NOTES ON THE MECHANICS OF RIGID BODY ROTATIONS, QUATERNIONS AND SOME ASSOCIATED MATHEMATICS

Donald J. Dichmann

March 19, 1999

1 Introduction

This is an informal set of notes on the mechanics of rigid bodies. The primary focus is on the representation of the orientation of a rigid body by Euler parameters. Such a representation also pertains to the director theory of rod mechanics. I also summarize some closely related mathematics, particularly the algebra and geometry of the quaternions.

2 Representations of Rotations

A parametrization of the directors $\{\mathbf{d}_k\}$ is equivalent to selecting a representation for the group of orthogonal transformations $SO(3)$. In fact, we will always regard an element $\mathbf{A} \in SO(3)$ as representing the components of the directors $\{\mathbf{d}_k\}$ with respect to the fixed frame $\{\mathbf{e}_k\}$, so that in matrix form $\mathbf{A} = (A_{ij})$, where $A_{ij} \equiv \mathbf{d}_i \cdot \mathbf{e}_j$. That is, the vector $\mathbf{d}_k$ forms the $k-$th row of the matrix $\mathbf{A}$.

In mechanics there are several common choices for such a representation. (Cf. e.g. [5].) Euler angles are often used because this representation uses the the minimum number of parameters, three. There are two significant drawbacks in using Euler angles. First, due to the topology of $SO(3)$ a set of three parameters cannot provide a global representation. In the case of Euler angles there is a direction in space at which the representation becomes singular. This singularity makes the study of dynamical systems a delicate matter, since the particular set of Euler angles must then be selected to avoid the singularity, or different charts must be used as the dynamics evolve. A second drawback in the use of Euler angles is the appearance of trigonometric functions, which can be computationally expensive.

In these notes we instead focus upon Euler parameters, which provide a four-dimensional, global, two-to-one parametrization of $SO(3)$. In addition, only polynomial and rational functions arise in this representation. The difficulty with the Euler parameter representation is that the squares of the parameters must sum to one, which imposes a constraint upon the dynamics. An elementary discussion of Euler parameters can be found in [5]. Further discussions of Euler parameters as they apply to the motion of rigid spacecraft can also be found in [10] and [25]. The kinematics of Euler parameters are summarized in section 2.1. For completeness, we also discuss a representation by Euler angles in section 2.2.

2.1 Euler Parameters

A set of Euler parameters is a quadruple of real numbers $\mathbf{q} = (q_1, q_2, q_3, q_4)^T$ that satisfies the identity

$$\mathbf{q} \cdot \mathbf{q} = {q_1}^2 + {q_2}^2 + {q_3}^2 + q_4^2 = 1. \tag{2.1.1}$$

The set of all quadruples satisfying (2.1.1) forms the unit sphere $\mathbf{S}^3$ in $\mathbf{R}^4$. A set of Euler parameters $\mathbf{q}$ is often called a unit quaternion, in connection with the algebra of quaternions [21]. In fact one reason that Euler parameters are so useful for the representation of $SO(3)$ is because the composition of two rotations may be expressed an elegant manner which is closely related to the multiplication of quaternions. Euler parameters are also closely associated with the Cayley-Klein parameters used to represent $SU(2)$ [14]. The relationship with Cayley-Klein parameters is discussed later in these notes.

To obtain a representation of the directors in terms of Euler parameters, first recall that by Euler's Theorem each element of $\mathbf{A} \in SO(3)$ is equivalent to a rotation through an angle Φ about an axis determined by a unit vector $\mathbf{k}$. Specifically, $\mathbf{A}$ is given in terms of $\mathbf{k} = k_j\mathbf{e}_j$ and Φ by

$$
\begin{aligned}
\mathbf{A} &= (\cos\Phi)\mathbf{1} + (1-\cos\Phi)\mathbf{k}\otimes\mathbf{k} - \sin\Phi[\mathbf{k}\times] \\
&= \begin{pmatrix} c\Phi + k_1^2(1-c\Phi) & k_1k_2(1-c\Phi)+k_3 s\Phi & k_1k_3(1-c\Phi)-k_2 s\Phi \\ k_1k_2(1-c\Phi)-k_3 s\Phi & c\Phi + k_2^2(1-c\Phi) & k_2k_3(1-c\Phi)+k_1 s\Phi \\ k_1k_3(1-c\Phi)+k_2 s\Phi & k_2k_3(1-c\Phi)-k_1 s\Phi & c\Phi + k_3^2(1-c\Phi) \end{pmatrix},
\end{aligned}
\tag{2.1.2}
$$

where $c\Phi$ is $\cos\Phi$, $s\Phi$ is $\sin\Phi$, $\mathbf{1}$ is the identity matrix, $\otimes$ is the outer product[1], and $[\mathbf{k}\times]$ is the skew matrix

$$
[\mathbf{k}\times] \equiv \begin{pmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{pmatrix}
\tag{2.1.3}
$$

corresponding to the cross product with the vector $\mathbf{k}$.

The Euler parameters (q_1, q_2, q_3, q_4) are defined in terms of $\mathbf{k}$ and Φ by

$$
q_j = k_j \sin(\Phi/2), \qquad \text{for } j = 1, 2, 3
\tag{2.1.4}
$$

and

$$
q_4 = \cos(\Phi/2).
\tag{2.1.5}
$$

Euler parameters have the unusual feature that if $\mathbf{q}$ corresponds to the pair $(\mathbf{k}, \Phi)$, then $-\mathbf{q}$ corresponds to the pair $(\mathbf{k}, \Phi + 2\pi)$ which represents a coterminal rotation. Thus there is a two-to-one correspondence between Euler parameters and rotations. Exploiting the identities $\cos\Phi = 2\cos^2(\Phi/2) - 1$ and $\sin\Phi = 2\sin(\Phi/2)\cos(\Phi/2)$ together with equations (2.1.1), (2.1.5), (2.1.4) and (2.1.2) one can represent the rotation in terms of $\mathbf{q}$ as

$$
\mathbf{A}(\mathbf{q};\beta) = |\mathbf{q}|^{\beta-2} \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_1q_4 + q_2q_3) \\ 2(q_1q_3 + q_2q_4) & 2(-q_1q_4 + q_2q_3) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix},
\tag{2.1.6}
$$

where $|\mathbf{q}| \equiv \sqrt{\mathbf{q}\cdot\mathbf{q}}$.

The factor of $|\mathbf{q}|^{\beta-2}$ we have introduced in equation (2.1.6) is not generally used in this representation of the rotation matrix, but indicates that there is actually a family of representations of a rotation in terms of a quaternion $\mathbf{q}$, all of which are equal if $|\mathbf{q}| = 1$. The parameter β is called the *degree of homogeneity* of the rotation matrix, since the multiplication of the quaternion $\mathbf{q}$ by a scalar ϵ causes the matrix to scale as $\mathbf{A}(\epsilon\mathbf{q};\beta) = \epsilon^\beta\mathbf{A}(\mathbf{q};\beta)$. Most commonly, one sets $\beta = 2$ so that the factor $1/|\mathbf{q}|^{2-\beta}$ is identically one. Alternatively, if we choose $\beta = 0$, the matrix $\mathbf{A}(\mathbf{q};0)$ is *explicitly* orthogonal for all $\mathbf{q} \neq 0$, from which follows the identity

$$
\mathbf{A}(\mathbf{q};\beta) = |\mathbf{q}|^\beta \mathbf{A}(\mathbf{q};0)\ .
\tag{2.1.7}
$$

As we have noted, the choice of β is of no consequence if the normality condition $|\mathbf{q}| = 1$ is satisfied. However, during computer simulations of rotational dynamics the normality

[1]If vectors $\mathbf{a}$ and $\mathbf{b}$ are thought of as column vectors, possibly of different dimensions, then their outer product $\mathbf{a}\otimes\mathbf{b}$ is the given by the matrix $\mathbf{a}\,\mathbf{b}^T$. Put differently, if vector $\mathbf{c}$ is of the same dimension as $\mathbf{b}$, then $(\mathbf{a}\otimes\mathbf{b})\,\mathbf{c} = \mathbf{a}\,(\mathbf{b}\cdot\mathbf{c})$.

condition is often violated due to numerical errors (although columns and rows of the matrix $\mathbf{A}(\mathbf{q};\beta)$ are always orthogonal). We therefore seek to determine the effects when $|\mathbf{q}| \neq 1$ by considering various choices of β. We will however assume that $\mathbf{q} \neq 0$, so that the matrix $\mathbf{A}(\mathbf{q};\beta)$ remains well-defined.

We use the matrix $\mathbf{A}(\mathbf{q};\beta)$ to denote the transformation from the reference frame $\mathcal{O} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to the director frame $\mathcal{D} = \{\mathbf{d}_1(\mathbf{q};\beta), \mathbf{d}_2(\mathbf{q};\beta), \mathbf{d}_3(\mathbf{q};\beta)\}$. Thus

$$\mathbf{d}_1(\mathbf{q};\beta) = |\mathbf{q}|^{\beta-2} \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 \\ 2(q_1q_2 + q_3q_4) \\ 2(q_1q_3 - q_2q_4) \end{pmatrix} , \tag{2.1.8}$$

$$\mathbf{d}_2(\mathbf{q};\beta) = |\mathbf{q}|^{\beta-2} \begin{pmatrix} 2(q_1q_2 - q_3q_4) \\ -q_1^2 + q_2^2 - q_3^2 + q_4^2 \\ 2(q_2q_3 + q_1q_4) \end{pmatrix} , \tag{2.1.9}$$

$$\mathbf{d}_3(\mathbf{q};\beta) = |\mathbf{q}|^{\beta-2} \begin{pmatrix} 2(q_1q_3 + q_2q_4) \\ 2(q_2q_3 - q_1q_4) \\ -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix} . \tag{2.1.10}$$

For any β and for any $\mathbf{q} \neq 1$, the set of directors $\{\mathbf{d}_j(\mathbf{q};\beta)\}$ is orthogonal, but the directors are orthonormal only if $|\mathbf{q}| = 1$ or $\beta = 0$ (or both). Instead we have the identities

$$\mathbf{d}_j(\mathbf{q};\beta) = |\mathbf{q}|^{\beta}\mathbf{d}_j(\mathbf{q};0) \tag{2.1.11}$$

analogous to (2.1.7), and

$$|\mathbf{d}_j(\mathbf{q};\beta)| = |\mathbf{q}|^{\beta} . \tag{2.1.12}$$

We now consider the problem of expressing the angular velocities in the body frame $\{\omega_j\}$ as a function of the Euler parameters. First, because the directors $\left\{\mathbf{d}_j^{\beta}\right\}$ are not normalized for $\beta \neq 0$, the kinematic relation

$$\dot{\mathbf{d}}_k = \boldsymbol{\omega} \times \mathbf{d}_k , \tag{2.1.13}$$

does not hold in general, and so the definition of angular velocity must be generalized slightly.

We believe that it is must reasonable to represent a vector in the body frame with respect to the explicitly orthonormal basis $\{\mathbf{d}_j(\mathbf{q};0)\}$, so that the angular velocity is written as

$$\boldsymbol{\omega} = \omega_1\,\mathbf{d}_1(\mathbf{q};0) + \omega_2\,\mathbf{d}_2(\mathbf{q};0) + \omega_3\,\mathbf{d}_3(\mathbf{q};0) \tag{2.1.14}$$

We can use the angular velocity (2.1.14) to write the time derivative of $\mathbf{d}_3(\mathbf{q};\beta)$ as

$$\begin{aligned} \dot{\mathbf{d}}_3(\mathbf{q};\beta) &= [\,\omega_1\,\mathbf{d}_1(\mathbf{q};0) + \omega_2\,\mathbf{d}_2(\mathbf{q};0) + \omega_3\,\mathbf{d}_3(\mathbf{q};0)\,] \times \mathbf{d}_3(\mathbf{q};\beta) \\ &= -\omega_1\mathbf{d}_2(\mathbf{q};\beta) + \omega_2\mathbf{d}_1(\mathbf{q};\beta) , \end{aligned} \tag{2.1.15}$$

where we have used the identity

$$\mathbf{d}_i(\mathbf{q};0) \times \mathbf{d}_j(\mathbf{q};\beta) = \sum_{k=1}^{3} \epsilon_{ijk}\mathbf{d}_k(\mathbf{q};\beta) . \tag{2.1.16}$$

Taking the dot product of $\mathbf{d}_2(\mathbf{q};\beta)$ with equation (2.1.16), we can express express ω_1 as

$$\omega_1 = -\frac{\mathbf{d}_2(\mathbf{q};\beta)\cdot\dot{\mathbf{d}}_3(\mathbf{q};\beta)}{|\mathbf{q}|^{2\beta}} = -\frac{\mathbf{d}_2(\mathbf{q};\beta)}{|\mathbf{q}|^{2\beta}} \cdot \mathbf{D}_3(\mathbf{q};\beta)^T\,\dot{\mathbf{q}} \tag{2.1.17}$$

where the 4-by-3 matrix $\mathbf{D}_3(\mathbf{q};\beta)$ is defined by

$$\mathbf{D}_k(\mathbf{q};\beta) \equiv \left(\frac{\partial \mathbf{d}_k(\mathbf{q};\beta)}{\partial \mathbf{q}}\right)^T \quad . \tag{2.1.18}$$

Equation (2.1.17) can be rewritten as

$$\omega_1 = \frac{2}{|\mathbf{q}|^2}\mathbf{B}_1\mathbf{q}\cdot\dot{\mathbf{q}} \ , \tag{2.1.19}$$

where the 4-vector $\mathbf{B}_1\mathbf{q}$ is is defined to be

$$\mathbf{B}_1\mathbf{q} \equiv -\frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_3(\mathbf{q};\beta)\,\mathbf{d}_2(\mathbf{q};\beta) \ . \tag{2.1.20}$$

Using equation (2.1.10), one can write $\mathbf{D}_3(\mathbf{q};\beta)$ explicitly as

$$\mathbf{D}_3(\mathbf{q};\beta) = 2|\mathbf{q}|^{\beta-2}\begin{pmatrix} q_3 & -q_4 & -q_1 \\ q_4 & q_3 & -q_2 \\ q_1 & q_2 & q_3 \\ q_2 & -q_1 & q_4 \end{pmatrix} + (\beta-2)\frac{\mathbf{q}}{|\mathbf{q}|^2}\otimes\mathbf{d}_3(\mathbf{q};\beta) \ . \tag{2.1.21}$$

It follows from equation (2.1.21) that the vector $\mathbf{B}_1\mathbf{q}$ has an extremely simple form, namely

$$\mathbf{B}_1\mathbf{q} = (\quad q_4, \quad q_3, -q_2, -q_1)^T \quad . \tag{2.1.22}$$

We remark that time differentiation of the identity

$$\mathbf{d}_i(\mathbf{q};\beta)\cdot\mathbf{d}_j(\mathbf{q};\beta) = |\mathbf{q}|^{2\beta}\delta_{ij} \tag{2.1.23}$$

with $i=2$ and $j=3$ implies

$$\mathbf{d}_3(\mathbf{q};\beta)\cdot\dot{\mathbf{d}}_2(\mathbf{q};\beta) = -\mathbf{d}_2(\mathbf{q};\beta)\cdot\dot{\mathbf{d}}_3(\mathbf{q};\beta) \ , \tag{2.1.24}$$

which in turn implies that the vector $\mathbf{B}_1\mathbf{q}$ can be expressed alternatively as

$$\mathbf{B}_1\mathbf{q} \equiv \frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_2(\mathbf{q};\beta)\,\mathbf{d}_3(\mathbf{q};\beta) \ . \tag{2.1.25}$$

Also, because the components of the vector $\mathbf{B}_1\mathbf{q}$ depend linearly upon $\mathbf{q}$, we can interpret $\mathbf{B}_1\mathbf{q}$ as the product of a 4-by-4 matrix $\mathbf{B}_1$ with $\mathbf{q}$. More will be said about the matrix $\mathbf{B}_1$ below.

Using computations analogous to those above, one can express the other components of angular velocity as

$$\omega_j = \frac{2}{|\mathbf{q}|^2}\dot{\mathbf{q}}\cdot\mathbf{B}_j\mathbf{q} \tag{2.1.26}$$

for $j=2$ and 3, where

$$\begin{aligned} \mathbf{B}_2\mathbf{q} &= \frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_3(\mathbf{q};\beta)\,\mathbf{d}_1(\mathbf{q};\beta) \\ &= -\frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_1(\mathbf{q};\beta)\,\mathbf{d}_3(\mathbf{q};\beta) \\ &= (-q_3, \quad q_4, \quad q_1, -q_2)^T \ , \end{aligned} \tag{2.1.27}$$

and

$$\begin{aligned} \mathbf{B}_3\mathbf{q} &= \frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_2(\mathbf{q};\beta)\,\mathbf{d}_3(\mathbf{q};\beta) \\ &= -\frac{1}{2}\,|\mathbf{q}|^{2-2\beta}\,\mathbf{D}_3(\mathbf{q};\beta)\,\mathbf{d}_2(\mathbf{q};\beta) \\ &= (\quad q_2, -q_1, \quad q_4, -q_3)^T \ . \end{aligned} \tag{2.1.28}$$

The vectors $\mathbf{B}_j\mathbf{q}$ are all linear in $\mathbf{q}$. Moreover, it is readily shown that the set of vectors $\{\mathbf{q}, \mathbf{B}_1\mathbf{q}, \mathbf{B}_2\mathbf{q}, \mathbf{B}_3\mathbf{q}\}$ is mutually orthogonal, and forms an orthonormal basis for $\mathbf{R}^4$ if $|\mathbf{q}| = 1$.

One can use the basis of vectors $\{\mathbf{q}, \mathbf{B}_1\mathbf{q}, \mathbf{B}_2\mathbf{q}, \mathbf{B}_3\mathbf{q}\}$ in $\mathbf{R}^4$ together with the basis $\{\mathbf{d}_j\}$ in $\mathbf{R}^3$ to rewrite the matrices $\mathbf{D}_k$ in a geometrically more meaningful way. We first note that because $|\mathbf{d}_k(\mathbf{q};\beta)| = |\mathbf{q}|^\beta$ implies

$$\frac{1}{2}\frac{\partial}{\partial \mathbf{q}}|\mathbf{d}_k(\mathbf{q};\beta)|^2 = \mathbf{D}_k(\mathbf{q};\beta)\mathbf{d}_k(\mathbf{q};\beta) = \beta|\mathbf{q}|^{2\beta-2}\mathbf{q} \ . \tag{2.1.29}$$

Then, because $\{|\mathbf{q}|^{-\beta}\mathbf{d}_k(\mathbf{q};\beta)\}$ forms an orthonormal basis for $\mathbf{R}^3$, we can use equations (2.1.20), (2.1.28) and (2.1.29) to rewrite $\mathbf{D}_3(\mathbf{q};\beta)$ as

$$\mathbf{D}_3(\mathbf{q};\beta) = \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_2\mathbf{q}\otimes\mathbf{d}_1(\mathbf{q};\beta) - \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_1\mathbf{q}\otimes\mathbf{d}_2(\mathbf{q};\beta) + \frac{\beta}{|\mathbf{q}|^2}\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) \ . \tag{2.1.30}$$

Similarly, one can show that

$$\mathbf{D}_1(\mathbf{q};\beta) = \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_3\mathbf{q}\otimes\mathbf{d}_2(\mathbf{q};\beta) - \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_2\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) + \frac{\beta}{|\mathbf{q}|^2}\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) \ , \tag{2.1.31}$$

$$\mathbf{D}_2(\mathbf{q};\beta) = \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_1\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) - \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}_3\mathbf{q}\otimes\mathbf{d}_1(\mathbf{q};\beta) + \frac{\beta}{|\mathbf{q}|^2}\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) \ . \tag{2.1.32}$$

which can be summarized as

$$\mathbf{D}_k(\mathbf{q}) = \sum_{i,j=1}^{3}\frac{2}{|\mathbf{q}|^2}\,\epsilon_{ijk}\mathbf{B}_j\mathbf{q}\otimes\mathbf{d}_i + \frac{\beta}{|\mathbf{q}|^2}\mathbf{q}\otimes\mathbf{d}_3(\mathbf{q};\beta) \ . \tag{2.1.33}$$

Equation (2.1.33) yields a geometric interpretation of the vector $\mathbf{B}_j\mathbf{q}$: for small ϵ, a change in $\mathbf{q}$ by $\epsilon\mathbf{B}_k\mathbf{q}$ produces a rotation about the $\mathbf{d}_k$ axis through an angle 2ϵ. In particular we have the identity

$$\mathbf{D}_k(\mathbf{q};\beta)^T\mathbf{B}_k\mathbf{q} = \mathbf{0} \ , \tag{2.1.34}$$

which reflects the fact that $\mathbf{d}_k$ itself is unchanged by a rotation about $\mathbf{d}_k$. We also observe from equation (2.1.33) that

$$\mathbf{D}_k(\mathbf{q};\beta)^T\mathbf{q} = \beta\mathbf{d}_k(\mathbf{q};\beta) \ , \tag{2.1.35}$$

a consequence of Euler's theorem and the fact that $\mathbf{d}_k$ is homogeneous with degree β with respect to $\mathbf{q}$.

Equations (2.1.26) and (2.1.31) –(2.1.30) seem to indicate that the set of vectors $\{2\mathbf{q}/|\mathbf{q}|^2, 2\mathbf{B}_j\mathbf{q}/|\mathbf{q}|^2\}$, rather than $\{\mathbf{q}, \mathbf{B}_j\mathbf{q}\}$, is the more natural basis for the space tangent to $\mathbf{S}^3$. In this connection, we note that the dual basis to $\{2\mathbf{q}/|\mathbf{q}|^2, 2\mathbf{B}_j\mathbf{q}/|\mathbf{q}|^2\}$ is $\{\frac{1}{2}\mathbf{q}, \frac{1}{2}\mathbf{B}_j\mathbf{q}\}$, which arises in a Hamiltonian formulation of rigid body dynamics using Euler parameters.

Using equation (2.1.26), we can express components of the angular velocity in the rotating frame collectively as

$$\underline{\boldsymbol{\omega}} = \frac{2}{|\mathbf{q}|^2}\,\mathbf{B}(\mathbf{q})\,\dot{\mathbf{q}}, \tag{2.1.36}$$

where

$$\mathbf{B}(\mathbf{q}) = \sum_{j=1}^{3} \mathbf{e}_j \otimes \mathbf{B}_j\mathbf{q} = \begin{pmatrix} q_4 & q_3 & -q_2 & -q_1 \\ -q_3 & q_4 & q_1 & -q_2 \\ q_2 & -q_1 & q_4 & -q_3 \end{pmatrix} . \tag{2.1.37}$$

This is the relationship between angular velocity and Euler parameters that one commonly sees in the literature. I am not aware of any prior work in the kinematics of rods or of rigid bodies where the vectors $\mathbf{B}_j\mathbf{q}$ have been utilized explicitly. However, I believe the basis vectors $\mathbf{B}_j\mathbf{q}$ are in general more useful for a geometric comprehension and for detailed calculations.

Using calculation analogous to those as above, we can express the components of angular velocity with respect the *fixed* frame as

$$\omega \cdot \mathbf{e}_j = \frac{2}{|\mathbf{q}|^2}\dot{\mathbf{q}} \cdot \mathbf{F}_j\mathbf{q} \ , \tag{2.1.38}$$

where

$$\mathbf{F}_1\mathbf{q} = (\ q_4, -q_3, \ q_2, -q_1)^T \ , \tag{2.1.39}$$
$$\mathbf{F}_2\mathbf{q} = (\ q_3, \ q_4, -q_1, -q_2)^T \ , \tag{2.1.40}$$
$$\mathbf{F}_3\mathbf{q} = (-q_2, \ q_1, \ q_4, -q_3)^T \ . \tag{2.1.41}$$

This equation are often written collectively as

$$\boldsymbol{\omega} = \frac{2}{|\mathbf{q}|^2}\,\mathbf{F}(\mathbf{q})\,\dot{\mathbf{q}} \ , \tag{2.1.42}$$

where

$$\mathbf{F}(\mathbf{q}) = \sum_{j=1}^{3} \mathbf{e}_j \otimes \mathbf{B}_j\mathbf{q} = \begin{pmatrix} q_4 & -q_3 & q_2 & -q_1 \\ q_3 & q_4 & -q_1 & -q_2 \\ -q_2 & q_1 & q_4 & -q_3 \end{pmatrix} . \tag{2.1.43}$$

By the definition of the transformation matrix $\mathbf{A}$, the components of angular momentum with respect the fixed frame and the rotating frame are given by

$$\boldsymbol{\omega} \cdot \mathbf{e}_j = \sum_{i=1}^{3}(\boldsymbol{\omega} \cdot \mathbf{d}_i)(\mathbf{d}_i \cdot \mathbf{e}_j) = \sum_{i=1}^{3} A_{ij}(\boldsymbol{\omega} \cdot \mathbf{d}_i) \ , \tag{2.1.44}$$

or

$$\boldsymbol{\omega} = \mathbf{A}^T(\mathbf{q})\,\underline{\boldsymbol{\omega}} \ . \tag{2.1.45}$$

It therefore follows that

$$\mathbf{F}(\mathbf{q}) = \mathbf{A}(\mathbf{q})^T\mathbf{B}(\mathbf{q}) \tag{2.1.46}$$

It is a remarkable property of Euler parameters that the matrices $\mathbf{B}(\mathbf{q})$ and $\mathbf{F}(\mathbf{q})$ both take such a simple form. We now note a few properties of the matrices $\mathbf{B}$ and $\mathbf{F}$ which will be used later. Observe first that the 4-by-4 matrix given by

$$\mathbf{R}(\mathbf{q}) \equiv \frac{1}{|\mathbf{q}|}\begin{pmatrix} \mathbf{B}(\mathbf{q}) \\ \mathbf{q}^T \end{pmatrix} \tag{2.1.47}$$

is orthogonal. It then follows from the identity $\mathbf{R}(\mathbf{q})\mathbf{R}(\mathbf{q})^T = \mathbf{1}$ that

$$\mathbf{B}(\mathbf{q})\mathbf{B}^T(\mathbf{q}) = |\mathbf{q}|^2\mathbf{1}, \tag{2.1.48}$$

and that the 3-by-4 matrix $\mathbf{B}(\mathbf{q})$ has rank 3, with its null space spanned by $\mathbf{q}$:

$$\mathbf{B}(\mathbf{q})\mathbf{q} = \mathbf{0}. \tag{2.1.49}$$

Similarly the identity $\mathbf{R}(\mathbf{q})^T\mathbf{R}(\mathbf{q}) = \mathbf{1}$ implies

$$\mathbf{B}^T(\mathbf{q})\mathbf{B}(\mathbf{q}) = |\mathbf{q}|^2\mathbf{\Pi}(\mathbf{q}), \tag{2.1.50}$$

where

$$\mathbf{\Pi}(\mathbf{q}) \equiv \mathbf{1} - \frac{\mathbf{q}\otimes\mathbf{q}}{|\mathbf{q}|^2} \tag{2.1.51}$$

is the projection onto the subspace orthogonal to $\mathbf{q}$. Notice that equations (2.1.37), (2.1.43) and (2.1.48) imply

$$\frac{1}{|\mathbf{q}|^2}\mathbf{B}(\mathbf{q})\,\mathbf{F}(\mathbf{q})^T = \mathbf{A}(\mathbf{q}). \tag{2.1.52}$$

Equation (2.1.52) shows that the matrices $\mathbf{B}(\mathbf{q})$ and $\mathbf{F}(\mathbf{q})$ provide a factorization of the rotation matrix $\mathbf{A}(\mathbf{q})$, analogous to the factorization of a rotation matrix in terms of three elementary rotations using Euler angles. Apparently Euler first introduced Euler parameters precisely for the purpose of producing such a factorization. (Cf. [9], [21].) Refer to the end of section 4 for further remarks on this subject.

The matrices which determine the vectors $\mathbf{B}_j\mathbf{q}$ are

$$\mathbf{B}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{2.1.53}$$

$$\mathbf{B}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \tag{2.1.54}$$

and

$$\mathbf{B}_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{2.1.55}$$

The skew matrices $\mathbf{B}_k$ satisfy the commutation relations

$$(\mathbf{B}_k)^2 = -\mathbf{1} \qquad \text{for } k = 1, 2, 3, \tag{2.1.56}$$
$$\mathbf{B}_1\mathbf{B}_2 = -\mathbf{B}_3, \tag{2.1.57}$$
$$\mathbf{B}_2\mathbf{B}_3 = -\mathbf{B}_1, \tag{2.1.58}$$
$$\mathbf{B}_3\mathbf{B}_1 = -\mathbf{B}_2, \tag{2.1.59}$$

which can be summarized as

$$\mathbf{B}_j\mathbf{B}_k = -\delta_{jk}\mathbf{1} - \sum_{i=1}^{3} \epsilon_{ijk}\mathbf{B}_i, \tag{2.1.60}$$

where δ_{jk} is the Kronecker delta, and ϵ_{ijk} is the permutation tensor. Thus group of matrices $\{\pm\mathbf{1}, \pm\mathbf{B}_k\}$ is isomorphic to the quaternion group. (Cf. [11].) The skewness of the matrices $\mathbf{B}_j$ together with the property (2.1.56) imply that each exponential $\boldsymbol{\mathcal{B}}_j \equiv \exp(\frac{1}{2}\epsilon\mathbf{B}_j)$ is an orthogonal transformation in $\mathbf{R}^4$ given by

$$\boldsymbol{\mathcal{B}}_j = \exp(\tfrac{1}{2}\epsilon\mathbf{B}_j) = \cos(\tfrac{1}{2}\epsilon)\mathbf{1} + \sin(\tfrac{1}{2}\epsilon)\mathbf{B}_j. \tag{2.1.61}$$

In fact each transformation $\boldsymbol{\mathcal{B}}_j$ on the Euler parameters corresponds to a rotation $\tilde{\boldsymbol{\mathcal{B}}}_j$ in $\mathbf{R}^3$ through an angle ϵ about the axis $\mathbf{d}_j$. In particular,

$$\begin{aligned} \mathbf{d}_1(\boldsymbol{\mathcal{B}}_3\mathbf{q}) &= \cos(\epsilon)\,\mathbf{d}_1(\mathbf{q}) + \sin(\epsilon)\,\mathbf{d}_2(\mathbf{q}), \\ \mathbf{d}_2(\boldsymbol{\mathcal{B}}_3\mathbf{q}) &= -\sin(\epsilon)\,\mathbf{d}_1(\mathbf{q}) + \cos(\epsilon)\,\mathbf{d}_2(\mathbf{q}), \\ \mathbf{d}_3(\boldsymbol{\mathcal{B}}_3\mathbf{q}) &= \mathbf{d}_3(\mathbf{q}). \end{aligned} \tag{2.1.62}$$

The transformation $\boldsymbol{\mathcal{B}}_3$ describes the rotational symmetry about the $\mathbf{d}_3$ axis in the dynamics of an isotropic rod.

If $\mathbf{q}(t)$ is a unit quaternion for all time t, then the time differentiation of the norm condition (2.1.1) implies

$$\mathbf{q} \cdot \dot{\mathbf{q}} = 0 \ . \tag{2.1.63}$$

Equation (2.1.26) for the angular velocity can be inverted using the orthogonality condition (2.1.63) to obtain

$$\dot{\mathbf{q}} = \sum_{j=1}^{3} \tfrac{1}{2}\omega_j\mathbf{B}_j\mathbf{q} \tag{2.1.64}$$

For given $\{\omega_j(t)\}$, equation (2.1.64) is a first-order linear differential equation for $\mathbf{q}$ which is extremely convenient in spacecraft dynamics. We can use (2.1.64) to obtain a relationship between a given set of Euler angles describing a rotation and the equivalent set of Euler parameters. For example, an orientation given by the $3-2-3$ Euler angle sequence (ϕ, θ, ψ) can be reached through a sequence of three rotations

$$\begin{aligned} (\omega_1, \omega_2, \omega_3) &= (0, 0, \phi) \text{ for } t \in [0, 1), \\ (\omega_1, \omega_2, \omega_3) &= (0, \theta, 0) \text{ for } t \in [1, 2), \\ (\omega_1, \omega_2, \omega_3) &= (0, 0, \psi) \text{ for } t \in [2, 3). \end{aligned} \tag{2.1.65}$$

Because the differential equation (2.1.64) is linear and autonomous on each of these intervals, the solution at the completion of the rotation sequence is simply

$$\begin{aligned}\mathbf{q} &= \exp(\tfrac{1}{2}\psi\mathbf{B}_3)\,\exp(\tfrac{1}{2}\theta\mathbf{B}_2)\,\exp(\tfrac{1}{2}\phi\mathbf{B}_3)\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\\ &= \left[c(\tfrac{1}{2}\psi)\mathbf{1}+s(\tfrac{1}{2}\psi)\mathbf{B}_3\right]\left[c(\tfrac{1}{2}\theta)\mathbf{1}+s(\tfrac{1}{2}\theta)\mathbf{B}_2\right]\left[c(\tfrac{1}{2}\phi)\mathbf{1}+s(\tfrac{1}{2}\phi)\mathbf{B}_3\right]\begin{pmatrix}0\\0\\0\\1\end{pmatrix}.\end{aligned} \tag{2.1.66}$$

This calculation appears (somewhat mysteriously) in Junkins and Turner [12], who in turn credit Harold Morton. The results table of Euler parameters for each of the twelve Euler angle sequences also appears in [12].

Similarly the vectors $\mathbf{F}_j\mathbf{q}$ are given by the matrices

$$\mathbf{F}_1 = \begin{pmatrix}0&0&0&1\\0&0&-1&0\\0&1&0&0\\-1&0&0&0\end{pmatrix}, \tag{2.1.67}$$

$$\mathbf{F}_2 = \begin{pmatrix}0&0&1&0\\0&0&0&1\\-1&0&0&0\\0&-1&0&0\end{pmatrix}, \tag{2.1.68}$$

and

$$\mathbf{F}_3 = \begin{pmatrix}0&-1&0&0\\1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}. \tag{2.1.69}$$

The skew matrices $\mathbf{F}_k$ satisfy the commutation relations

$$(\mathbf{F}_k)^2 = -\mathbf{1} \qquad \text{for } k=1,2,3, \tag{2.1.70}$$

$$\mathbf{F}_1\mathbf{F}_2 = \mathbf{F}_3, \tag{2.1.71}$$

$$\mathbf{F}_2\mathbf{F}_3 = \mathbf{F}_1, \tag{2.1.72}$$

$$\mathbf{F}_3\mathbf{F}_1 = \mathbf{F}_2, \tag{2.1.73}$$

which can be summarized as

$$\mathbf{F}_j\mathbf{F}_k = -\delta_{jk}\mathbf{1} + \sum_{i=1}^{3}\epsilon_{ijk}\mathbf{F}_i. \tag{2.1.74}$$

The group of matrices $\{\pm\mathbf{1}, \pm\mathbf{F}_k\}$ is also isomorphic to the quaternion group. The matrices $\mathbf{B}_k$ are also closely related to the Pauli spin matrices. (Cf. e.g. [5].) In addition, while we are not aware of their prior application in rod dynamics, Seywald *et al.* [20] have exploited the matrices $\mathbf{F}_k$ and their commutation relations (2.1.74) to analyze the optimal control problem of reorienting a rigid spacecraft using the least amount of fuel. Each orthogonal transformation $\boldsymbol{\mathcal{F}}_j \equiv \exp(\frac{1}{2}\epsilon\mathbf{F}_j)$ corresponds to a rotation $\tilde{\boldsymbol{\mathcal{F}}}_j$ in $\mathbf{R}^3$ about the axis $\mathbf{e}_j$ through an angle ϵ. These transformations also have the property that

$$(\boldsymbol{\mathcal{F}}_j)^T \, \mathbf{B}_k \, \boldsymbol{\mathcal{F}}_j = \mathbf{B}_k. \tag{2.1.75}$$

This invariance of the $\mathbf{B}_k$ matrices is related to the fact that the angular velocities ω_j relative to the $\{\mathbf{d}_k\}$ basis are unchanged by a rigid rotation of the body.

Using the matrices $\mathbf{B}_k$ and $\mathbf{F}_k$ one can rewrite equation (2.1.52) in terms of the entries of matrix $\mathbf{A}$ as

$$\mathbf{d}_j \cdot \mathbf{e}_k = (\mathbf{d}_j)_k = |\mathbf{q}|^{\beta-2}\, \mathbf{q}\cdot\mathbf{D}_{jk}\mathbf{q}\ , \tag{2.1.76}$$

where the 4-by-4 matrix $\mathbf{D}_{jk}$ is given by

$$\mathbf{D}_{jk} \equiv (\mathbf{B}_j)^T\, \mathbf{F}_k = -\mathbf{B}_j\, \mathbf{F}_k\ . \tag{2.1.77}$$

One can show that each matrix $\mathbf{D}_{jk}$ is symmetric, which together with the skew-symmetry of the matrices $\mathbf{B}_k$ and $\mathbf{F}_k$ implies

$$\mathbf{D}_{jk} = (\mathbf{F}_k)^T\, \mathbf{B}_j = -\mathbf{F}_k\, \mathbf{B}_j\ . \tag{2.1.78}$$

$$\mathbf{D}_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.1.79}$$

$$\mathbf{D}_{12} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{2.1.80}$$

$$\mathbf{D}_{13} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \tag{2.1.81}$$

$$\mathbf{D}_{21} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{2.1.82}$$

$$\mathbf{D}_{22} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.1.83}$$

$$\mathbf{D}_{23} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{2.1.84}$$

$$\mathbf{D}_{31} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} , \tag{2.1.85}$$

$$\mathbf{D}_{32} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} , \tag{2.1.86}$$

$$\mathbf{D}_{33} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.1.87}$$

$$\tag{2.1.88}$$

The matrices $\{\pm\mathbf{1}, \pm\mathbf{B}_k, \pm\mathbf{F}_k, \pm\mathbf{D}_{jk}\}$ form a group under multiplication with 32 elements, in which the algebraic properties of $\mathbf{B}_j$ and $\mathbf{B}_k$ and the definition (2.1.77) of $\mathbf{D}_{jk}$ imply

$$\mathbf{D}_{jk}\,\mathbf{B}_j = \mathbf{F}_k \ , \tag{2.1.89}$$

$$\mathbf{D}_{jk}\,\mathbf{B}_l = -\sum_{m=1}^{3} \epsilon_{jlm}\,\mathbf{D}_{mk} \ , \qquad \text{for } l \neq j, \tag{2.1.90}$$

$$\mathbf{D}_{jk}\,\mathbf{F}_k = \mathbf{B}_j \ , \tag{2.1.91}$$

$$\mathbf{D}_{jk}\,\mathbf{F}_l = -\sum_{m=1}^{3} \epsilon_{klm}\,\mathbf{D}_{jm} \ , \qquad \text{for } l \neq k, \tag{2.1.92}$$

$$\mathbf{D}_{ij}\,\mathbf{D}_{ij} = \mathbf{1} \ , \tag{2.1.93}$$

$$\mathbf{D}_{ij}\,\mathbf{D}_{il} = -\sum_{m=1}^{3} \epsilon_{jlm}\,\mathbf{F}_m \ , \qquad \text{for } l \neq j, \tag{2.1.94}$$

$$\mathbf{D}_{ij}\,\mathbf{D}_{kj} = \sum_{m=1}^{3} \epsilon_{ikm}\,\mathbf{B}_m \ , \qquad \text{for } k \neq i, \tag{2.1.95}$$

$$\mathbf{D}_{ij}\,\mathbf{D}_{kl} = -\sum_{m,n=1}^{3} \epsilon_{ikm}\,\epsilon_{jln}\,\mathbf{D}_{mn} \ , \qquad \text{for } k \neq i, \ l \neq j. \tag{2.1.96}$$

I am not aware of any prior work in kinematics where this group has been recorded.

2.2 Euler Angles

There are 12 distinct Euler angle sequences. Here we will consider only the sequence generated by a rotation about the 3-axis through an angle ϕ, followed by a rotation about the resulting 2-axis through an angle θ, followed by a rotation about the resulting 3-axis through an angle ψ. The Euler angle sequence will be denoted by $\boldsymbol{\phi} = (\phi, \theta, \psi)^T$, where the superscript T denotes the transpose. This representation the singularity is avoided so long as the angle θ is near $\pi/2$.

For our choice of coordinates,

$$\mathbf{A} = \mathbf{e}_j \otimes \mathbf{d}_j = \begin{pmatrix} c\psi c\theta c\phi - s\psi s\phi & c\psi c\theta s\phi + s\psi c\phi & -c\psi s\theta \\ -s\psi c\theta c\phi - c\psi s\phi & -s\psi c\theta s\phi + c\psi c\phi & s\psi s\theta \\ s\theta c\phi & s\theta s\phi & c\theta \end{pmatrix}, \tag{2.2.1}$$

where c denotes the cosine and s denotes the sine. The three rows of $\mathbf{A}$ specify the coordinates in the reference frame of $\mathbf{d}_1$, $\mathbf{d}_2$, and $\mathbf{d}_3$, respectively. Note that $\mathbf{d}_3$ is independent of the angle ψ.

Following the same procedure used in the previous section, the body components of angular velocity can be related to the Euler angles as

$$\omega_j = \mathbf{b}^j[\mathbf{q}] \cdot \dot{\boldsymbol{\phi}} , \tag{2.2.2}$$

where

$$\begin{aligned} \mathbf{b}^1(\boldsymbol{\phi}) &= (-c\psi s\theta, \quad s\psi, \quad 0)^T , &(2.2.3)\\ \mathbf{b}^2(\boldsymbol{\phi}) &= (\ \ s\psi s\theta, \quad c\psi, \quad 0)^T , &(2.2.4)\\ \mathbf{b}^3(\boldsymbol{\phi}) &= (\qquad c\theta, \quad 0, \quad 1)^T . &(2.2.5)\end{aligned}$$

Equation (2.2.2) is commonly written in the form

$$\underline{\boldsymbol{\omega}} = \mathbf{B}[\boldsymbol{\phi}]\,\dot{\boldsymbol{\phi}}, \tag{2.2.6}$$

where

$$\mathbf{B}[\boldsymbol{\phi}] = \mathbf{e}_j \otimes \mathbf{b}^j(\boldsymbol{\phi}) = \begin{pmatrix} -c\psi s\theta & s\psi & 0 \\ s\psi s\theta & c\psi & 0 \\ c\theta & 0 & 1 \end{pmatrix}. \tag{2.2.7}$$

Using the bases $\{\mathbf{d}_k\}$ and $\{\mathbf{b}^k[\boldsymbol{\phi}]\}$, we can express the derivative of each vector $\mathbf{d}_k$ as

$$\mathbf{D}_k[\boldsymbol{\phi}] \equiv \left(\frac{\partial \mathbf{d}_1}{\partial \boldsymbol{\phi}}\right)^T = \epsilon_{ijk}\mathbf{B}_j\mathbf{q} \otimes \mathbf{d}_i . \tag{2.2.8}$$

In analogy with equation (2.1.29), we observe the identity

$$\mathbf{D}_k[\boldsymbol{\phi}]\mathbf{d}_k(\boldsymbol{\phi}) = \mathbf{0} \qquad \text{(no sum)}. \tag{2.2.9}$$

which follows from the fact that $\mathbf{d}_k$ is a unit vector.

Assuming $\sin\theta \neq 0$, the vectors $\{\mathbf{b}^j(\boldsymbol{\phi})\}$ define a basis for $\mathbf{R}^3$. However this basis is not orthonormal in general, so it is convenient to also define its dual basis $\{\mathbf{b}_j(\boldsymbol{\phi})\}$ which satisfies the identity

$$\mathbf{b}^j \cdot \mathbf{b}_k = \delta_{jk}. \tag{2.2.10}$$

The vectors $\{\mathbf{b}_j(\boldsymbol{\phi})\}$ are therefore the columns of the inverse of $\mathbf{B}[\boldsymbol{\phi}]$, so that

$$\mathbf{B}[\boldsymbol{\phi}]^{-1} = \mathbf{b}_j(\boldsymbol{\phi}) \otimes \mathbf{e}_j = \begin{pmatrix} \mathbf{b}_1(\boldsymbol{\phi}) & \mathbf{b}_2(\boldsymbol{\phi}) & \mathbf{b}_3(\boldsymbol{\phi}) \end{pmatrix} \tag{2.2.11}$$

and

$$\mathbf{b}_1(\boldsymbol{\phi}) = \begin{pmatrix} -c\psi/s\theta \\ s\psi \\ c\psi c\theta/s\theta \end{pmatrix}, \quad \mathbf{b}_2(\boldsymbol{\phi}) = \begin{pmatrix} s\psi/s\theta \\ c\psi \\ -s\psi c\theta/s\theta \end{pmatrix}, \quad \mathbf{b}_3(\boldsymbol{\phi}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.2.12}$$

The identity (2.2.10) together with the forms of the matrices $\mathbf{D}_j[\boldsymbol{\phi}]$ imply

$$\mathbf{D}_k[\boldsymbol{\phi}]^T \mathbf{b}_k(\boldsymbol{\phi}) = \mathbf{0} \qquad \text{(no sum)} . \tag{2.2.13}$$

We infer from equation (2.2.8) that, for small ϵ, a change in $\boldsymbol{\phi}$ along the direction $\epsilon\mathbf{b}_k(\boldsymbol{\phi})$ produces a rotation about the axis $\mathbf{d}_k(\boldsymbol{\phi})$ through an angle ϵ. For example, $\mathbf{D}_1[\boldsymbol{\phi}]^T\,\epsilon\mathbf{b}_1(\boldsymbol{\phi}) = \mathbf{0}$ while $\mathbf{D}_2[\boldsymbol{\phi}]^T\,\epsilon\mathbf{b}_1(\boldsymbol{\phi}) = \epsilon\mathbf{d}_3(\boldsymbol{\phi})$ and $\mathbf{D}_3[\boldsymbol{\phi}]^T\,\epsilon\mathbf{b}_1(\boldsymbol{\phi}) = -\epsilon\mathbf{d}_2(\boldsymbol{\phi})$.

Using equation (2.2.6) we can express the angular momentum with respect the reference frame as

$$\boldsymbol{\omega} = \mathbf{F}[\boldsymbol{\phi}]\,\dot{\boldsymbol{\phi}} \tag{2.2.14}$$

where

$$\mathbf{F}[\boldsymbol{\phi}] \equiv \mathbf{A}^T[\boldsymbol{\phi}]\mathbf{B}[\boldsymbol{\phi}] = \begin{pmatrix} 0 & -s\phi & s\theta c\phi \\ 0 & c\phi & s\theta s\phi \\ 1 & 0 & c\theta \end{pmatrix}. \tag{2.2.15}$$

3 Rodrigues' Composition of Rotations Formula

In this section we present a derivation of Rodrigues' formula of the composition of two rotations. This formula leads naturally to the definition of the Euler-Rodrigues parameters. The derivation parallels that of Rodrigues in his 1840 paper.

3.1 One Plane Rotation = Two Reflections

The key observation is the following proposition:

Proposition 1. *A plane rotation through an angle ϕ is equivalent to two successive reflections across lines which are separated by an angle $\phi/2$.*

Figure 1 indicates graphically why Proposition 1 is plausible.

Without loss of generality, we can choose the first line, L_1, to be the horizontal axis. Let $\mathbf{b}_1$ be the unit vector along the positive horizontal axis, and let $\mathbf{a}$ be the unit vector at an angle θ from $\mathbf{b}_1$. Then $\mathbf{a}_1$, the reflection of $\mathbf{a}$ across L_1, lies at an angle $-\theta$ from $\mathbf{b}_1$. Now let $\mathbf{b}_2$ be the unit vector along L_2, lying at an angle $\phi/2$ from $\mathbf{b}_1$, and let $\mathbf{a}_2$ be the reflection of $\mathbf{a}_1$ across L_2. Because $\mathbf{a}_1$ lies at an angle $-\theta - \phi/2$ from $\mathbf{b}_2$, its reflection $\mathbf{a}_2$ lies at an angle $\theta + \phi/2$ from $\mathbf{b}_2$, or equivalently, at an angle $\phi/2 + \theta + \phi/2 = \theta + \phi$ from $\mathbf{b}_1$. That is, $\mathbf{a}$ has been rotated to $\mathbf{a}_2$ through an angle ϕ.

We now prove Proposition 1 using vector algebra. Let $\mathbf{k}$ be the unit vector normal to the plane of rotation and let $\mathbf{b} \equiv \mathbf{k} \times \mathbf{n}$ so that $\{\mathbf{n}, \mathbf{b}, \mathbf{k}\}$ forms a right-handed orthonormal basis. The reflection across the plane P orthogonal to $\mathbf{n}$ is given by the operator

$$\mathbf{R}(\mathbf{n}) \equiv \mathbf{1} - 2\mathbf{n} \otimes \mathbf{n}. \tag{3.1.1}$$

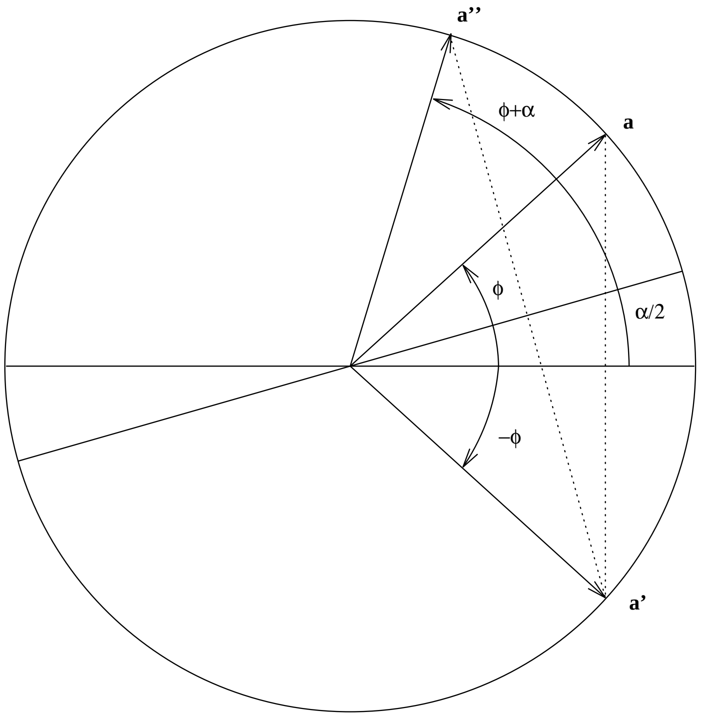

Figure 1.

The vector $\mathbf{n}'$ is given by

$$\mathbf{n}' = \mathbf{n}\cos(\phi/2) + \mathbf{b}\sin(\phi/2). \tag{3.1.2}$$

Therefore

$$\begin{aligned} \mathbf{R}(\mathbf{n}') &= \mathbf{1} - 2\mathbf{n}' \otimes \mathbf{n}' \\ &= 1 - 2\cos^2(\phi/2)\mathbf{n} \otimes \mathbf{n} - 2\sin(\phi/2)\cos(\phi/2)\mathbf{n} \otimes \mathbf{b} \\ &\quad -2\sin(\phi/2)\cos(\phi/2)\mathbf{b} \otimes \mathbf{n} - 2\sin^2(\phi/2)\mathbf{b} \otimes \mathbf{b}. \end{aligned} \tag{3.1.3}$$

Using the identities $2\cos^2(\phi/2) = 1 + \cos\phi$, $2\sin^2(\phi/2) = 1 - \cos\phi$ and $2\sin(\phi/2)\cos(\phi/2) = \sin\phi$, we then obtain

$$\begin{aligned} \mathbf{R}(\mathbf{n}') &= \mathbf{1} - (1 + \cos\phi)\mathbf{n} \otimes \mathbf{n} - \sin(\phi)\mathbf{n} \otimes \mathbf{b} \\ &\quad - \sin(\phi)\mathbf{b} \otimes \mathbf{n} - (1 - \cos\phi)\mathbf{b} \otimes \mathbf{b}. \end{aligned} \tag{3.1.4}$$

After simplification, the composition of the two reflections (3.1.1) and (3.1.4) can be expressed as

$$\begin{aligned} \mathbf{Q} \equiv \mathbf{R}(\mathbf{n}')\mathbf{R}(\mathbf{n}) &= \mathbf{1} - (1 - \cos\phi)(\mathbf{n} \otimes \mathbf{n} + \mathbf{b} \otimes \mathbf{b}) - \sin(\phi)(\mathbf{b} \otimes \mathbf{b} - \mathbf{n} \otimes \mathbf{b} \\ &= \cos(\phi)\mathbf{1} + (1 - \cos\phi)\mathbf{k} \otimes \mathbf{k} + \sin(\phi)\,[\mathbf{k}\times]. \end{aligned} \tag{3.1.5}$$

A comparison of equation (3.1.5) with equation (2.1.2) shows that $\mathbf{Q}$ is a rotation about the axis $\mathbf{k}$ through an angle ϕ. We will denote this rotation by $\mathbf{Q}(\mathbf{k}, \phi)$.

3.2 The Composition of Rotations

We can use Proposition 1 to obtain a simple expression for the composition of two rotations

$$\mathbf{Q}(\mathbf{k}, \phi) \equiv \mathbf{Q}(\mathbf{k}_2, \phi_2)\mathbf{Q}(\mathbf{k}_1, \phi_1). \tag{3.2.6}$$

Figure 2 illlustrates the process.

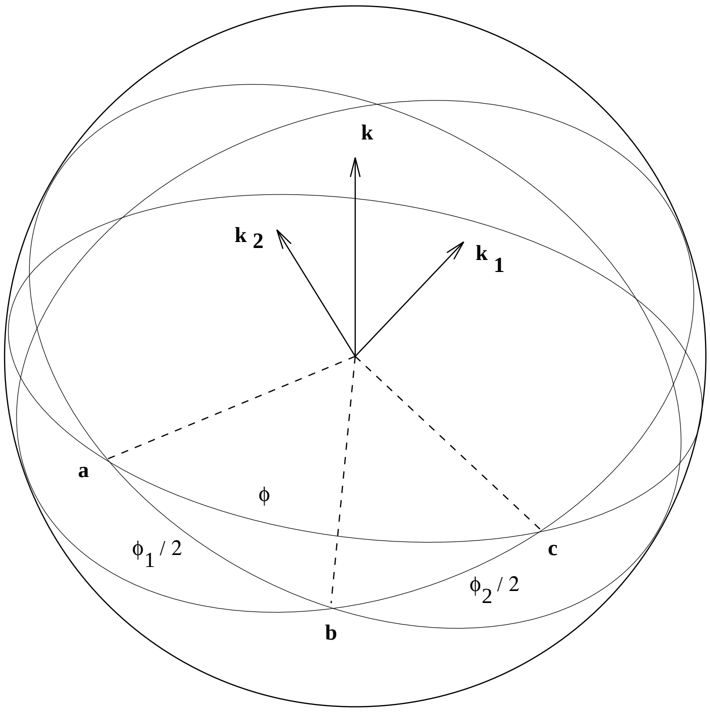

Figure 2.

The calculation is trivial if $\mathbf{k}_1$ and $\mathbf{k}_2$ are collinear, so assume $\mathbf{k}_1 \times \mathbf{k}_2 \neq \mathbf{0}$ and define the unit vector

$$\mathbf{b} \equiv (\mathbf{k}_1 \times \mathbf{k}_2) / |\mathbf{k}_1 \times \mathbf{k}_2| \tag{3.2.7}$$

which lies in the planes orthogonal to both $\mathbf{k}_1$ and $\mathbf{k}_2$. Let us also define the unit vectors

$$\mathbf{a} \equiv \mathbf{b}\cos(\tfrac{1}{2}\phi_1) - (\mathbf{k}_1 \times \mathbf{b})\sin(\tfrac{1}{2}\phi_1), \tag{3.2.8}$$
$$\mathbf{c} \equiv \mathbf{b}\cos(\tfrac{1}{2}\phi_2) + (\mathbf{k}_2 \times \mathbf{b})\sin(\tfrac{1}{2}\phi_2). \tag{3.2.9}$$

We can use Proposition 1 to observe that

$$\mathbf{Q}(\mathbf{k}_1, \phi_1) = \mathbf{R}(\mathbf{b})\mathbf{R}(\mathbf{a}) \tag{3.2.10}$$

and

$$\mathbf{Q}(\mathbf{k}_2, \phi_2) = \mathbf{R}(\mathbf{c})\mathbf{R}(\mathbf{b}). \tag{3.2.11}$$

Because the composition of the reflection $R(\mathbf{c})$ with itself is the identity, equations (3.2.6), (3.2.10) and (3.2.11) together imply

$$\mathbf{Q}(\mathbf{k}, \phi) = \mathbf{R}(\mathbf{c})\mathbf{R}(\mathbf{a}). \tag{3.2.12}$$

We can now apply Proposition 1 again to equation (3.2.12) to conclude that the net angle of rotation ϕ is twice the angle between $\mathbf{a}$ and $\mathbf{c}$, or

$$\cos(\tfrac{1}{2}\phi) = \mathbf{a} \cdot \mathbf{c}, \tag{3.2.13}$$

and the net axis of rotation $\mathbf{k}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{c}$, so

$$\mathbf{k} = (\mathbf{a} \times \mathbf{c}) \,/ |\mathbf{a} \times \mathbf{c}|. \tag{3.2.14}$$

Equation (3.2.13) implies that $\sin(\frac{1}{2}\phi) = |\mathbf{a} \times \mathbf{c}|$ (note there is no sign ambiguity, because $\sin(\frac{1}{2}\phi) \geq 0$), which together with equation (3.2.14) implies

$$\sin(\tfrac{1}{2}\phi)\mathbf{k} = \mathbf{a} \times \mathbf{c}. \tag{3.2.15}$$

Substituting equations (3.2.8) and (3.2.9) into equations (3.2.13) and (3.2.15) we obtain

$$\cos(\tfrac{1}{2}\phi) = \cos(\tfrac{1}{2}\phi_1)\cos(\tfrac{1}{2}\phi_2) - \sin(\tfrac{1}{2}\phi_1)\sin(\tfrac{1}{2}\phi_2)\,(\mathbf{k}_1 \times \mathbf{b}) \cdot (\mathbf{k}_2 \times \mathbf{b}) \tag{3.2.16}$$

and

$$\begin{aligned} \sin(\tfrac{1}{2}\phi)\mathbf{k} \;\; = \;\; & \cos(\tfrac{1}{2}\phi_1)\sin(\tfrac{1}{2}\phi_2)\,\mathbf{b} \times (\mathbf{k}_2 \times \mathbf{b}) \\ & - \cos(\tfrac{1}{2}\phi_2)\sin(\tfrac{1}{2}\phi_1)\;(\mathbf{k}_1 \times \mathbf{b}) \times \mathbf{b} \\ & - \sin(\tfrac{1}{2}\phi_1)\sin(\tfrac{1}{2}\phi_2)\;(\mathbf{k}_1 \times \mathbf{b}) \times (\mathbf{k}_2 \times \mathbf{b})\,. \end{aligned} \tag{3.2.17}$$

To simplify equations (3.2.16) and (3.2.17) we use the vector identities

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \;\; = \;\; \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) \tag{3.2.18}$$
$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) \;\; = \;\; (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}), \tag{3.2.19}$$

and the definition (3.2.7) of $\mathbf{b}$ to obtain

$$\mathbf{b} \times (\mathbf{k}_2 \times \mathbf{b}) \;\; = \;\; \mathbf{k}_2, \tag{3.2.20}$$
$$(\mathbf{k}_1 \times \mathbf{b}) \times \mathbf{b} \;\; = \;\; -\mathbf{k}_1, \tag{3.2.21}$$
$$(\mathbf{k}_1 \times \mathbf{b}) \times (\mathbf{k}_2 \times \mathbf{b}) \;\; = \;\; \mathbf{k}_1 \cdot \mathbf{k}_2 \tag{3.2.22}$$

Finally we can insert the identities (3.2.20) – (3.2.22) into equations (3.2.16) and (3.2.17) to obtain

$$\cos(\tfrac{1}{2}\phi) = \cos(\tfrac{1}{2}\phi_1)\cos(\tfrac{1}{2}\phi_2) - \left[\sin(\tfrac{1}{2}\phi_1)\mathbf{k}_1\right] \cdot \left[\sin(\tfrac{1}{2}\phi_2)\mathbf{k}_2\right] \tag{3.2.23}$$

and

$$\begin{aligned} \sin(\tfrac{1}{2})\mathbf{k} \;\; = \;\; & \cos(\tfrac{1}{2}\phi_1)\left[\sin(\tfrac{1}{2}\phi_2)\,\mathbf{k}_2\right] + \cos(\tfrac{1}{2}\phi_2)\left[\sin(\tfrac{1}{2}\phi_1)\,\mathbf{k}_1\right] \\ & + \left[\sin(\tfrac{1}{2}\phi_1)\,\mathbf{k}_1\right] \times \left[\sin(\tfrac{1}{2}\phi_2)\,\mathbf{k}_2\right]. \end{aligned} \tag{3.2.24}$$

Equations (3.2.23) and (3.2.24) provided the initial motivation for the definition of the Euler-Rodrigues parameters, because with them one can compute the composition of two rotations in a very simple manner in terms of the variables $\cos(\frac{1}{2}\phi_j)$ and $\sin(\frac{1}{2}\phi_j)\,\mathbf{k}_j$.

4 Factorizations of Orthogonal Transformations

In mechanics it is sometimes desirable to represent an orthogonal transformation as a product of transformations which are in some sense simpler. Perhaps the most famous factorization of an orthogonal transformation in $\mathbf{R}^3$ is the Euler angle representation, as a composition of three elementary rotations. Less well-known is Rodrigues' representation of a rotation in $\mathbf{R}^3$ as a composition of two reflections.

Yet another factorization which is relevant to our discussion is credited to Cayley (reference to be found), who showed that, for any skew transformation $\mathbf{C}$ acting on $\mathbf{R}^n$, the so-called Cayley transformation

$$\mathbf{A} \equiv (\mathbf{1} + \mathbf{C})^{-1}(\mathbf{1} - \mathbf{C}) \tag{4.0.1}$$

is an orthogonal transformation. However, the converse is not true. In fact, an orthogonal transformation $\mathbf{A}$ can be represented in the form (4.0.1) if and only if $\mathbf{A}$ does not have -1 as an eigenvalue. To see this, suppose $\mathbf{A}$ is given by equation (4.0.1), and that $\mathbf{A}$ has eigenvalue -1 with eigenvector $\mathbf{n}$. Then if we premultiply (4.0.1) by $(\mathbf{1} + \mathbf{C})$ and postmultiply by $\mathbf{n}$, we obtain $bn = \mathbf{n}$, a contradiction. Conversely, if $\mathbf{A}$ does not have an eigenvalue -1, then the skew transformation $\mathbf{C}$ given by

$$\mathbf{C} \equiv (\mathbf{1} + \mathbf{A})^{-1}(\mathbf{1} - \mathbf{A}) \tag{4.0.2}$$

satisfies equation (4.0.1).

In these notes we are concerned primarily with $\mathbf{R}^3$, in which a skew transformation can be expressed as $[\mathbf{c}\times]$ for some vector $\mathbf{c}$, so that equation (4.0.1) takes the form

$$\mathbf{A} \equiv (\mathbf{1} + [\mathbf{c}\times])^{-1}(\mathbf{1} - [\mathbf{c}\times]). \tag{4.0.3}$$

In the trivial case $\mathbf{c} = \mathbf{0}$, equation (4.0.3) yields $\mathbf{A} = \mathbf{1}$. We will therefore restrict our attention to the case where $\mathbf{c} \neq \mathbf{0}$. We first note that one can write the inverse of $\mathbf{1} + [\mathbf{c}\times]$ explicitly as

$$(\mathbf{1} + [\mathbf{c}\times])^{-1} = \frac{1}{1 + |\mathbf{c}|^2}(\mathbf{1} - [\mathbf{c}\times] + \mathbf{c} \otimes \mathbf{c}). \tag{4.0.4}$$

Equation (4.0.4) can be obtained formally from the series expansion

$$(\mathbf{1} + [\mathbf{c}\times])^{-1} = \sum_{j=0}^{\infty}(-1)^j\,[\mathbf{c}\times]^j, \tag{4.0.5}$$

together with the vector identity (3.2.18) which implies

$$\begin{aligned} [\mathbf{c}\times]^2 &= \mathbf{c} \otimes \mathbf{c} - |\mathbf{c}|^2\mathbf{1}, && (4.0.6)\\ [\mathbf{c}\times]^3 &= -|\mathbf{c}|^2\,[\mathbf{c}\times]\,, && (4.0.7) \end{aligned}$$

and in general

$$\begin{aligned} [\mathbf{c}\times]^{2n} &= (-1)^{(n-1)}|\mathbf{c}|^{2(n-1)}\left(\mathbf{c} \otimes \mathbf{c} - |\mathbf{c}|^2\mathbf{1}\right), && (4.0.8)\\ [\mathbf{c}\times]^{(2n-1)} &= (-1)^n|\mathbf{c}|^{2(n-1)}\,[\mathbf{c}\times]\,, && (4.0.9) \end{aligned}$$

for $n > 1$. If we then introduce equation (4.0.4) into equation (4.0.3) and simplify, we find that $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \left(\frac{1-|\mathbf{c}|^2}{1+|\mathbf{c}|^2}\right)\mathbf{1} + \left(\frac{2|\mathbf{c}|^2}{1+|\mathbf{c}|^2}\right)\mathbf{k}\otimes\mathbf{k} - \left(\frac{2|\mathbf{c}|}{1+|\mathbf{c}|^2}\right)[\mathbf{k}\times], \tag{4.0.10}$$

where $\mathbf{k}$ is the unit vector

$$\mathbf{k} \equiv \mathbf{c}/|\mathbf{c}|. \tag{4.0.11}$$

A comparison with equation (2.1.2) shows that in (4.0.10), $\mathbf{k}$ is the axis of rotation and the angle of rotation in given by

$$\tan(\phi/2) = |\mathbf{c}|. \tag{4.0.12}$$

Equation (4.0.12) confirms our previous observation that a rotation $\mathbf{A}$ through an angle π, which has an eigenvalue -1, cannot be represented in the form (4.0.1). The vector

$$\mathbf{c} = (c_1, c_2, c_3) = \tan(\phi/2)\,\mathbf{k} \tag{4.0.13}$$

is called the Gibbs vector, and its components are called the Rodrigues parameters. The singularity we have just observed is yet another manifestation of the topological theorem that any three-parameter representation of $SO(3)$ must have a singularity. The Rodrigues parameters are related to the Euler parameters by

$$c_j = q_j/q_4. \tag{4.0.14}$$

Finally, the Cayley transform (4.0.3) is closely related to the factorization (2.1.52), since we can express the 3-by-4 matrices $\mathbf{B}[\mathbf{q}]$ and $\mathbf{B}^o[\mathbf{q}]$ in the partitioned forms

$$\mathbf{B}[\mathbf{q}] = (\cos(\phi/2)\,\mathbf{1} + [\sin(\phi/2)\,\mathbf{k}\times] \quad | \quad -\sin(\phi/2)\,\mathbf{k})\,, \tag{4.0.15}$$

$$\mathbf{B}[\mathbf{q}] = (\cos(\phi/2)\,\mathbf{1} - [\sin(\phi/2)\,\mathbf{k}\times] \quad | \quad -\sin(\phi/2)\,\mathbf{k})\,. \tag{4.0.16}$$

Question for further research: It was Jacobi who pointed out to Cayley, I think, that the parameters Rodrigues obtained in is analysis of rigid body motions were first developed by Euler. I am told Euler developed his parametrization of orthogonal transformation as a mans to study Diophantine equations in 4 dimensions, and he did not recognize their geometrical significance. Where does this work appear?

5 Algebras

This section contains some observations about the algebras arising in dimensions two, three and four, respectively complex algebra, vector algebra (which forms a Lie algebra), and quaternion algebra. Most of the remarks regarding complex and vector algebra are made for comparison and contrast with the quaternions algebra.

5.1 Dimension 2: Complex Algebra

A complex number can be written in the form $\mathbf{q} = q_1 i + q_2$, where q_1 and q_2 are real numbers, and $i^2 = -1$. It follows that the general multiplication rule for complex numbers is given by

$$\mathbf{q}\,\mathbf{p} = (q_1 i + q_2)(p_1 i + p_2) = (q_2 p_1 + q_1 p_2)i + (q_2 p_2 - q_1 p_1) \ . \tag{5.1.1}$$

In particular, complex multiplication is commutative.

The product $\mathbf{q}\,\mathbf{p}$ can be viewed as a bilinear operation on $\mathbf{q}$ and $\mathbf{p}$. We can therefore express the product as

$$\mathbf{q}\,\mathbf{p} = \mathbf{L}(\mathbf{q}) \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \ , \tag{5.1.2}$$

where

$$\mathbf{L}(\mathbf{q}) \equiv \begin{pmatrix} q_2 & q_1 \\ -q_1 & q_2 \end{pmatrix} \tag{5.1.3}$$

The matrix $\mathbf{L}(\mathbf{q})$ can be decomposed as

$$\mathbf{L}(\mathbf{q}) = q_1 \mathbf{J} + q_2 \mathbf{1} \ , \tag{5.1.4}$$

where $\mathbf{1}$ is the 2-by-2 identity matrix, and

$$\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{5.1.5}$$

Because multiplication is commutative, we could also express the product as $\mathbf{q}\,\mathbf{p} = L(\mathbf{p})\,\mathbf{q}$.

If we further identify each complex number $q_1 i + q_2$ with a 2-by-2 real matrix of the form (5.1.3), then the standard matrix product of $\mathbf{L}(\mathbf{q})$ with $\mathbf{L}(\mathbf{p})$ is isomorphic with the complex product. That is,

$$\mathbf{L}(\mathbf{q}\,\mathbf{p}) = \mathbf{L}(\mathbf{q})\,\mathbf{L}(\mathbf{p}) \ . \tag{5.1.6}$$

5.2 Dimension 3: Vector Algebra

To be completed. We will summarize the vector (cross) and scalar (dot) products for 3-vectors, and note the well-known inequalities

$$|\mathbf{a} \times \mathbf{b}| \leq |\mathbf{a}|\,|\mathbf{b}| \ , \tag{5.2.1}$$

with equality when $\mathbf{a} \perp \mathbf{b}$,

$$|\mathbf{a} \cdot \mathbf{b}| \leq |\mathbf{a}|\,|\mathbf{b}| \ , \tag{5.2.2}$$

with equality when $\mathbf{a} = \pm\mathbf{b}$, as well as the equality

$$|\mathbf{a} \times \mathbf{b}|^2 + |\mathbf{a} \cdot \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \ . \tag{5.2.3}$$

5.3 Dimension 4: Quaternion Algebra

The set of quaternions $\mathbf{H}$ is defined as a skew field (i.e. an algebra which satisfies all the properties of a field except the commutativity of multiplication) with elements of the form

$$\mathbf{q} \equiv q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} + q_4\mathbf{1}, \tag{5.3.1}$$

where the q_1, q_2, q_3 and q_4 are real numbers, $\mathbf{1}$ is the identity element, and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfy

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1. \tag{5.3.2}$$

The relations (5.3.2) imply the multiplication rules

$$\begin{array}{lll} \mathbf{i}\,\mathbf{j} = \mathbf{k}\,, & \mathbf{j}\,\mathbf{k} = \mathbf{i}\,, & \mathbf{k}\,\mathbf{i} = \mathbf{j}\,, \\ \mathbf{j}\,\mathbf{i} = -\mathbf{k}\,, & \mathbf{k}\,\mathbf{j} = -\mathbf{i}\,, & \mathbf{i}\,\mathbf{k} = -\mathbf{j}\,. \end{array} \tag{5.3.3}$$

It follows that the product of two general quaternions has the form

$$(q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} + q_4\mathbf{1})(p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k} + p_4\mathbf{1}) = r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k} + r_4\mathbf{1}\ , \tag{5.3.4}$$

where

$$r_1 = q_4p_1 - q_3p_2 + q_2p_3 + q_1p_4\ , \tag{5.3.5}$$
$$r_2 = q_3p_1 + q_4p_2 - q_1p_3 + q_2p_4\ , \tag{5.3.6}$$
$$r_3 = -q_2p_1 + q_1p_2 + q_4p_3 + q_3p_4\ , \tag{5.3.7}$$
$$r_4 = -q_1p_1 - q_2p_2 - q_3p_3 + q_4p_4\ . \tag{5.3.8}$$

This multiplication has the property that it preserves the modulus of the quaternions:

$$|\mathbf{q}\mathbf{p}| = |\mathbf{q}||\mathbf{p}|, \tag{5.3.9}$$

where

$$|\mathbf{q}| \equiv \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}. \tag{5.3.10}$$

Hamilton called equation (5.3.9) the “Law of Moduli”. It was this desired property primarily which suggested to Hamilton the multiplication rule for quaternions which is given by equations (5.3.4) - (5.3.8). (Cf. [24].)

It is sometimes convenient denote the quaternion by $(\vec{\mathbf{q}}, \mathbf{q}_4)$ where $\vec{\mathbf{q}} \equiv \mathbf{q}_1\mathbf{i} + \mathbf{q}_2\mathbf{j} + \mathbf{q}_3\mathbf{k}$ is the *vector* part of the quaternion and q_4 is the *scalar* part. Which decomposition is especially appropriate when we associate Euler parameters with quaternions. Using this notation the quaternion product can be rewritten as

$$(\vec{\mathbf{q}}, \mathbf{q}_4)(\vec{\mathbf{p}}, \mathbf{p}_4) = (\mathbf{q}_4\vec{\mathbf{p}} + \mathbf{p}_4\vec{\mathbf{q}} + \vec{\mathbf{q}} \times \vec{\mathbf{p}},\ \mathbf{q}_4\mathbf{p}_4 - \vec{\mathbf{q}} \cdot \vec{\mathbf{p}}). \tag{5.3.11}$$

In particular, the product of two purely imaginary quaternions, for which the scalar part is zero, is

$$(\vec{\mathbf{q}}, 0)(\vec{\mathbf{p}}, 0) = (\vec{\mathbf{q}} \times \vec{\mathbf{p}},\ -\vec{\mathbf{q}} \cdot \vec{\mathbf{p}}). \tag{5.3.12}$$

Historically, this observation led J.W. Gibbs to the modern formulation of vector analysis. (Cf [3].)

We can also exploit the bilinear structure of the product $\mathbf{qp}$ to express it in the form

$$\mathbf{q}\,\mathbf{p} = \mathbf{L}(\mathbf{q}) \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} , \tag{5.3.13}$$

where

$$\mathbf{L}(\mathbf{q}) = \begin{pmatrix} q_4 & -q_3 & q_2 & q_1 \\ q_3 & q_4 & -q_1 & q_2 \\ -q_2 & q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{pmatrix} \tag{5.3.14}$$

Unlike the cases of complex and vector multiplication, however, we can also express the same product in a distinctly different form, as

$$\mathbf{q}\,\mathbf{p} = \mathbf{R}(\mathbf{p}) \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} , \tag{5.3.15}$$

where

$$\mathbf{R}(\mathbf{p}) = \begin{pmatrix} p_4 & p_3 & -p_2 & p_1 \\ -p_3 & p_4 & p_1 & p_2 \\ p_2 & -p_1 & p_4 & p_3 \\ -p_1 & -p_2 & -p_3 & p_4 \end{pmatrix} \tag{5.3.16}$$

The matrices $\mathbf{L}(\mathbf{q})$ and $\mathbf{R}(\mathbf{p})$ can be decomposed with respect to the bases $\{\mathbf{B}_j\}$ and $\{\mathbf{F}_j\}$, as

$$\mathbf{L}(\mathbf{q}) = q_1\mathbf{F}_1 + q_2\mathbf{F}_2 + q_3\mathbf{F}_3 + q_4\mathbf{1} \tag{5.3.17}$$

and

$$\mathbf{R}(\mathbf{p}) = -p_1\mathbf{B}_1 - p_2\mathbf{B}_2 - p_3\mathbf{B}_3 + p_4\mathbf{1}\ . \tag{5.3.18}$$

Because the basis $\{\mathbf{F}_j\}$ arises in the matrix $\mathbf{L}(\mathbf{q})$, we are motivated to denote the quaternion product $\mathbf{q}\,\mathbf{p}$ using an alternative notation, as $(\mathbf{q},\mathbf{p})_f$. The product $(\mathbf{q},\mathbf{p})_f$ therefore describes a rotation of $\mathbf{p}$, described by the quaternion $\mathbf{q}$, with respect to the fixed frame. This notation will be used below.

The quaternion multiplication just described provides a second way to express the rotation defined by the quaternion $\mathbf{q}$. Suppose that the application of the rotation $\mathbf{A}(\mathbf{q})$ defined by equation (2.1.6) to the vector $\vec{\mathbf{v}} \in \mathbf{R}^3$ yields vector

$$\vec{\mathbf{w}} = \mathbf{A}(\mathbf{q})\vec{\mathbf{v}} \tag{5.3.19}$$

If we define the quaternions $\mathbf{v} = (\vec{\mathbf{v}}, 0)$ and $\mathbf{w} = (\vec{\mathbf{w}}, 0)$, then the relation (5.3.19) is equivalent to

$$\mathbf{w} = \mathbf{q}\,\mathbf{v}\,\mathbf{q}^{-1} = \frac{\mathbf{q}\,\mathbf{v}\,\bar{\mathbf{q}}}{|\mathbf{q}|^2}\ . \tag{5.3.20}$$

To see that (5.3.20) is equivalent to (5.3.19), we observe that

$$\frac{1}{|\mathbf{q}|^2}\,\mathbf{q}\,\mathbf{v}\,\bar{\mathbf{q}} = \frac{1}{|\mathbf{q}|^2}\,\mathbf{R}\left(\bar{\mathbf{q}}\right)\,\mathbf{L}\left(\mathbf{q}\right)\mathbf{v}\ . \tag{5.3.21}$$

Moreover,

$$\mathbf{L}(\mathbf{q}) = \begin{pmatrix} \mathbf{F}(\mathbf{q})^T & \mathbf{q} \end{pmatrix} \tag{5.3.22}$$

and

$$\mathbf{R}\left(\bar{\mathbf{q}}\right) = \begin{pmatrix} \mathbf{B}(\mathbf{q}) \\ \mathbf{q}^T \end{pmatrix} \tag{5.3.23}$$

It then follows from equations (5.3.21) and (2.1.52) and the definition of $\mathbf{v}$ that

$$\frac{1}{|\mathbf{q}|^2}\,\mathbf{q}\,\mathbf{v}\,\bar{\mathbf{q}} = \frac{1}{|\mathbf{q}|^2}\begin{pmatrix} \mathbf{A}(\mathbf{q}) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}\begin{pmatrix} \vec{\mathbf{v}} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{A}(\mathbf{q})\vec{\mathbf{v}} \\ \mathbf{0} \end{pmatrix}\ , \tag{5.3.24}$$

which was to be proved.

There are various matrix representations of the quaternions. For example, we can identify each quaternion with a matrix of the form

$$\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}, \tag{5.3.25}$$

where

$$\begin{aligned} z &\equiv q_4 + iq_3, & (5.3.26)\\ w &\equiv q_2 + iq_1. & (5.3.27) \end{aligned}$$

The complex numbers z and w are called the *Cayley-Klein parameters* [5]. The matrix given in (5.3.25) can be decomposed as

$$i\left(q_1\boldsymbol{\sigma}_1 + q_2\boldsymbol{\sigma}_2 + q_3\boldsymbol{\sigma}_3\right) + q_4\mathbf{1}, \tag{5.3.28}$$

in which $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3$ are the Pauli spin matrices

$$\boldsymbol{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.3.29}$$

$$\boldsymbol{\sigma}_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{5.3.30}$$

and

$$\boldsymbol{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.3.31}$$

The set of matrices given by (5.3.25) equipped with the usual matrix algebra is isomorphic with the quaternion algebra. If we further identify each complex number with a 2-by-2 real matrix as in equation (5.1.3), then the quaternions can be represented by the set of 4-by-4 real matrices of the form

$$\begin{pmatrix} q_4 & q_3 & q_2 & q_1 \\ -q_3 & q_4 & -q_1 & q_2 \\ -q_2 & q_1 & q_4 & -q_3 \\ -q_1 & -q_2 & q_3 & q_4 \end{pmatrix} = q_1\mathbf{F}_1 + q_2\mathbf{F}_2 - q_3\mathbf{F}_3 + q_4\mathbf{1}, \tag{5.3.32}$$

where the matrices $\mathbf{F}_k$ are given by equations (2.1.67) - (2.1.69). A comparison of equation (5.3.32) with the definition (5.3.1) of the quaternion $\mathbf{q}$ shows that we can identify the matrices $\mathbf{F}_1$, $\mathbf{F}_1$ and $-\mathbf{F}_3$ respectively with the quaternion basis elements $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. The fact that $-\mathbf{F}_3$ rather $\mathbf{F}_3$ is identified with $\mathbf{k}$ can be understood by the fact that the matrices $\mathbf{F}_j$ satisfy $\mathbf{F}_1\,\mathbf{F}_2\,\mathbf{F}_3 = \mathbf{1}$, while by definition the elements $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ satisfy $\mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1$. An alternative representation is

$$\begin{pmatrix} q_4 & q_1 & -q_2 & -q_3 \\ -q_1 & q_4 & -q_3 & q_2 \\ q_2 & q_3 & q_4 & q_1 \\ q_3 & -q_2 & -q_1 & q_4 \end{pmatrix} = q_1\mathbf{B}_1 + q_2\mathbf{B}_2 - q_3\mathbf{B}_3 + q_4\mathbf{1}. \tag{5.3.33}$$

Unlike the complex product, where $\mathbf{p}\,\mathbf{q} = \mathbf{q}\,\mathbf{p}$, and vector product, where $\mathbf{p}\times\mathbf{q} = -\mathbf{q}\times\mathbf{p}$, the quaternion product $\mathbf{p}\,\mathbf{q}$ is distinctly different the product $\mathbf{q}\,\mathbf{p}$. Much as we have denoted the standard product by $(\mathbf{q},\mathbf{p})_f \equiv \mathbf{q}\,\mathbf{p}$, we will now denote the alternative product by $(\mathbf{q},\mathbf{p})_b \equiv \mathbf{p}\,\mathbf{q}$. We use the notation $(\mathbf{q},\mathbf{p})_b$ because

$$(\mathbf{q},\mathbf{p})_b = \mathbf{p}\,\mathbf{q} = \mathbf{R}(\mathbf{q})\,\mathbf{p} \tag{5.3.34}$$

and equation (5.3.18) shows that $\mathbf{R}(\mathbf{q})$ is a linear combination of the basis $\{\mathbf{B}_j\}$ with coefficients $\mathbf{q}_j$. The product $(\mathbf{q},\mathbf{p})_b$ therefore describes a rotation of $\mathbf{p}$, described by the quaternion $\mathbf{q}$, with respect to the body frame.

Put another way, it is not essential that the quaternion multiplication be required to satisfy the relation $\mathbf{i}\,\mathbf{j}\,\mathbf{k} = -1$. One can define the product $(\mathbf{q},\mathbf{p})_b$ by the relations

$$(\mathbf{i},\mathbf{i})_b = (\mathbf{j},\mathbf{j})_b = (\mathbf{k},\mathbf{k})_b = -1 \tag{5.3.35}$$

together with

$$((\mathbf{i},\mathbf{j})_b,\mathbf{k})_b = +1\ . \tag{5.3.36}$$

It then follows that

$$\begin{array}{lll} (\mathbf{i},\ \mathbf{j})_b = \ \ \,\mathbf{k}\ , & (\mathbf{j},\ \mathbf{k})_b = \ \ \,\mathbf{i}\ , & (\mathbf{k},\ \mathbf{i})_b = \ \ \,\mathbf{j}\ , \\ (\mathbf{j},\ \mathbf{i})_b = -\mathbf{k}\ , & (\mathbf{k},\ \mathbf{j})_b = -\mathbf{i}\ , & (\mathbf{i},\ \mathbf{k})_b = -\mathbf{j}\ . \end{array} \tag{5.3.37}$$

A quaternion product cannot however be defined arbitrarily, if the Law of Moduli is to be satisfied. Hamilton originally considered a multiplication rule with only two imaginary elements $\mathbf{i}$ and $\mathbf{j}$, with a multiplication rule given by $\mathbf{i}\,\mathbf{j} = \mathbf{1}$, but he could not satisfy the Law of Moduli with such a product. It would appear then that the multiplication tables (5.3.3) and (5.3.37) are the only two possibilities for multiplication rules.

Equations (5.3.32) and (5.3.33) suggest there is a close connection between the algebra of quaternions and spatial rotations. In particular the group $U \equiv (\mathbf{S}^3, \cdot)$ of unit quaternions under multiplication is isomorphic with $SU(2)$ because we can identify each element of U uniquely with an element of $SU(2)$ by equations (5.3.25) - (5.3.27), and under this mapping matrix multiplication faithfully represents quaternion multiplication. Moreover the group U is homomorphic with $SO(3)$, because under the two-to-one mapping of $\mathbf{q}$ to $\mathbf{A}(\mathbf{q})$ we have

$$\mathbf{A}(\mathbf{q}\,\mathbf{p}) = \mathbf{A}(\mathbf{q})\,\mathbf{A}(\mathbf{p}) \ . \tag{5.3.38}$$

5.4 Dimension 8: Cayley Algebra

We have remarked that it was the Law of Moduli (5.3.9) which led Hamilton to his definition of multiplication law on $\mathbf{R}^4$. Shortly afterward Cayley followed the same path to define an analogous multiplication law on $\mathbf{R}^8$, producing the *Cayley numbers*. We will follow [2] and denote the Cayley numbers by **Cay**. The Law of Moduli in this case takes the form

$$\begin{aligned}&(q_1^2+q_2^2+q_3^2+q_4^2+q_5^2+q_6^2+q_7^2+q_8^2)\\ &\qquad(p_1^2+p_2^2+p_3^2+p_4^2+p_5^2+p_6^2+p_7^2+p_8^2)\\ &=r_1^2+r_2^2+r_3^2+r_4^2+r_5^2+r_6^2+r_7^2+r_8^2,\end{aligned} \tag{5.4.1}$$

where

$$\begin{aligned}
r_1 &= q_8p_1 + q_1p_8 + q_2p_3 - q_3p_2 + q_4p_5 - q_5p_4 + q_7p_6 - q_7p_6,\\
r_2 &= q_8p_2 + q_2p_8 + q_3p_1 - q_1p_3 + q_4p_6 - q_6p_4 + q_5p_7 - q_7p_5,\\
r_3 &= q_8p_3 + q_3p_8 + q_1p_2 - q_2p_1 + q_4p_7 - q_7p_4 + q_6p_5 - q_5p_6,\\
r_4 &= q_8p_4 + q_4p_8 + q_5p_1 - q_1p_5 + q_6p_2 - q_2p_6 + q_7p_3 - q_3p_7,\\
r_5 &= q_8p_5 + q_5p_8 + q_1p_4 - q_4p_1 + q_7p_2 - q_2p_7 + q_3p_6 - q_6p_3,\\
r_6 &= q_8p_6 + q_6p_8 + q_1p_7 - q_7p_1 + q_2p_4 - q_4p_2 + q_5p_3 - q_3p_5,\\
r_7 &= q_8p_7 + q_7p_8 + q_6p_1 - q_1p_6 + q_2p_5 - q_5p_2 + q_3p_4 - q_4p_3,\\
r_8 &= q_8p_8 - q_1p_1 - q_2p_2 - q_3p_3 - q_4p_4 - q_5p_5 - q_6p_6 - q_7p_7.
\end{aligned} \tag{5.4.2}$$

6 Special Properties of Dimensions 1, 2, 4 and 8

In this section we summarize some of the significant topological and algebraic features which are connected with the one-dimensional real numbers **R**, the two-dimensional complex numbers **C**, the four-dimensional quaternions **H** and the eight-dimensional Cayley numbers **Cay** (also called octonions).

Two of the important algebraic properties of the quaternions are that multiplication is associative, is that it is a division algebra. Frobenius (1878) showed that the quaternions, along with its subalgebras, the real numbers and complex numbers, are essentially the only finite dimensional algebras with these properties.

Theorem 1 *(Frobenius) Every finite-dimensional associative division algebra is isomorphic with the algebra of the real numbers, the complex numbers or the quaternions.*

(Cf. e.g. [2], [13].)

Frobenius also proved a generalization of this theorem where the requirement that multiplication be associative is relaxed. An algebra is called *alternative* if the identities $a(bb) = (ab)b$ and $(bb)a = b(ba)$ hold. In this case we have the following result:

Theorem 2 *(Frobenius) Every alternative division algebra is isomorphic with the algebra of the real numbers, the complex numbers, the quaternions or the Cayley numbers.*

Another significant property which the quaternions share with the real numbers, the complex numbers and the Cayley numbers is the Law of Moduli (5.3.9). An algebra $\mathcal{A}$ is said to be *normed* if there exists an inner product $\langle\cdot,\cdot\rangle$ such that

$$\langle \mathbf{qp}, \mathbf{qp}\rangle = \langle \mathbf{q}, \mathbf{q}\rangle\langle \mathbf{p}, \mathbf{p}\rangle \tag{6.0.1}$$

for all $\mathbf{q}, \mathbf{p} \in \mathcal{A}$.

Two related questions arise:

1. Can we classify all normed algebras of finite dimension?

2. For which values of n can we construct bilinear forms $r_j(q_1, ..., q_n, p_1, ..., p_n)$ for $j = 1$ to n such that

$$(q_1^2 + ... + q_n^2)(p_1^2 + ... + p_n^2) = r_1^2 + ... + r_n^2. \tag{6.0.2}$$

The two questions are closely related. One can show that any set of n bilinear forms r_j satisfying (6.0.2) defines a normed algebra. Conversely, given an n-dimensional normed algebra $\mathcal{A}$ one can construct an orthonormal basis $\mathbf{i}_1, ..., \mathbf{i}_n$ and then define the bilinear forms $r_j(\mathbf{q}, \mathbf{p}) \equiv (\mathbf{qp}, \mathbf{i}_j)$, which necessarily satisfy (6.0.2).

In 1898 Hurwitz proved

Theorem 3 *(Hurwitz) Every normed algebra with identity is isomorphic to one of the algebras* **R**, **C**, **H** *or* **Cay**.

The requirement that the algebra have an identity element is essential in Hurwitz' theorem. However, given a normed algebra $\mathcal{A}$ over a vector space $\mathbf{V}$ one can construct a new normed algebra over $\mathbf{V}$ by defining the product

$$\mathbf{q} \odot \mathbf{p} \equiv A(\mathbf{q})\, B(\mathbf{p}), \tag{6.0.3}$$

where A and B are orthogonal transformations on $\mathcal{A}$. One can show (cf. [13], p.**132**) that *every* normed algebra over the vector space $\mathbf{V}$ can be constructed in this manner. In particular, from a normed algebra $\mathcal{A}$ of dimension n one can always construct a normed algebra $\mathcal{A}_0$ with identity and having the same dimension n. This fact together with Hurwitz' Theorem answers question 1 above. Moreover, by our earlier observation about the connection between questions 1 and 2 we see that the answer to question is $n = 1$, 2, 4 or 8.

Related to the Law of Moduli is Lagrange's Theorem [8], which states that any natural number can be expressed as a sum of four squares (of natural numbers). The proof of this theorem depends upon the fact that the set of all sums of four squares is closed under multiplication, which follows from the Law of Moduli for quaternions. To complete the proof it is necessary to show that every prime can be expressed as a sum of four squares. One can see that four squares are necessary in general because $(2n)^2 = 0 \pmod 4$ and $(2n+1)^2 = 1 \pmod 8$, so one cannot construct a number of the form $8n+7$ as a sum of only three squares. In fact, Legendre and Gauss both showed that sums of three squares generate all numbers *except* those of the form $4^m(8n+7)$. One can also show that sums of two squares generate all numbers of the for $n_1^2 n_2$, where has n_2 has *no* prime factors of the form $4n+3$.

Finally, Milnor and Kervaire showed that there is an important geometric property of the spheres $\mathbf{S}^{n-1} \in \mathbf{R}^n$ for $n = 2$, 4 and 8. The Hedgehog Theorem states that for each even-dimensional sphere $\mathbf{S}^{2k}$ there is no smooth nonzero vector field. Using a result of Bott, Milnor and Kervaire were able to show that $\mathbf{S}^k$ is parallelizable if and only if $k = 1$, 3 or 7. ($\mathbf{S}^k$ is parallelizable if there exist k linearly independent vector fields.) When we consider the circle $\mathbf{S}^1$ in $\mathbf{C}$, it is easy to see that we can simply select the vector field defined by iz at each $z \in \mathbf{S}^1$. Analogously, we can consider $\mathbf{S}^3$ in $\mathbf{H}$, and then we can define three linearly independent vector fields $\mathbf{iq}$, $\mathbf{jq}$ and $\mathbf{kq}$ at each point $\mathbf{q} \in \mathbf{S}^3$. The ability to explicitly write down an orthonormal basis for the tangent plane is a powerful tool, and is extremely useful in stability analysis, where perturbations must lie in the tangent plane. One can similarly generate 7 linearly independent vector fields on $\mathbf{S}^7$ in **Cay** by premultiplying by the standard seven "complex" basis elements.

References

[1] Altmann, S.L. **Rotations, Quaternions and Double Groups** (Clarendon, 1986).

[2] Artmann, B. **The Concept of Number: from quaternions to monads to topological fields** (Halsted Press, 1988).

Chapter 5 contains a valuable discussion, particularly of the nnnmatrix representation of quaternions.

[3] Crowe, M.J. **A History of Vector Analysis** (Notre Dame, 1964).

Traces the early history of quaternions and their influence on the development of modern vector analysis.

[4] DuVal, P. **Homographies, Quaternions and Rotations** (Oxford, 1964).

A homography is a Möbius transformation of the form $t \to (at + b)/(ct + d)$, where $d = \bar{a}$ and $c = -\bar{b}$ and $|a|^2 + |b|^2 = 1$. This monograph draws the connections between homographies, unit quaternions and rotations.

[5] Goldstein, H. **Classical Mechanics**, Second Edition (Addison-Wesley, 1981).

Refers to quaternions as "musty mathematics", but describes the associated structure of Cayley-Klein parameters.

[6] Gray, J.J. Olinde Rodrigues' Paper of 1840 on Transformation Groups. *Archive for the History of the Exact Sciences* **21** (1980) 376-385.

[7] Hamilton, W.R. **Elements of Quaternions** (Chelsea, 1969).

[8] Hardy, G.H. & E.M. Wright. **An Introduction to the Theory of Numbers** (Oxford, 1954).

Treats Lagrange's (or the Four Squares) Theorem in the context of quaternions.

[9] Haug, E.J. **Computer-Aided Kinematics and Dynamics of Mechanical Systems VolumeI: Basic Methods** (Allyn and Bacon, 1989).

Summarizes kinematics and dynamics in Euler parameters, discusses enforcement of the associated constraint.

[10] Hughes, P. **Spacecraft Attitude Dynamics** (Wiley, 1986).

[11] Hungerford, T.W. **Algebra** (Springer-Verlag, New York, 1974).

[12] Junkins, J. & J. Turner. **Optimal Spacecraft Rotational Maneuvers** (Elsevier, 1986).

Discusses relationships between quaternions and Euler angles, and provides examples where quaternions used in modelling spacecraft dynamics.

[13] Kantor, I.L. & A.S. Solodovnikov. **Hypercomplex Numbers: An Elementary Introduction to Algebras** (Springer-Verlag, 1989).

[14] Klein, F. & A. Sommerfeld.

[15] Kyrala, A. **Theoretical Physics: Applications of Vectors, Matrics, Tensors and Quaternions** (W.B. Saunders, 1967).

Applies quaternions to formulate Maxwell's equations and quantum mechanics.

[16] Milnor, J. Some Consequences of a Theorem of Bott. *Annals of Mathematics* **68** (1958), 444-449.

Proves that the sphere $\mathbf{S}^k$ is parallelizable if and only if $k = 1$, 3 or 7. Proof given independently by Kervaire in Proc. Nat. Acad. Sci. U.S.A. **44** (1958), 280-283.

[17] Morton, H.S. A Formulation of Rigid-Body Rotational Dynamics Based on Generalized Angular Momentum Variables Corresponding to Euler Parameters. AIAA/Astrodynamics Conference 1984. AIAA-84-2023.

[18] Morton, H.S., Jr. Hamiltonian and Lagrangian Formulations of Rigid-Body Rotational Dynamics Based on the Euler Parameters. *J. Astronautical Sciences* **41** (1994) 569-592.

A Hamiltonian formulation, precursor to impetus-striction.

[19] Penrose, R. & W. Rindler. **Spinors and Space-Time** (Cambridge, 1984).

Here spinors are related to Cayley-Klein parameters and quaternions, and used to represent relativistic fields.

[20] Seywald, H., R.R. Kumar, S.M. Deshpande & M.L. Heck.

[21] Shuster, M.D. A Survey of Attitude Representations. *J. Astronautical Sciences* **41** (1994) 439-518.

[22] Shuster, M.D. & G.A. Natanson. Quaternion Computation from a Geometric Point of View. *J. Astronautical Sciences* **41** (1994) 545-556.

Typical algorithms to obtain attitude representations from Euler parameters suffer singularities when q_4 is near zero. However the unit norm constraint implies that at least one of the four parameters is greater than or equal to 1/4 is magnitude. This paper presents a consistent method by means of successive rotations to minimize the risk of round-off error.

[23] van der Waerden, B.L. **Group Theory and Quantum Mechanics** (Springer-Verlag, 1974).

[24] van der Waerden, B.L. Hamilton's Discovery of Quaternions. *Mathematics Magazine* **49** (1976), 227-234.

[25] Wertz, J. **Spacecraft Attitude Determination and Control** (D. Reidel, 1985).

Summarizes fundamentals of the algebra of quaternions and their applications to rotational dynamics.