reference [2]. That work has been cited very frequently within the astrodynamics community over the past fifteen years, and its formulation seems to have become standard there. The traditional presentation of the quaternion can be found in great detail and in great extent in reference [3]. The notation of this article is that of references [2]. The literature on quaternions is plentiful. A chapter on quaternions can be found in nearly every advanced textbook which treats rotational mechanics. For convenience, we limit ourselves to citations of reference [2] whenever possible.

An Historical Perspective

The Euler symmetric parameters appeared first in publication in 1770 [4], five years before Euler’s formula for the direction-cosine matrix [5]. However, the Euler symmetric parameters weren’t developed as a parameterization of rotations. Euler (1707–1783) was interested simply in finding a parameterization of any orthogonal matrix, and examined specifically square matrices of dimension 3, 4 and 5. His published results in 1770 for the 3×3 orthogonal matrix were only for improper orthogonal matrices. Euler was hardly interested in rotations at the time.

The quaternion as a static attitude representation was more fully developed by Rodrigues (1795–1851) in 1840 [6]. Rodrigues discovered the connection (the Rodrigues symmetric parameters) to the attitude matrix, a geometrical multiplication rule for quaternions, and the Rodrigues vector, often called the Gibbs vector, after J. Willard Gibbs, who popularized its use some 60 years later [7].

Hamilton (1805–1865) enters the picture in 1843, only a few years after Rodrigues. While Rodrigues was interested in attitude, Hamilton’s primary interest was in developing a theory of hypercomplex numbers, that is, in extending the complex-number system. Rodrigues’ interest was in Geometry, Hamilton’s in Algebra. For a long time, Hamilton struggled with a system with only two imaginary numbers, i and j , but ran into problems when trying to find an expression for ij as the linear combination of 1, i and j , a task in which he never succeeded. He arrived finally at the idea that ij must be a separate imaginary k , and the quaternion algebra with its three imaginaries was born [8]. The principal interest of algebraists today in Hamilton’s discovery was that the field generated by the set $\{1, i, j, k\}$, together with the binary operations of hypercomplex addition and multiplication, was the first example of a *skew field*, a field in which multiplication is not commutative. With the advent of matrices and matrix multiplication, skew fields became a commonplace. It is Hamilton who coined the word *quaternion*. (For a linguistic discussion of “quaternion,”

see footnote 4 of reference [9].) For us, the interest in quaternions comes from the identification of the Hamiltonian imaginaries with coordinate axes and their utility in describing rotations.

Hamilton never acknowledged the work of Rodrigues, and his use of quaternions as a description of rotations was very wrong [2]. He believed that the expression for a rotated vector was linear in the quaternion rather than quadratic. It is Cayley (1821–1895) whom we must thank for the further development of quaternions as a representation of attitude. It is Cayley who is responsible for the familiar kinematic equation for the quaternion [10],

$$\frac{d}{dt} \bar{\eta}(t) = \Xi(\bar{\eta}(t)) \boldsymbol{\omega}(t) \quad (1)$$

where $\boldsymbol{\omega}$ is the angular-velocity column vector, and $\Xi(\bar{\eta})$ is the matrix

$$\Xi(\bar{\eta}) = \begin{bmatrix} \eta_4 & -\eta_3 & \eta_2 \\ \eta_3 & \eta_4 & -\eta_1 \\ -\eta_2 & \eta_1 & \eta_4 \\ -\eta_1 & -\eta_2 & -\eta_3 \end{bmatrix} \quad (2)$$

It was Cayley, in fact, who invented matrix algebra in 1855 [11, 12]. The formula for matrix multiplication

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj} \quad (3)$$

appears first in Cayley's work.

Hamilton's error concerning rotations arose from the analogy of the Argand diagram for complex numbers with the geometrical plane. In the Argand diagram, multiplication of a complex number by i is equivalent to a "rotation" of the complex number in the Argand diagram by $\pi/2$. Hamilton then concluded that rotation using quaternions must be a operation linear in the quaternion in general, although both he and Cayley had showed that the operation was necessarily bilinear [13]. In fact, "rotation by i " in the Argand diagram was more akin to rotation by a rotation matrix. Despite this, Hamilton persisted in his error, and the statement is to be found even in his final great (posthumous) work on quaternions [14]. Hamilton's quaternion ideas became something of a cult after his death, with the rôle of high priest filled by the Scottish mathematician Peter Guthrie Tait (1831–1901), one of many mathematicians whose name has been attached to the antisymmetric Euler angles [2]. Hamilton's mistaken ideas concerning quaternions and rotations could not, of course, survive the need for useful applications and are now forgotten, except by historians of science. Today, Hamilton's approach to rotations belongs with phlogiston, the luminiferous ether, the Dean drive and cold fusion.

Quaternions had largely disappeared from view by the beginning of the twentieth century, having been replaced by the simpler and more transparent vectors, which first appeared in the works of J. Willard Gibbs (1839–1903) [15] and Oliver Heaviside (1850–1925) [16]. About the time, matrices were beginning to receive wide application in Engineering, particularly in studies of

elasticity, and in Physics, especially with the advent of Matrix Mechanics, an early formulation of Quantum Mechanics, in 1925. This fading of interest in quaternions is responsible for one of the inconsistencies that occurs in the two current presentations of quaternions.

The quotation of Hamilton's at the beginning of this article is revealing about quaternions. When Hamilton invented his quaternion in terms of imaginaries, it was the first time, perhaps, that a single symbol substituted itself generally for a group of components. One might today have made the same comment about vectors. One of the great advances of Hamilton's work on quaternions was the use of a single symbol to denote a vector as a quaternion which is a linear combination only of the imaginaries. Hamilton's displeasure expressed in the opening quotation of this article is similar to our own when we look at works in Physics from 150 years ago, in which every component of a vector and every entry in a matrix is expressed by a different letter of the alphabet.

For greater detail on the history of vectors and quaternions the reader is referred to references [1], [17] and [18].

The Problem of Quaternion Multiplication

Beginning with Hamilton, quaternions have been multiplied in the opposite order than rotation matrices are today. Thus, in the traditional formulation of quaternions, one writes

$$R(\vec{\eta}') R(\vec{\eta}) = R(\vec{\eta} \circ \vec{\eta}') \quad (4)$$

Here, R denotes the rotation matrix (attitude matrix, direction-cosine matrix), $\vec{\eta}$ the quaternion, and “ \circ ” the multiplication operation for quaternions.²As in reference [2], we use $\vec{\eta}$ to denote an element of the multiplication group of quaternions with unit norm (the quaternions of rotation) and \bar{q} to denote an element of the quaternion algebra, in which the quaternion may have any norm. We refer to the order of quaternion multiplication in equation (4) as the *traditional order*. This is the order found in reference [3].

More recently, in spacecraft work, the order of quaternion multiplication has been chosen to satisfy

$$R(\vec{\eta}') R(\vec{\eta}) = R(\vec{\eta}' \circ \vec{\eta}) \quad (5)$$

which we call the *natural order* in this article, because the order of quaternion multiplication is the same as that of matrix multiplication. This is the order of reference [2]. The “natural” order for quaternion multiplication seems more reasonable within the framework of spacecraft attitude studies, where both the rotation matrix and the quaternion receive frequent use. We shall refer to the two formulations of the the quaternion as the *traditional* formulation and the imore recent formulation.

Had the quaternion still been of great interest in the early twentieth century, the natural order might have been adopted for quaternions a century ago, but Tait, the great champion of quaternions after Hamilton, was gone by then, and quaternions remained stagnant.

² R is a proper direction-cosine matrix, that is, a direction-cosine of positive determinant. We sometimes call R the attitude matrix and denote it by A .

The confusion in quaternion multiplication is most apparent in the presentations of quaternion multiplication in the book, *Spacecraft Attitude Determination and Control* [19], which appeared in 1978, and which was an important landmark in the development of spacecraft attitude estimation [20]. Reference [19] bestrides the transition from traditional order to natural order in quaternion multiplication. In the appendix of reference [19] on “Quaternions” [21], the quaternions are presented in a manner consistent with equation (4), the traditional order, while in the section within the main text of reference [19] on “Parameterization of the Attitude” [22], quaternion multiplication is presented in a manner consistent with equation (5), the natural order.³ Due to the brevity of references [21] and [22] and the non-identical selection of material, the two presentations are not trivially comparable, so it has not been necessarily evident to readers that they are mutually inconsistent.⁴ Reference [2] is consistent with equation (5). Of the more recent texts, four [23–26] follow the conventions of reference [2] and cite it, and one [3] follows the traditional approach typified by equation (4). The material on the quaternion in reference [27] is insufficient to determine which formulation was employed. Not all of these works state equation (4) or (5) explicitly, and the nature of their approach to quaternion multiplication must be inferred from other equations in those works, in some cases from only one equation.

Thus, one dilemma of the quaternion. Many authors, especially first authors on spacecraft attitude, take equations from both parts of *Spacecraft Attitude Determination and Control* [19] and from other works, unaware of the inconsistency. One purpose of the present work is to retrace the steps of the traditional formulation of quaternions critically, and then to present the more recent approach, which is more in tune with the more modern needs of attitude studies created by the application to spacecraft. There are several ramifications of the difference between equations (4) and (5), particularly in the nature of the fundamental relationship of traditional quaternion multiplication $\mathbf{ij} = \mathbf{k}$, discovered by Hamilton in 1843.

The second dilemma concerns the very nature of \mathbf{i} , \mathbf{j} , and \mathbf{k} in quaternion theory. Are these the directions of real coordinate axes or something else? The connection of the multiplication rule for Hamilton’s three imaginaries and the multiplication rule for rotations will be carefully explored in this work.

Hamilton’s approach to quaternions leads to equation (4), seemingly in universal use until the publication of reference [22], and still in almost universal use until the publication of reference [2], which, probably, more than any other work, has been responsible for the change to the natural order of quaternion multiplication in spacecraft attitude estimation and control. This was, in fact, an avowed purpose of the author of reference [2]. But although nearly every writer on spacecraft attitude is aware now of reference [2], which is cited

³Markley claims that he first encountered the natural order of quaternion multiplication in the technical report of a prime contractor.

⁴The author of the present work may be partly responsible for the simultaneous appearance of the two inconsistent treatments in reference [19], since he was an official proofreader for *Spacecraft Attitude Determination and Control* in 1977 and 1978. However, his proofreading took place during his very first year in Engineering, and while he was able to find many typographical errors, fundamental inconsistencies obviously eluded him.

frequently, he or she may not be aware of the inconsistency of reference [2] with many other works on quaternions.

The impetus, at least for the present writer, to change to the natural order for quaternion multiplication came not only from aesthetic considerations but also from the needs of real mission support in spacecraft attitude determination and control, especially, in spacecraft attitude determination. In purely theoretical studies of spacecraft attitude dynamics and control, one seldom has more than one quaternion, because one seldom has more than two reference frames, inertial and body. In theoretical studies of spacecraft attitude estimation, which includes also the study of sensor alignment estimation, the number of reference frames is increased by the number of sensor reference frames, which can be large. This cannot compare, however, with the number of reference frames needed in real spacecraft attitude mission support, which may be several dozen. A large part of the functional-specification document for spacecraft attitude mission support software consists of the specification of the very large number of reference frames and the transformations connecting them. System development will be less prone to error if the multiplication rule for quaternions has the same order as that for the corresponding rotation matrices.

Vectors and Attitude

As in reference [2], we distinguish between physical vectors, their column-vector representations, and numerical column vectors, which have a constant value independent of the choice of basis.

A *physical vector*, as defined in this article, is coordinate-free. It is not composed of three components and is specified by some physical property (for example, as the position vector of a particle) or in terms of other physical vectors. If we assume that the vector space \mathcal{V} of physical vectors has a scalar product (inner product, dot product), that is, that \mathcal{V} is an inner-product space, then it also possess a right-handed orthonormal basis [2], which we denote by $\mathcal{E} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$.

We define the components of a physical vector \mathbf{u} with respect to the basis \mathcal{E} as

$$\varepsilon_{u_k} \equiv \hat{\mathbf{e}}_k \cdot \mathbf{u}, \quad k = 1, 2, 3 \quad (6)$$

where “ \cdot ” denotes the scalar product operation. The *column vector representation* of \mathbf{u} with respect to \mathcal{E} is then

$${}^{\mathcal{E}}\mathbf{u} \equiv \begin{bmatrix} \varepsilon_{u_1} \\ \varepsilon_{u_2} \\ \varepsilon_{u_3} \end{bmatrix} \quad (7)$$

For a given \mathcal{E} , there is an isomorphism between \mathcal{V} and ${}^{\mathcal{E}}\mathcal{V}$, the vector space of column-vector representations with respect to \mathcal{E} [2].

Of particular importance is the *autorepresentation* of a basis, that is, the representation of a basis with respect to itself. Clearly,

$${}^{\mathcal{E}}\hat{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \equiv \hat{\mathbf{1}}, \quad {}^{\mathcal{E}}\hat{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \equiv \hat{\mathbf{2}}, \quad {}^{\mathcal{E}}\hat{\mathbf{e}}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \equiv \hat{\mathbf{3}}, \quad (8abc)$$

The autorepresentation of a basis is the same for every basis, even for bases which are not orthonormal. The *numerical column vector* is a column vector which is specified entirely by the specific values of its entries. The vectors $\hat{\mathbf{1}}$, $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are the only numerical column vectors that appear in this work.

The transformation of a physical vector \mathbf{a} into a physical vector \mathbf{b} of equal magnitude under the influence of a (counterclockwise) rotation about a physical axis vector $\hat{\mathbf{n}}$ through an angle θ is given by [2]

$$\begin{aligned} \mathbf{b} &= (\cos \theta) \mathbf{a} + (1 - \cos \theta) (\hat{\mathbf{n}} \cdot \mathbf{a}) \hat{\mathbf{n}} + (\sin \theta) \hat{\mathbf{n}} \times \mathbf{a} \\ &= \mathbf{a} + (\sin \theta) \hat{\mathbf{n}} \times \mathbf{a} + (1 - \cos \theta) \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{a}). \end{aligned} \quad (9)$$

This is Euler's formula [5].

The *attitude matrix* $\mathbf{A}^{\mathcal{E}'/\mathcal{E}}$ transforming a right-handed orthonormal basis \mathcal{E} into a right-handed orthonormal basis \mathcal{E}' is defined as the direction-cosine matrix

$$A_{ij}^{\mathcal{E}'/\mathcal{E}} \equiv \hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j \quad (10)$$

It follows that

$$\hat{\mathbf{e}}'_i = \sum_{j=1}^3 A_{ij}^{\mathcal{E}'/\mathcal{E}} \hat{\mathbf{e}}_j, \quad i = 1, 2, 3 \quad (11)$$

The transformation of column-vector representations under the same rotation is given by [2]

$${}^{\mathcal{E}'}\mathbf{u} = \mathbf{A}^{\mathcal{E}'/\mathcal{E}} {}^{\mathcal{E}}\mathbf{u} \quad (12)$$

Note also

$${}^{\mathcal{E}'}\hat{\mathbf{e}}'_i = \sum_{j=1}^3 A_{ij}^{\mathcal{E}'/\mathcal{E}} {}^{\mathcal{E}}\hat{\mathbf{e}}_j, \quad i = 1, 2, 3 \quad (13)$$

$${}^{\mathcal{E}'}\hat{\mathbf{e}}_i = \mathbf{A}^{\mathcal{E}'/\mathcal{E}} {}^{\mathcal{E}}\hat{\mathbf{e}}_i, \quad i = 1, 2, 3 \quad (14)$$

$${}^{\mathcal{E}'}\hat{\mathbf{e}}'_i = {}^{\mathcal{E}}\hat{\mathbf{e}}_i, \quad i = 1, 2, 3 \quad (15)$$

In a lax notation one might wish to call all the left members of equations (13) through (15) $\hat{\mathbf{e}}'_i$, $i = 1, 2, 3$. It is important always to denote the basis of representation in the symbol for a column-vector representation. Note that the operation in the right member of equation (13) is the multiplication of a matrix by a scalar, the operation in equation (14) is matrix multiplication.

From equation (9) with \mathbf{a} replaced by $\hat{\mathbf{e}}_k$ and \mathbf{b} replaced by $\hat{\mathbf{e}}_k$, successively for $k = 1, 2, 3$, it follows that for a rotation through an angle θ counterclockwise about a physical rotation vector $\hat{\mathbf{n}}$, the attitude matrix is given by [2]

$$\begin{aligned} \mathbf{A} &= (\cos \theta) \mathbf{I}_{3 \times 3} + (1 - \cos \theta) \hat{\mathbf{n}} \hat{\mathbf{n}}^T - (\sin \theta) [\hat{\mathbf{n}} \times] \\ &= \mathbf{I}_{3 \times 3} - (\sin \theta) [\hat{\mathbf{n}} \times] + (1 - \cos \theta) [\hat{\mathbf{n}} \times]^2 \end{aligned} \quad (16)$$

where $\mathbf{I}_{3 \times 3}$ is the 3×3 identity matrix, given by

$$\mathbf{I}_{3 \times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (17)$$

and the antisymmetric matrix $[\hat{\mathbf{n}} \times]$ is defined by

$$[\mathbf{u} \times] \equiv \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \quad (18)$$

for any column vector \mathbf{u} . In terms of individual entries, equation (15) can be written as

$$\mathbf{A} = \begin{bmatrix} c + n_1^2(1-c) & n_1 n_2(1-c) + n_3 s & n_1 n_3(1-c) - n_2 s \\ n_2 n_1(1-c) - n_3 s & c + n_2^2(1-c) & n_2 n_3(1-c) + n_1 s \\ n_3 n_1(1-c) + n_2 s & n_3 n_2(1-c) - n_1 s & c + n_3^2(1-c) \end{bmatrix} \quad (19)$$

where $c \equiv \cos \theta$ and $s \equiv \sin \theta$. We have not written the basis of representation on $\hat{\mathbf{n}}$, because

$${}^{\mathcal{E}}\hat{\mathbf{n}} = {}^{\mathcal{E}'}\hat{\mathbf{n}} \quad (20)$$

since

$$\mathbf{A}(\hat{\mathbf{n}}, \theta) \hat{\mathbf{n}} = \hat{\mathbf{n}} \quad (21)$$

See reference [2] for further details.

Note that

$$\mathbf{u} = {}^{\mathcal{E}}u_1 \hat{\mathbf{e}}_1 + {}^{\mathcal{E}}u_2 \hat{\mathbf{e}}_2 + {}^{\mathcal{E}}u_3 \hat{\mathbf{e}}_3, \quad (22a)$$

$${}^{\mathcal{E}}\mathbf{u} = {}^{\mathcal{E}}u_1 {}^{\mathcal{E}}\hat{\mathbf{e}}_1 + {}^{\mathcal{E}}u_2 {}^{\mathcal{E}}\hat{\mathbf{e}}_2 + {}^{\mathcal{E}}u_3 {}^{\mathcal{E}}\hat{\mathbf{e}}_3, \quad (22b)$$

$$= {}^{\mathcal{E}}u_1 \hat{\mathbf{i}} + {}^{\mathcal{E}}u_2 \hat{\mathbf{j}} + {}^{\mathcal{E}}u_3 \hat{\mathbf{k}} \quad (22c)$$

Despite the similarities of equations (22a) and (22b) in appearance, they are very different in nature.

The Traditional Formulation of Quaternions

Quaternion Algebra

Since Hamilton, the quaternions have consisted of the field generated by the real numbers and the three imaginaries \mathbf{i} , \mathbf{j} and \mathbf{k} . The multiplication rule for the imaginaries, as proposed by Hamilton, is

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j} \quad (23abc)$$

$$\mathbf{i}\mathbf{i} = \mathbf{j}\mathbf{j} = \mathbf{k}\mathbf{k} = \mathbf{i}\mathbf{j}\mathbf{k} = -1 \quad (24)$$

and naturally,

$$1\mathbf{i} = \mathbf{i}1 = \mathbf{i}, \quad 1\mathbf{j} = \mathbf{j}1 = \mathbf{j}, \quad 1\mathbf{k} = \mathbf{k}1 = \mathbf{k} \quad (25abc)$$

$$0\mathbf{i} = \mathbf{i}0 = \mathbf{0}, \quad 0\mathbf{j} = \mathbf{j}0 = \mathbf{0}, \quad 0\mathbf{k} = \mathbf{k}0 = \mathbf{0} \quad (26abc)$$

Equations (23) bear a striking resemblance to those for the vector product. The most general quaternion has the form

$$\bar{q} = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} + q_4 \quad (27)$$

The conjugate quaternion \bar{q}^* is defined as

$$\bar{q}^* \equiv -q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k} + q_4 \quad (28)$$

and

$$\bar{q}^*\bar{q} = q_1^2 + q_2^2 + q_3^2 + q_4^2 \equiv |\bar{q}|^2 \quad (29)$$

The quantity $|\bar{q}|$ is the *length* of the quaternion. The identity element of the quaternion (skew) field is 1 and the inverse quaternion is given by

$$\bar{q}^{-1} = |\bar{q}|^{-2} \bar{q}^* \quad (30)$$

The imaginaries \mathbf{i} , \mathbf{j} and \mathbf{k} are interpreted generally also as the basis vectors of a right-handed orthogonal physical vector space. Thus, we have also

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1 \quad (31)$$

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i} = 0, \quad \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{j} = 0, \quad \mathbf{k} \cdot \mathbf{i} = \mathbf{i} \cdot \mathbf{k} = 0 \quad (32abc)$$

and

$$\mathbf{i} \times \mathbf{j} = -\mathbf{j} \times \mathbf{i} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = -\mathbf{k} \times \mathbf{j} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = -\mathbf{i} \times \mathbf{k} = \mathbf{j} \quad (33abc)$$

The quaternion space in the traditional formulation contains the real numbers (scalars) q_4 and \mathcal{V} , the physical vector space in three dimensions, with elements

$$\mathbf{q} = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \quad (34)$$

The general quaternion may be written

$$\bar{q} = \mathbf{q} + q_4 \quad (35)$$

In the traditional formulation of quaternions, the binary multiplicative operation does not have a special symbol but is simply written by juxtaposing the two quaternions, as is done for matrices.

Rotations

For a rotation, one has

$$\boldsymbol{\eta} = \hat{n} \sin(\theta/2) = \eta_1\mathbf{i} + \eta_2\mathbf{j} + \eta_3\mathbf{k} \quad (36a)$$

X **Rev 1g** **Shuster**

$$\eta_4 = \cos(\theta/2) \quad (36b)$$

$$\bar{\eta} = \boldsymbol{\eta} + \eta_4 \quad (36c)$$

with $\hat{\boldsymbol{n}}$ and θ the axis and angle of rotation.

$$\hat{\boldsymbol{n}} = n_1 \mathbf{i} + n_2 \mathbf{j} + n_3 \mathbf{k} \quad (37)$$

$$n_1^2 + n_2^2 + n_3^2 = 1 \quad (38)$$

Thus, $|\bar{\eta}| = 1$. Rotation by quaternions in the traditional formulation is effected by the quaternion as in references [3] and [21]

$$\bar{\eta}^* \mathbf{u} \bar{\eta} = \mathbf{u}' \quad (39)$$

If, following equation (22b), we write the vector \mathbf{u} as

$$\mathbf{u} = \varepsilon_{u_1} \mathbf{i} + \varepsilon_{u_2} \mathbf{j} + \varepsilon_{u_3} \mathbf{k} \quad (40)$$

where \mathcal{E} is the prior basis, $\mathcal{E} = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ and ε_{u_1} , ε_{u_2} and ε_{u_3} denote the components of \mathbf{u} with respect to the basis vectors \mathbf{i} , \mathbf{j} and \mathbf{k} , respectively, then evaluation of equation (39) leads to

$$\mathbf{u}' = \varepsilon'_{u_1} \mathbf{i} + \varepsilon'_{u_2} \mathbf{j} + \varepsilon'_{u_3} \mathbf{k} \quad (41)$$

where \mathcal{E}' is the posterior basis, which satisfies equation (11). The result given by equation (41) does not correspond to the transformation of a physical vector, as given by equation (22a). If it did, one would have

$$\varepsilon_{\mathbf{u}'} = \varepsilon'_{\mathbf{u}}, \quad (42)$$

which is not true generally. However, if we interpret \mathbf{i} , \mathbf{j} , and \mathbf{k} as the numerical column vectors $\hat{\mathbf{1}}$, $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$, respectively, then we can rewrite equations (40) and (41) as

$$\mathbf{u} = \varepsilon_{u_1} \hat{\mathbf{1}} + \varepsilon_{u_2} \hat{\mathbf{2}} + \varepsilon_{u_3} \hat{\mathbf{3}} \quad (43a)$$

$$\mathbf{u}' \equiv \varepsilon'_{\mathbf{u}} = \varepsilon'_{u_1} \hat{\mathbf{1}} + \varepsilon'_{u_2} \hat{\mathbf{2}} + \varepsilon'_{u_3} \hat{\mathbf{3}} \quad (43b)$$

which make perfect sense, because

$$\varepsilon'_{\hat{\mathbf{e}}_1} = \varepsilon_{\hat{\mathbf{e}}_1} = \hat{\mathbf{1}}, \quad \varepsilon'_{\hat{\mathbf{e}}_2} = \varepsilon_{\hat{\mathbf{e}}_2} = \hat{\mathbf{2}}, \quad \varepsilon'_{\hat{\mathbf{e}}_3} = \varepsilon_{\hat{\mathbf{e}}_3} = \hat{\mathbf{3}} \quad (44)$$

Thus, we are forced to interpret the three Hamilton imaginaries, \mathbf{i} , \mathbf{j} and \mathbf{k} , not as physical basis vectors but as the autorepresentation of a physical basis.

If we consider the action of two successive quaternions on a vector \mathbf{u} (which we should write more correctly now as \mathbf{u} , but we will retain the traditional notation for the few remaining equations of this section), then

$$\mathbf{u}'' = \bar{\eta}''^* \mathbf{u}' \bar{\eta}'' = \bar{\eta}''^* (\bar{\eta}'^* \mathbf{u} \bar{\eta}') \bar{\eta}'' = (\bar{\eta} \bar{\eta}')^* \mathbf{u} (\bar{\eta} \bar{\eta}') \quad (45)$$

so that

$$\bar{\eta}'' = \bar{\eta} \bar{\eta}' \quad (46)$$

which is equivalent to equation (4).

The traditional formulation of quaternions presents many difficulties for modern applications in Engineering. First is the fact that the Hamilton imaginaries must be interpreted as column vectors rather than as physical vectors within a formulation that distinguishes between the two. This makes the addition of a scalar and a vector in the quaternion system impossible, because the addition of a scalar and a column vector is not allowed in matrix algebra. Secondly, the multiplication of quaternions is necessarily in the opposite order to that of rotation matrices, which is, at least, a nuisance. Thirdly, if we wish to adapt the traditional quaternion to computations and write

$$\bar{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \quad (47)$$

then one has

$$1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (48)$$

which equates a scalar to a 4×1 column vector. There is a certain enchantment to adding things that normally are not supposed to be added together, like scalars and vectors, but the system it leads to is simply wrong.

Hamilton would not have been bothered by the difficulties raised above, since matrix algebra was invented only after the publication of his *Lectures on Quaternions* [28]. In fact, we know from Tait (in his discussion on quaternions with Cayley in 1894 [29, 30]) that Hamilton had considered abandoning the imaginaries in favor of what must have been the ordered-list representation of quaternions, namely, the notation $\bar{q} = (q_1, q_2, q_3, q_4)$, but was unable in the end to abandon his beloved imaginaries.

The More Recent Formulation of Quaternions

Reference [2] take the easy way out of the problem of the limitations and confusions of the traditional formulation of quaternions and simply dispenses with i , j and k either as imaginaries or as basis vectors of physical space and instead considers the quaternion simply as a column vector in a four-dimensional quaternion column-vector space \mathcal{Q} . There is no need for the “vectors” of the quaternion space to be also vectors in physical space. It is sufficient that there be an isomorphism, as there is between physical vectors and their column-vector representations. As shown in the previous section, the quaternion acts on column-vector representations, not on physical vectors, and the axes of the space of column-vector representations are not the axes of the real physical vector space.

The four basis vectors of the quaternion vector space are

$$\bar{1} = \begin{bmatrix} \hat{1} \\ 0 \end{bmatrix}, \quad \bar{2} = \begin{bmatrix} \hat{2} \\ 0 \end{bmatrix}, \quad \bar{3} = \begin{bmatrix} \hat{3} \\ 0 \end{bmatrix}, \quad \bar{0} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad (49abc)$$

where $\hat{\mathbf{1}}$, $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are the familiar 3×1 matrices of equations (8). and

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (50)$$

We may write these four quaternion basis vectors also as $\bar{\mathbf{e}}_1$, $\bar{\mathbf{e}}_2$, $\bar{\mathbf{e}}_3$ and $\bar{\mathbf{e}}_4$. Any quaternion may be written then as

$$\begin{aligned} \bar{q} &= q_1 \bar{\mathbf{1}} + q_2 \bar{\mathbf{2}} + q_3 \bar{\mathbf{3}} + q_4 \bar{\mathbf{1}} \\ &= \sum_{k=1}^4 q_k \bar{\mathbf{e}}_k = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} \end{aligned} \quad (51)$$

There is no longer any mystical mystery of adding scalars and vectors. All elements of the quaternion space are 4×1 column vectors.

The quaternion multiplication rule has not yet been specified. Reference [2] determines that multiplication rule for quaternions by specifying the conditions that must be satisfied: (a) The quaternion of rotation must have the form

$$\boldsymbol{\eta} = \hat{\mathbf{n}} \sin(\theta/2), \quad \eta_4 = \cos(\theta/2) \quad (52ab)$$

where $\hat{\mathbf{n}}$ (more exactly, ${}^{\mathcal{E}}\hat{\mathbf{n}}^{\mathcal{E}'/\mathcal{E}}$) is the representation with respect to the prior basis \mathcal{E} of the axis of rotation for a rotation of the prior basis \mathcal{E} into the posterior basis \mathcal{E}' , and θ (more exactly $\theta^{\mathcal{E}'/\mathcal{E}}$) is the angle of rotation; (b) the multiplication of quaternions must satisfy equation (5); and (c) $\bar{\mathbf{1}}$ must be the identity element for quaternion multiplication. (Both $\bar{\mathbf{1}}$ and $-\bar{\mathbf{1}}$ correspond to the identity rotation matrix, but only one of these can be the identity element of the quaternion multiplication group.) These three conditions are sufficient to specify the multiplication rule for quaternions completely. The result [2] is

$$\bar{\eta}' \circ \bar{\eta} = \begin{bmatrix} \eta_4 \eta'_4 + \eta'_4 \eta - \boldsymbol{\eta}' \times \boldsymbol{\eta} \\ \eta'_4 \eta_4 - \boldsymbol{\eta}' \cdot \boldsymbol{\eta} \end{bmatrix}, \quad (53)$$

which may be written as

$$\bar{\eta}' \circ \bar{\eta} = \{ \bar{\eta}' \}_L \bar{\eta} = \{ \bar{\eta} \}_R \bar{\eta}'. \quad (54)$$

Explicitly,

$$\{ \bar{\eta} \}_L \equiv \begin{bmatrix} \eta_4 & \eta_3 & -\eta_2 & \eta_1 \\ -\eta_3 & \eta_4 & \eta_1 & \eta_2 \\ \eta_2 & -\eta_1 & \eta_4 & \eta_3 \\ -\eta_1 & -\eta_2 & -\eta_3 & \eta_4 \end{bmatrix}, \quad \{ \bar{\eta} \}_R \equiv \begin{bmatrix} \eta_4 & -\eta_3 & \eta_2 & \eta_1 \\ \eta_3 & \eta_4 & -\eta_1 & \eta_2 \\ -\eta_2 & \eta_1 & \eta_4 & \eta_3 \\ -\eta_1 & -\eta_2 & -\eta_3 & \eta_4 \end{bmatrix}. \quad (55ab)$$

From the above, it follows that the multiplication rule for the the basis vectors of the 4×1 quaternion space \mathcal{Q} is

$$\bar{\mathbf{1}} \circ \bar{\mathbf{2}} = -\bar{\mathbf{2}} \circ \bar{\mathbf{1}} = -\bar{\mathbf{3}}, \quad \bar{\mathbf{2}} \circ \bar{\mathbf{3}} = -\bar{\mathbf{3}} \circ \bar{\mathbf{2}} = -\bar{\mathbf{1}}, \quad \bar{\mathbf{3}} \circ \bar{\mathbf{1}} = -\bar{\mathbf{1}} \circ \bar{\mathbf{3}} = -\bar{\mathbf{2}} \quad (56abc)$$

and also

$$\bar{\mathbf{1}} \circ \bar{\mathbf{1}} = \bar{\mathbf{2}} \circ \bar{\mathbf{2}} = \bar{\mathbf{3}} \circ \bar{\mathbf{3}} = \bar{\mathbf{3}} \circ \bar{\mathbf{2}} \circ \bar{\mathbf{1}} = -\bar{\mathbf{1}} \quad (57)$$

Instead of $\mathbf{j}\mathbf{j} = \mathbf{k}$, etc., for the traditional development of the quaternion we have instead for the more recent formulation $\bar{\mathbf{1}} \circ \bar{\mathbf{2}} = -\bar{\mathbf{3}}$, or equivalently, $\bar{\mathbf{2}} \circ \bar{\mathbf{1}} = \bar{\mathbf{3}}$. Not surprisingly, since $\bar{\mathbf{1}}$, $\bar{\mathbf{2}}$ and $\bar{\mathbf{3}}$, as unit quaternions, are quaternions of rotation, a change in the order of quaternion multiplication must entail a change in the order of multiplication for the “imaginaries.”

The quaternion representation of a 3×1 column vector is given by

$$\bar{\mathbf{v}} \equiv \begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix} \quad (58)$$

whence,

$$\bar{\eta} \circ \bar{\mathbf{v}} \circ \bar{\eta}^* = \begin{bmatrix} \mathbf{R}(\bar{\eta}) \mathbf{v} \\ 0 \end{bmatrix} \quad (59)$$

or equivalently,

$$\bar{\eta} \circ \overset{\mathcal{E}}{\bar{\mathbf{v}}} \circ \bar{\eta}^* = \begin{bmatrix} \overset{\mathcal{E}}{\mathbf{v}} \\ 0 \end{bmatrix} = \overset{\mathcal{E}}{\bar{\mathbf{v}}} \quad (60)$$

in analogy with equation (25) above. Note the difference in order between equations (60) and (39). Explicitly, the rotation matrix as a function of the quaternion is given by

$$\mathbf{R}(\bar{\eta}) = (\eta_4^2 - |\boldsymbol{\eta}|^2) \mathbf{I}_{3 \times 3} + 2 \boldsymbol{\eta} \boldsymbol{\eta}^T + 2 \eta_4 [\boldsymbol{\eta} \times] \quad (61)$$

Generally, the computational burden for the transformation of column vectors is less if $\mathbf{R}(\bar{\eta})$ is first calculated using equation (61) and applied to the column vector than if equation (60) is used. The temporal evolution of the quaternion is described by equations (1) and (2).

The more recent formulation of quaternions avoids the pitfalls signalled by Lord Kelvin in an opening quote and those in the discussion of Tait and Cayley [29, 30], which dealt in part with the nature of the Hamilton imaginaries as coordinate axes.

Discussion

For the study of spacecraft attitude estimation and control and for spacecraft mission support, it is clear that the more recent formulation of the quaternion is the more apt. Writing a quaternion as a 4×1 column vector is not new and can be found in reference [21], which adheres to the traditional approach to quaternions. What the formulation of quaternions in reference [2] did was to make the multiplication rule the starting point of the formulation.

The biggest difference of the recent approach to quaternions in spacecraft attitude work and the traditional approach is not the use of 4×1 column vectors but the elimination of the Hamilton imaginaries altogether, and the modification of the quaternion multiplication rule so that quaternion multiplication and the corresponding matrix multiplication are performed in the same order. As we

have seen, this meant abandoning the traditional order for the multiplication of the Hamilton imaginaries, with $\mathbf{i}\mathbf{j} = \mathbf{k}$ being replaced by $\bar{\mathbf{e}}_2 \circ \bar{\mathbf{e}}_1 = \bar{\mathbf{e}}_3$.

The appearance of the quaternion basis vectors $\bar{\mathbf{e}}_1$, $\bar{\mathbf{e}}_2$, $\bar{\mathbf{e}}_3$ and $\bar{\mathbf{e}}_4$ are reminiscent of the autorepresentation of a basis in physical three-space. Thus, we might regard them as the autorepresentation of a “physical” quaternion basis $\bar{\mathbf{e}}_k$, $k = 1, 2, 3, 4$, in some kind of “physical” quaternion space, and these might bear a connection to the attitude dyadic [2] similar to that of the quaternion column-vector to the attitude matrix. One can indeed construct such a “physical” space of quaternions, but they serve no practical value.

It might seem from the perspective of this article that the traditional formulation of the quaternion is a rather ramshackle affair with scalars and vectors participating in some sort of clumsy and unholy marriage. That opinion would be very unfair. While the formulation of reference [2] is certainly better attuned to engineering and the needs of automatic computation, the place and purpose of quaternions in pure mathematics is rather different. The traditional formulation of quaternions is very much alive in books on Algebra as the first example of a skew field. The concern of pure mathematics is not in representing physical reality efficiently but in exploring mathematical structures. The deep concern of the present article with the representation of physical processes and the duality of physical vectors and column-vectors might seem to be a needless distraction in pure mathematics. The recent formulation for the quaternion seeks only to make the quaternion a more pedestrian object, a worthy goal, but not the only possible goal. As engineers, our interest is in “impure” mathematics, contaminated by the needs of practical application.

The need to abandon the traditional formulation of quaternions is revealed, as we have seen, in the traditional formulation itself. The Hamiltonian imaginaries cannot be interpreted as physical vectors but, at best, only as column vectors. Thus, they cannot label coordinate axes. That internal inconsistency, however, would not be apparent in a formulation that does not distinguish between physical vectors and column vector, but without that distinction, attitude estimation would be confused.

Interest in the quaternion has been renewed during the past two decades under the rubric of Geometric Algebra, in which quaternions and Grassman’s algebra play a fundamental rôle in the description of physical processes. These ideas have gained only a small foothold in Physics and Mathematics. The great champion of Geometric Algebra is David Hestenes (b. 1933), who is also the author of the most extensive work on that subject [31].

Some of the information in this article and the opening quotations have been gleaned from the Wikipedia articles on “Arthur Cayley,” “William Rowan Hamilton,” and “Quaternions” (<http://en.wikipedia.org>).

Acknowledgment

The author is grateful to Yang Cheng and F. Landis Markley for helpful comments.

References

- [1] ALTMANN, S. L., *Rotations, Quaternions, and Double Groups*, Oxford, Oxford University Press, 1986.
- [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] KUIPERS, J. B. *Quaternions and Rotation Sequences*, Princeton, Princeton University Press, 1999.
- [4] EULER, L., “Problema Algebraicum ob Affectiones Prorsus Singulares Memorabile,” *Novi Commentari Academiae Scientiarum Imperialis Petropolitanae*, Vol. 15, 1770, pp. 75–106.
- [5] EULER, L., “Nova Methodus Motum Corporum Rigidorum Determinandi,” *Novi Commentari Academiae Scientiarum Imperialis Petropolitanae*, Vol. 20, 1775, pp. 208–238, also *Leonhardi Euleri Opera Omnia, Series Secunda, Opera Mechanica Et Astronomica*, Basel, 1968, Vol. 9, pp. 99–125.
- [6] RODRIGUES, O., “Des lois géométriques qui régissent les déplacements d’un système solide dans l’espace, et de la variation des coordonnées provenant de ses déplacements considérés indépendamment des causes qui peuvent les produire,” *Journal des Mathématiques Pures et Appliquées*, Vol. 5, 1840, pp. 380–440.
- [7] GIBBS, J. W., *Vector Analysis*, E. B. Wilson, ed., New York, Scribners, 1901.
- [8] HAMILTON, W. R., “On Quaternions; or a new system of imaginaries in algebra,” *Philosophical Magazine*, Vol. 25, 1844, pp. 489–495.
- [9] SHUSTER, M. D. “Attitude Analysis in Flatland: the Plane Truth,” *The Journal of the Astronautical Sciences*, Vol. 52, Nos. 1 and 2, January/June 2004, pp. 195–209.
- [10] CAYLEY, A., “On the Motion of Rotation of a Solid Body,” *Cambridge Mathematics Journal*, Vol. 3, 1843, pp. 224–232.
- [11] CAYLEY, A., “Recherches sur les Matrices dont les termes sont des fonctions linéaires d’une seule Indeterminée,” *Journal für die reine und angewandte Mathematik*, Berlin, Crelle, 1855, pp. 313–317.
- [12] CAYLEY, A., “Memoire on the Theory of Matrices,” *Philosophical Transactions* 1, Vol. CXLVIII, 1858, pp. 17–37.
- [13] CAYLEY, A., “On certain results relating to quaternions,” *Philosophical Magazine*, Vol. 26, 1845, pp. 141–145.
- [14] HAMILTON, W. R., *Elements of Quaternions*, London, Longmans, Green & Co., 1866.
- [15] GIBBS, J. W., *Scientific Papers*, Vol. II, New York, Dover, 1961.
- [16] HEAVISIDE, O., *Electromagnetic Theory*, London, 1893 (Vol. 1), 1899 (Vol. 2), 1912 (Vol. 3).
- [17] CROWE, M. J., *A History of Vector Analysis*, New York, Dover, 1967.
- [18] VAN DER WAERDEN, B. L., *A History of Algebra*, New York and Berlin, Springer Scientific + Business Media, 1985.
- [19] WERTZ, J. R. (ed.) *Spacecraft Attitude Determination and Control*, New York and Berlin, Springer Science + Business Media, 1978.

- [20] SHUSTER, M. D. “The Youngest Quadrant,” in “In my estimation,” *The Journal of the Astronautical Sciences*, Vol. 54, Nos. 3–4, July–December 2006, pp. 273–297.
- [21] FALLON, L. S. III “Quaternions,” Appendix D in J. R. Wertz (ed.), *op. cit.*, pp. 758–59.
- [22] MARKLEY, F. L. “Parameterization of the attitude,” Section 12.1 in J. R. Wertz (ed.), *op. cit.*, pp. 410–420.
- [23] WIE, B. *Space Vehicle Dynamics and Control*, Reston, Virginia, American Institute of Aeronautics and Astronautics, 2008.
- [24] SCHAUB H. and JUNKINS, J. L. *Analytical Mechanics of Space Systems*, Reston, Virginia, American Institute of Aeronautics and Astronautics, 2003.
- [25] CRASSIDIS, J. L. and JUNKINS J. L. *Optimal Estimation of Dynamic Systems*, New York, Chapman and Hall/CRC, 2004.
- [26] SHUSTER, M. D. and DELLINGER, W. Chapter 5, “Spacecraft Attitude Determination and Control,” in PISCANE, V. L. (ed.), *Fundamentals of Space Systems*, Oxford, Oxford University Press, 2005, pp. 236–325.
- [27] HUGHES, P. C. *Spacecraft Attitude Dynamics*, New York, John Wiley & Sons, 1986; reprint: Mineola, New York, Dover Publications, 2004.
- [28] HAMILTON, W. R., *Lectures on Quaternions*, Dublin, Hodges & Smith, 1853.
- [29] CAYLEY, A. “Coordinates versus Quaternions,” *Proceedings of the Royal Society of Edinburgh*, vol. 20, pp. 271–275, 1893–1894 session (published 1895).
- [30] TAIT, P. G. “On the Intrinsic Nature of the Quaternion,” *Proceedings of the Royal Society of Edinburgh*, vol. 20, pp. 276–284, 1893–1894 session (published 1895).
- [31] HESTENES, D. *New Foundations for Classical Mechanics*, second edition, New York and Berlin, Springer Scientific + Business Media, 1999.