

## Chapter 2

# Matrices, Vectors, Frames, Transforms

This chapter begins with an overview of matrices and vectors, which are used extensively in attitude analysis. We assume that the reader has some familiarity with this material, so the account is not completely self-contained. The principal objective of this section is to define our notation and conventions.

We next discuss a special category of four-component vectors, which we refer to as quaternions although they differ conceptionally from the quaternions introduced by W. R. Hamilton in 1844. They perform the same function as Hamilton's quaternions in applications, however, and have proved to be extremely useful in attitude analysis.

We then move on to a discussion of rotations in three-dimensional space and the most common parameter sets that have been used to specify these rotations: the Euler axis and angle, the rotation vector, the quaternion, the Rodrigues parameters, the modified Rodrigues parameters, and the Euler angles. The last section in this chapter addresses the representation of attitude errors. A more extensive treatment of the material in this section, including historical references, can be found in Shuster's comprehensive review article [17].

## 2.1 Matrices

An  $m \times n$  matrix  $A$  is an array with  $m$  rows and  $n$  columns of *scalars*:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (2.1)$$

We will assume that the scalar matrix elements are real numbers. The results in this chapter can be generalized to matrices with complex elements, but we will rarely need to deal with complex matrices. If  $m = n$ , then the matrix  $A$  is *square*. We denote an  $m \times n$  matrix with all components equal to zero by  $0_{m \times n}$ , or sometimes simply by  $0$  if confusion is unlikely to result.

A *column vector*, or sometimes simply a vector, is an  $n \times 1$  matrix

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (2.2)$$

We denote an  $n$ -component vector with all components equal to zero by  $\mathbf{0}_n$ , or sometimes simply by  $\mathbf{0}$  if confusion is unlikely.

Matrices can be added, subtracted, or multiplied. For addition and subtraction, all matrices must have the same number of rows and columns. The elements of

$$C = A \pm B \quad (2.3)$$

are given by  $c_{ij} = a_{ij} \pm b_{ij}$ . Matrix addition and subtraction are both commutative,  $A \pm B = B \pm A$ , and associative,  $(A \pm B) \pm C = A \pm (B \pm C)$ . The multiplication of two matrices  $A$  and  $B$ :

$$C = A B \quad (2.4)$$

is valid only when the number of columns of  $A$  is equal to the number of rows of  $B$  (i.e.  $A$  and  $B$  must be *conformable*). The resulting matrix  $C$  will have rows equal to the number of rows of  $A$  and columns equal to the number of columns of  $B$ . Thus, if  $A$  has dimension  $m \times n$  and  $B$  has dimension  $n \times p$ , then  $C$  will have dimension  $m \times p$ . The elements of  $C$  are given by

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad (2.5)$$

for all  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, p$ . Matrix multiplication is associative,  $A(BC) = (AB)C$ , and distributive,  $A(B + C) = AB + AC$ , but not commutative in general,  $AB \neq BA$ . In those cases for which  $AB = BA$ , the matrices  $A$  and  $B$  are said to *commute*.

The transpose of a matrix, denoted  $A^T$ , has rows that are the columns of  $A$  and columns that are the rows of  $A$ . The transpose of the matrix defined by Eq. (2.1), for example, is

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \quad (2.6)$$

The transpose of a column vector is a *row vector*. The transpose operator has the following properties:

$$(\alpha A)^T = \alpha A^T, \text{ where } \alpha \text{ is a scalar} \quad (2.7a)$$

$$(A + B)^T = A^T + B^T \quad (2.7b)$$

$$(A B)^T = B^T A^T \quad (2.7c)$$

If  $A = A^T$ , then  $A$  is a *symmetric* matrix, if  $A = -A^T$ , then  $A$  is a *skew symmetric* matrix.

A *diagonal* matrix is a square matrix with nonzero elements only on the main diagonal and all other elements equal to zero. An  $n \times n$  diagonal matrix can be formed from an  $n$ -component row or column vector by

$$\text{diag}(\mathbf{x}) = \text{diag}(\mathbf{x}^T) = \text{diag}([x_1 \ x_2 \ \cdots \ x_n]) \equiv \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix} \quad (2.8)$$

An important special case of a diagonal matrix is the *identity* matrix:

$$I \equiv \text{diag}([1 \ 1 \ \cdots \ 1]) \quad (2.9)$$

It has the property that  $I A = A$  and  $B I = B$  if the matrices are conformable. We sometimes denote the  $n \times n$  identity matrix by  $I_n$  if supplying the subscript helps to remove ambiguity.

An *upper triangular* matrix is a matrix in which all the entries below the main diagonal are zero, i.e.  $a_{ij} = 0$  for  $i < j$ . A *lower triangular* matrix has all zeros above the main diagonal, i.e.  $a_{ij} = 0$  for  $i > j$ .

Two useful scalar quantities can be defined for square matrices, the *trace* and the *determinant*. The trace of an  $n \times n$  matrix is simply the sum of the diagonal elements:

$$\text{tr} A = \sum_{i=1}^n a_{ii} \quad (2.10)$$

Some useful identities involving the matrix trace are given by

$$\text{tr}(\alpha A) = \alpha \text{tr} A \quad (2.11a)$$

$$\text{tr}(A + B) = \text{tr} A + \text{tr} B \quad (2.11b)$$

$$\text{tr}(A B C) = \text{tr}(B C A) = \text{tr}(C A B) \quad (2.11c)$$

$$\text{tr}(A B) = \text{tr}(B A) \quad (2.11d)$$

$$\text{tr}(\mathbf{x} \mathbf{y}^T) = \mathbf{x}^T \mathbf{y} \quad (2.11e)$$

$$\text{tr}(A \mathbf{y} \mathbf{x}^T) = \mathbf{x}^T A \mathbf{y} \quad (2.11f)$$

Equation (2.11d)–(2.11f) are special cases of Eq. (2.11c), which shows the cyclic invariance of the trace. The operation  $\mathbf{y} \mathbf{x}^T$  is known as the *outer product* (note that  $\mathbf{y} \mathbf{x}^T \neq \mathbf{x} \mathbf{y}^T$  in general).

The determinant of an  $n \times n$  matrix can be computed using an expansion about any row  $i$  or any column  $j$ :

$$\det A = \sum_{k=1}^n (-1)^{i+k} a_{ik} m_{ik} = \sum_{k=1}^n (-1)^{k+j} a_{kj} m_{kj} \quad (2.12)$$

where  $m_{ij}$  is the *minor*, which is the determinant of the  $(n-1) \times (n-1)$  matrix resulting from deleting row  $i$  and column  $j$  of  $A$ . Some useful determinant identities are

$$\det I = 1 \quad (2.13a)$$

$$\det A^T = \det A \quad (2.13b)$$

$$\det(A B) = \det A \det B \quad (2.13c)$$

$$\det(\alpha A) = \alpha^n \det A \quad (2.13d)$$

The elements of the *adjoint* matrix  $\text{adj } A$  are defined in terms of the minors:

$$[\text{adj } A]_{ij} = (-1)^{i+j} m_{ji} \quad (2.14)$$

Note the reversed order of the subscripts on the left and right side of this equation. It can be shown that [21]

$$A(\text{adj } A) = (\text{adj } A)A = (\det A)I \quad (2.15)$$

Thus we see that a *nonsingular* matrix, which is a square matrix with a nonzero determinant, has an inverse

$$A^{-1} = \frac{\text{adj } A}{\det A} \quad (2.16)$$

that satisfies the identities

$$A^{-1}A = A A^{-1} = I \quad (2.17a)$$

$$(A^T)^{-1} = (A^{-1})^T \equiv A^{-T} \quad (2.17b)$$

If  $A$  and  $B$  are  $n \times n$  matrices, then the matrix product  $AB$  is nonsingular if and only if  $A$  and  $B$  are nonsingular. If these conditions are met, then

$$(AB)^{-1} = B^{-1}A^{-1} \quad (2.18)$$

## 2.2 Vectors

The *dot product*, *inner product*, or *scalar product* of two vectors of equal dimension,  $n \times 1$ , is given by

$$\mathbf{x} \cdot \mathbf{y} \equiv \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} = \sum_{i=1}^n x_i y_i \quad (2.19)$$

If the dot product is zero, then the vectors are said to be *orthogonal*. A measure of the length of a vector is given by its *Euclidean norm*, which is the square root of the inner product of the vector with itself:

$$\|\mathbf{x}\| \equiv \sqrt{\mathbf{x} \cdot \mathbf{x}} = \left[ \sum_{i=1}^n x_i^2 \right]^{1/2} \quad (2.20)$$

It is easily seen that  $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ . The norm obeys the inequality

$$\|\mathbf{x}\| \geq 0 \quad (2.21)$$

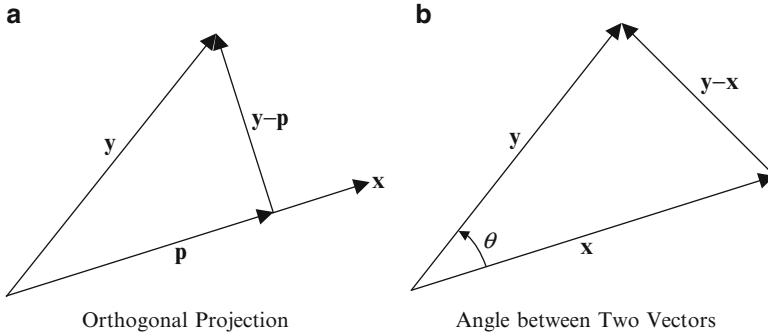
with equality only for  $\mathbf{x} = \mathbf{0}$ .

A vector with norm equal to unity is said to be a *unit vector*. Any nonzero vector can be made into a unit vector by dividing it by its norm:

$$\mathbf{x}_{\text{unit}} \equiv \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (2.22)$$

This is referred to as *normalizing* the vector  $\mathbf{x}$ .

Figure 2.1a shows two vectors,  $\mathbf{x}$  and  $\mathbf{y}$ , and the *orthogonal projection* of a vector  $\mathbf{y}$  onto a vector  $\mathbf{x}$ . The orthogonal projection of  $\mathbf{y}$  onto  $\mathbf{x} \neq \mathbf{0}$  is given by



**Fig. 2.1** Depiction of an orthogonal projection (a) and of the angle between two vectors (b)

$$\mathbf{p} = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2} \mathbf{x} \quad (2.23)$$

This projection yields  $(\mathbf{y} - \mathbf{p}) \cdot \mathbf{x} = 0$ . From Eqs. (2.21) and (2.23) we see that

$$0 \leq \|(\mathbf{y} - \mathbf{p})\|^2 = \|\mathbf{y}\|^2 - \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{x}\|^2} \quad (2.24)$$

This yields the *Cauchy-Schwarz inequality*:

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad (2.25)$$

The Cauchy-Schwarz inequality implies the *triangle inequality*

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (2.26)$$

and it allows us to define the angle  $\theta$  between the vectors  $\mathbf{x}$  and  $\mathbf{y}$  by

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \quad (2.27)$$

This angle is illustrated in Fig. 2.1b, as is the vector difference  $\mathbf{y} - \mathbf{x}$ .

## 2.3 Jacobian, Gradient, and Hessian

In this section we will introduce notation for partial derivatives with respect to components of a vector  $\mathbf{x}$ . If  $\mathbf{y}(\mathbf{x})$  is an  $m$ -component vector function of an  $n$ -component vector  $\mathbf{x}$ , we define the *Jacobian matrix* as the  $m \times n$  matrix

$$\frac{\partial \mathbf{y}(\mathbf{x})}{\partial \mathbf{x}} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \quad (2.28)$$

In particular, if we have a scalar  $f(\mathbf{x})$  in place of the vector function  $\mathbf{y}(\mathbf{x})$ , this reduces to the  $m \times 1$  row vector

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \equiv \left[ \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right] \quad (2.29)$$

The transpose of this, a  $1 \times m$  column vector, is known as the *gradient* of  $f(\mathbf{x})$ :

$$\nabla_{\mathbf{x}} f(\mathbf{x}) \equiv \left[ \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right]^T \equiv \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^T} \quad (2.30)$$

The *Hessian* is the  $n \times n$  symmetric matrix of second-order partial derivatives of  $f(\mathbf{x})$ :

$$\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T} \equiv \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix} \quad (2.31)$$

Some use the term *Hessian* to refer to the determinant of the Hessian matrix. The *Laplacian* is the trace of the Hessian matrix:

$$\nabla_{\mathbf{x}}^2 f(\mathbf{x}) \equiv \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = \text{tr} \left( \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T} \right) \quad (2.32)$$

## 2.4 Orthonormal Bases, Change of Basis

An *orthonormal* set of vectors is a set with inner products obeying

$$\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij} \quad (2.33)$$

where the Kronecker delta  $\delta_{ij}$  is defined as

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (2.34)$$

If the number of vectors in the set is equal to the dimension of the vectors, then the set is said to constitute an *orthonormal basis*. We will generally denote basis vectors (and some other unit vectors) by the letter  $\mathbf{e}$ . The orthonormal basis

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (2.35)$$

is called the *natural basis*. Any vector  $\mathbf{x}$  can be written as a linear superposition of basis vectors; using the natural basis gives

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^n x_j \mathbf{e}_j \quad (2.36)$$

An orthonormal set of basis vectors defines a *reference frame*. Let us refer to the frame defined by  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  as frame  $F$ ; and consider another frame  $F'$  defined by the orthonormal basis:  $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$ . We can express the vector  $\mathbf{x}$  in either basis

$$\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j = \sum_{k=1}^n x'_k \mathbf{e}'_k \quad (2.37)$$

Taking the dot product of this equation with  $\mathbf{e}_j$  or  $\mathbf{e}'_k$  gives

$$x_j = \mathbf{e}_j \cdot \mathbf{x} = \sum_{k=1}^n (\mathbf{e}_j \cdot \mathbf{e}'_k) x'_k \quad (2.38a)$$

$$x'_k = \mathbf{e}'_k \cdot \mathbf{x} = \sum_{j=1}^n (\mathbf{e}'_k \cdot \mathbf{e}_j) x_j \quad (2.38b)$$



Substituting the first equality of Eq. (2.38a) into the first equality of Eq. (2.37) gives

$$\mathbf{x} = \sum_{j=1}^n \mathbf{e}_j (\mathbf{e}_j^T \mathbf{x}) = \left( \sum_{j=1}^n \mathbf{e}_j \mathbf{e}_j^T \right) \mathbf{x} \quad (2.39)$$

It follows that the identity matrix can be expressed in terms of any orthonormal basis by

$$I_n = \sum_{i=1}^n \mathbf{e}_i \mathbf{e}_i^T \quad (2.40)$$

When considering transformations among different reference frames, it is very useful (almost indispensable, in fact) to regard  $\mathbf{x}$  as an *abstract* vector having an existence in  $n$ -dimensional space independent of any particular reference frame, and having *representations*  $\mathbf{x}_F$  in reference frame  $F$  and  $\mathbf{x}_{F'}$  in frame  $F'$ . Subscripts will be provided when it is important to indicate the reference frame explicitly or to carefully distinguish between abstract vectors and their representations, but subscripts will often be omitted when confusion is unlikely to arise. Arraying the components of  $\mathbf{x}$  in the reference frames  $F$  and  $F'$  as column vectors gives the representations

$$\mathbf{x}_F = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{x} \\ \mathbf{e}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{e}_n \cdot \mathbf{x} \end{bmatrix} \quad \text{and} \quad \mathbf{x}_{F'} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}'_1 \cdot \mathbf{x} \\ \mathbf{e}'_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{e}'_n \cdot \mathbf{x} \end{bmatrix} \quad (2.41)$$

Equation (2.38) can be used to express the relations between  $\mathbf{x}_F$  and  $\mathbf{x}_{F'}$  as matrix products:

$$\mathbf{x}_F = D_{FF'} \mathbf{x}_{F'} \quad (2.42a)$$

$$\mathbf{x}_{F'} = D_{F'F} \mathbf{x}_F \quad (2.42b)$$

where

$$D_{FF'} = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{e}'_1 & \mathbf{e}_1 \cdot \mathbf{e}'_2 & \cdots & \mathbf{e}_1 \cdot \mathbf{e}'_n \\ \mathbf{e}_2 \cdot \mathbf{e}'_1 & \mathbf{e}_2 \cdot \mathbf{e}'_2 & \cdots & \mathbf{e}_2 \cdot \mathbf{e}'_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}_n \cdot \mathbf{e}'_1 & \mathbf{e}_n \cdot \mathbf{e}'_2 & \cdots & \mathbf{e}_n \cdot \mathbf{e}'_n \end{bmatrix} \quad (2.43a)$$

$$D_{F'F} = \begin{bmatrix} \mathbf{e}'_1 \cdot \mathbf{e}_1 & \mathbf{e}'_1 \cdot \mathbf{e}_2 & \cdots & \mathbf{e}'_1 \cdot \mathbf{e}_n \\ \mathbf{e}'_2 \cdot \mathbf{e}_1 & \mathbf{e}'_2 \cdot \mathbf{e}_2 & \cdots & \mathbf{e}'_2 \cdot \mathbf{e}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}'_n \cdot \mathbf{e}_1 & \mathbf{e}'_n \cdot \mathbf{e}_2 & \cdots & \mathbf{e}'_n \cdot \mathbf{e}_n \end{bmatrix} \quad (2.43b)$$

The matrices  $D_{FF'}$  and  $D_{F'F}$  are known as *direction cosine matrices* (DCMs) because their elements are the cosines of the angles between the basis vectors in the two reference frames. These transformation equations hold for any two orthonormal bases. We see immediately that

$$D_{F'F} = D_{FF'}^T \quad (2.44)$$

which means that the matrix transforming vector representations from frame  $F$  to frame  $F'$  is the transpose of the matrix transforming from  $F'$  to  $F$ . Another way of stating the content of Eq. (2.43) is that the columns of  $D_{FF'}$  are the representations in frame  $F$  of the basis vectors of frame  $F'$  and *vice versa*:

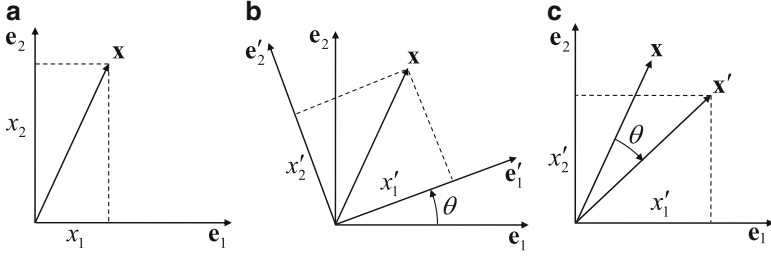
$$D_{FF'} = [\mathbf{e}'_{1F} \ \mathbf{e}'_{2F} \ \cdots \ \mathbf{e}'_{nF}] \quad (2.45a)$$

$$D_{F'F} = [\mathbf{e}_{1F'} \ \mathbf{e}_{2F'} \ \cdots \ \mathbf{e}_{nF'}] \quad (2.45b)$$

We emphasize that the transformations considered here are transformations of the representations of a fixed abstract vector resulting from a change in the reference frame. This is the *passive* interpretation of a transformation, also known as the *alias* sense (from the Latin word for “otherwise,” in the sense of “otherwise known as”) [17]. The alternative *active* interpretation (also known as the *alibi* sense from the Latin word for “elsewhere”) considers the representation in a fixed reference frame of an abstract vector that is rotated from  $\mathbf{x}$  to  $\mathbf{x}'$ . The difference between these two interpretations is illustrated in Fig. 2.2a–c. Figure 2.2a shows the components  $x_1$  and  $x_2$  of the vector  $\mathbf{x}$  in frame  $F$ . Figure 2.2b shows the components  $x'_1$  and  $x'_2$  of  $\mathbf{x}$  in the rotated frame  $F'$ , illustrating the alias sense. Figure 2.2c shows the components  $x'_1$  and  $x'_2$  of a rotated vector  $\mathbf{x}'$  in the unrotated frame  $F$ , illustrating the alibi sense. In this example, the values of  $x'_1$  and  $x'_2$  are the same in the alias and alibi interpretations, but the significance of these quantities is completely different in the two cases. They are the components of an unrotated vector in a rotated reference frame in the alias interpretation, while they are the components of a rotated vector in an unrotated reference frame in the alibi interpretation. When necessary, we can use a more precise notation to distinguish these two different interpretations, as in Eq. (2.45); but we can usually avoid this complication because this book will rarely, if ever, employ the active interpretation of transformations. In the example illustrated in Fig. 2.2b,c, it can be seen that the magnitude of the rotation is the same in the two interpretations but the sense of the rotation is opposite. It is important to remember this important difference in the sense of rotation in the two interpretations; it results in some unexpected minus signs, and overlooking them has led to actual errors in flight software.<sup>1</sup>

---

<sup>1</sup>One example is an incorrect sign for the velocity aberration correction for star tracker measurements on the WMAP spacecraft, which fortunately was easily corrected.



**Fig. 2.2** Alias and Alibi interpretations of a transformation. (a) Components. (b) Alias. (c) Alibi

Now consider that we have three reference frames, denoted  $F$ ,  $G$ , and  $H$ , and that we transform a vector representation from frame  $F$  to frame  $G$  and then from frame  $G$  to frame  $H$ . This is effected by the successive transformations

$$\mathbf{x}_H = D_{HG}\mathbf{x}_G = D_{HG}(D_{GF}\mathbf{x}_F) = (D_{HG}D_{GF})\mathbf{x}_F \quad (2.46)$$

We could have transformed directly from  $F$  to  $H$  by  $\mathbf{x}_H = D_{HF}\mathbf{x}_F$ . These transformations must be equivalent for any vector  $\mathbf{x}_F$ , so it must be true that

$$D_{HF} = D_{HG}D_{GF} \quad (2.47)$$

This says that successive transformations are accomplished by simple matrix multiplication of DCMs, which may appear to be an obvious result. It is not inconceivable, though, that the method for implementing successive transformations could have been more complex.

Transforming from frame  $F$  to frame  $G$  and back to frame  $F$  is effected by the matrix  $D_{FF}$ , which must be the identity matrix. But from Eqs. (2.44) and (2.47) this means that

$$I = D_{FG}D_{GF} = D_{FG}D_{FG}^T = D_{GF}^T D_{GF} \quad (2.48)$$

Matrices like DCMs for which  $I = D D^T = D^T D$  are called *orthogonal* matrices, or sometimes *orthonormal* matrices. The transpose of an orthogonal matrix is equal to its inverse; its columns constitute a set of orthonormal vectors, as do its rows. In fact, the columns and rows of a DCM are just the representations in one reference frame of the basis vectors in the other reference frame. Equation (2.7c) can be used to show that the product of two orthogonal matrices is orthogonal; i.e. if  $A$  and  $B$  are orthogonal, then

$$(A B)(A B)^T = A B B^T A^T = A A^T = I \quad (2.49)$$

This means that the set of  $n \times n$  orthogonal matrices form a *group*, which requires, among other things, that the product of two elements of the group is also an element.

The group of  $n \times n$  orthogonal matrices is called the *orthogonal group*  $O(n)$ . Because  $1 = \det I = \det(D^T D) = (\det D)^2$ , the determinant of an orthogonal matrix must be equal to  $\pm 1$ . It follows that the set of  $n \times n$  *proper* orthogonal matrices, which are those whose determinant is  $+1$ , also form a group, called the *special orthogonal group*  $SO(n)$ . The orthogonal matrices with determinant  $-1$  do not form a group, because the product of two matrices with determinant  $-1$  has determinant  $+1$ .

An important result follows from Eq. (2.48), namely that

$$\mathbf{x}_G \cdot \mathbf{y}_G = (D_{GF} \mathbf{x}_F)^T D_{GF} \mathbf{y}_F = \mathbf{x}_F^T D_{GF}^T D_{GF} \mathbf{y}_F = \mathbf{x}_F \cdot \mathbf{y}_F \quad (2.50)$$

This says that the value of the inner product of two vectors is independent of the reference frame in which they are represented, or equivalently that reference frame transformations preserve both lengths of vectors and angles between them. Inserting a matrix between two vectors leads to the relation

$$\mathbf{x}_F^T M_F \mathbf{y}_F = \mathbf{x}_F^T D_{GF}^T D_{GF} M_F D_{GF}^T D_{GF} \mathbf{y}_F = \mathbf{x}_G^T M_G \mathbf{y}_G \quad (2.51)$$

where

$$M_G \equiv D_{GF} M_F D_{GF}^T \quad (2.52)$$

This defines how matrices must transform under reference frame transformations for Eq. (2.51) to hold.

## 2.5 Vectors in Three Dimensions

The case of three dimensions is especially interesting because we and our vehicles live in three-dimensional space.<sup>2</sup> In three dimensions, the abstract vectors generally represent physical quantities that have both a magnitude and a direction, like displacements or velocities.

We can define a *vector product* or *cross product* for three-component vectors in terms of their components by

$$\mathbf{x} \times \mathbf{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = -\mathbf{y} \times \mathbf{x} \quad (2.53)$$

It is easily seen that the cross product  $\mathbf{x} \times \mathbf{y}$  is perpendicular to both  $\mathbf{x}$  and  $\mathbf{y}$ . The cross product can also be obtained using matrix multiplication:

---

<sup>2</sup>This is true in classical physics. Various contemporary physical theories indicate that we live in a space having anywhere from two to eleven dimensions.

$$\mathbf{x} \times \mathbf{y} = [\mathbf{x} \times] \mathbf{y} \quad (2.54)$$

where  $[\mathbf{x} \times]$  is the *cross product matrix*, defined by

$$[\mathbf{x} \times] \equiv \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad (2.55)$$

Note that  $[\mathbf{x} \times]$  is a skew symmetric matrix.<sup>3</sup>

The cross product and the cross product matrix obey the following relations:

$$\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{z} \quad (2.56a)$$

$$[\mathbf{x} \times] [\mathbf{y} \times] = -(\mathbf{x} \cdot \mathbf{y}) \mathbf{I} + \mathbf{y} \mathbf{x}^T \quad (2.56b)$$

$$[\mathbf{x} \times] [\mathbf{y} \times] - [\mathbf{y} \times] [\mathbf{x} \times] = \mathbf{y} \mathbf{x}^T - \mathbf{x} \mathbf{y}^T = [(\mathbf{x} \times \mathbf{y}) \times] \quad (2.56c)$$

$$\text{adj} [\mathbf{x} \times] = \mathbf{x} \mathbf{x}^T \quad (2.56d)$$

It follows from Eq. (2.56b) that

$$\|\mathbf{x} \times \mathbf{y}\|^2 = ([\mathbf{x} \times] \mathbf{y})^T ([\mathbf{x} \times] \mathbf{y}) = -\mathbf{y}^T [\mathbf{x} \times]^2 \mathbf{y} = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2 \quad (2.57)$$

With Eq. (2.27), this means that

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta \quad (2.58)$$

It is often convenient to express a  $3 \times 3$  matrix in terms of its columns:

$$\mathbf{M} \equiv [\mathbf{a} \ \mathbf{b} \ \mathbf{c}] \quad (2.59)$$

With this notation, the determinant is

$$\det \mathbf{M} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \quad (2.60)$$

and the adjoint is

$$\text{adj} \mathbf{M} \equiv \text{adj} ([\mathbf{a} \ \mathbf{b} \ \mathbf{c}]) = \begin{bmatrix} (\mathbf{b} \times \mathbf{c})^T \\ (\mathbf{c} \times \mathbf{a})^T \\ (\mathbf{a} \times \mathbf{b})^T \end{bmatrix} \quad (2.61)$$

We can also derive the useful identity

---

<sup>3</sup>A vector product of two vectors can be defined only in three dimensions because an  $n \times n$  skew-symmetric matrix has exactly  $n$  independent parameters only for  $n = 3$ .

$$\begin{aligned}
M^T [\mathbf{x} \times] M &= \begin{bmatrix} \mathbf{a} \cdot (\mathbf{x} \times \mathbf{a}) & \mathbf{a} \cdot (\mathbf{x} \times \mathbf{b}) & \mathbf{a} \cdot (\mathbf{x} \times \mathbf{c}) \\ \mathbf{b} \cdot (\mathbf{x} \times \mathbf{a}) & \mathbf{b} \cdot (\mathbf{x} \times \mathbf{b}) & \mathbf{b} \cdot (\mathbf{x} \times \mathbf{c}) \\ \mathbf{c} \cdot (\mathbf{x} \times \mathbf{a}) & \mathbf{c} \cdot (\mathbf{x} \times \mathbf{b}) & \mathbf{c} \cdot (\mathbf{x} \times \mathbf{c}) \end{bmatrix} \\
&= \begin{bmatrix} 0 & -(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{x} & (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{x} \\ (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{x} & 0 & -(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{x} \\ -(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{x} & (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{x} & 0 \end{bmatrix} = [ \{ (\text{adj } M) \mathbf{x} \} \times ]
\end{aligned} \tag{2.62}$$

Setting  $M = A^T$ , where  $A$  is a proper orthogonal  $3 \times 3$  matrix, gives  $\text{adj } M = A$  and

$$A [\mathbf{x} \times] A^T = [(A\mathbf{x}) \times], \quad \text{for } A \in \text{SO}(3) \tag{2.63}$$

This special case is much more useful than the general case. In the specific case of a reference frame transformation, we have

$$A_{GF} [\mathbf{x}_F \times] A_{GF}^T = [(A_{GF} \mathbf{x}_F) \times] = [\mathbf{x}_G \times] \tag{2.64}$$

which can be viewed as a special case of Eq. (2.52). This equation can be used to show that

$$\mathbf{x}_G \times \mathbf{y}_G = [(A_{GF} \mathbf{x}_F) \times] A_{GF} \mathbf{y}_F = A_{GF} (\mathbf{x}_F \times \mathbf{y}_F) \tag{2.65}$$

The significance of this is that the cross product of two vectors transforms exactly like any other vector under a reference frame rotation, which is what we want.

The discussion so far has been purely algebraic; it has said nothing about the *right hand rule*. Discussing this rule requires an intuitive picture of vectors in three-dimensional space. First note that the definition of the cross product means that the natural basis vectors defined by Eq. (2.35) satisfy the relation  $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$ . Now consider the possible orientation of these three basis vectors in physical space. The orientation of  $\mathbf{e}_1$  and  $\mathbf{e}_2$  is arbitrary, except that they must be orthogonal; but this leaves us only two choices for  $\mathbf{e}_3$ , which must be perpendicular to both  $\mathbf{e}_1$  and  $\mathbf{e}_2$ . We choose the orientation of  $\mathbf{e}_3$  to satisfy the right hand rule, i.e.: we place  $\mathbf{e}_1$  and  $\mathbf{e}_2$  tail-to-tail, flatten the right hand, extending it in the direction of  $\mathbf{e}_1$ , curl the fingers toward  $\mathbf{e}_2$  through the shortest angle, and choose  $\mathbf{e}_3$  to point along the direction indicated by the thumb. This defines a right-handed reference frame, and all cross products will obey the right hand rule if we restrict ourselves to right handed reference frames. To see this explicitly, consider two arbitrary vectors  $\mathbf{x}$  and  $\mathbf{y}$ . There is a reference frame in which

$$\mathbf{x}_F = \|\mathbf{x}\| \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{y}_F = \|\mathbf{y}\| \begin{bmatrix} \cos \theta \\ \sin \theta \\ 0 \end{bmatrix} \implies \mathbf{x}_F \times \mathbf{y}_F = \|\mathbf{x}\| \|\mathbf{y}\| \begin{bmatrix} 0 \\ 0 \\ \sin \theta \end{bmatrix} \tag{2.66}$$

where  $\theta$  is the angle between  $\mathbf{x}$  and  $\mathbf{y}$ . This cross product obeys the right-hand rule, and Eq. (2.65) guarantees that the right hand rule will then hold in any reference frame. Note that if  $A$  had been an improper orthogonal matrix with determinant  $-1$ , an undesirable minus sign would have appeared between the two sides of Eq. (2.63). Seen from this point of view, the problem with improper orthogonal matrices is that they would change a right-handed reference frame into a left-handed one; they would turn the reference frame inside out.

## 2.6 Some Useful Reference Frames

Several reference frames in three dimensions are of special interest for attitude analysis. We will discuss the most important of these in this section. In general, a reference frame is specified by the location of its origin and the orientation of its coordinate axes, with the orientation being much more important for attitude analysis.

### 2.6.1 *Spacecraft Body Frame*

A spacecraft body frame is defined by an origin at a specified point in the spacecraft body and three Cartesian axes. A body frame is used to align the various components during spacecraft assembly. Components will generally shift due to the large forces experienced during launch, though, and can also move while on orbit due to thermal deformations. Every effort is made to limit these motions, but they cannot always be neglected. Whether they are negligible or not depends on the pointing accuracy required of the spacecraft. As an additional complication, some components of the spacecraft, such as solar arrays or gimballed instruments, are moved quite deliberately. Therefore, it is quite common to define the body coordinate system operationally as the orientation of some sufficiently rigid *navigation base*, which is a subsystem of the spacecraft including the most critical attitude sensors and payload instruments. The navigation base often takes the form of a specially constructed *optical bench*, with its attached sensors and payload components. The purpose of attitude estimation and attitude control is to ascertain and to control the orientation of the navigation base relative to some external reference frame.

### 2.6.2 *Inertial Reference Frames*

An inertial reference frame is a frame in which Newton's laws of motion are valid. It is a well known fact of classical mechanics that any frame moving at constant velocity and without rotation with respect to an inertial frame is also inertial [5].

The existence of these dynamically preferred frames raises the interesting question of whether there is something with respect to which all inertial frames are non-rotating and unaccelerated. Weinberg concludes in [25, p. 474] that “...inertial frames are any reference frames that move at constant velocity, and without rotation, relative to frames in which the universe appears spherically symmetric.” This characterization is consistent with *Mach’s Principle*, the hypothesis that inertial frames are somehow determined by the mass of everything in the universe.

Celestial reference frames with their axes fixed relative to distant “fixed” stars are the best realizations of inertial frames. The standard as of this writing is the *International Celestial Reference System* (ICRF) with its axes fixed with respect to the positions of several hundred distant extragalactic sources of radio waves, determined by very long baseline interferometry [8, 11]. The  $z$  axis of this frame is aligned with the Earth’s North pole, and the  $x$  axis with the *vernal equinox*, the intersection of the Earth’s equatorial plane with the plane of the Earth’s orbit around the Sun, in the direction of the Sun’s position relative to the Earth on the first day of spring. Unfortunately, neither the polar axis nor the ecliptic plane is inertially fixed, so the ICRF axes are defined to be the *mean* orientations of the pole and the vernal equinox (the positions with short-period motions removed by dynamic models) at some fixed epoch time. The origin of the ICRF is at the center-of-mass of the solar system.

An approximate inertial frame, known as the *Geocentric Inertial Frame* (GCI) has its origin at the center of mass of the Earth. This frame has a linear acceleration because of the Earth’s circular orbit about the Sun, but this is unimportant for attitude analysis. The axes of a “mean of epoch” GCI frame are aligned with the mean North pole and mean vernal equinox at some epoch. The GCI frame is denoted by the triad  $\{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\}$ , as shown in Fig. 2.3.

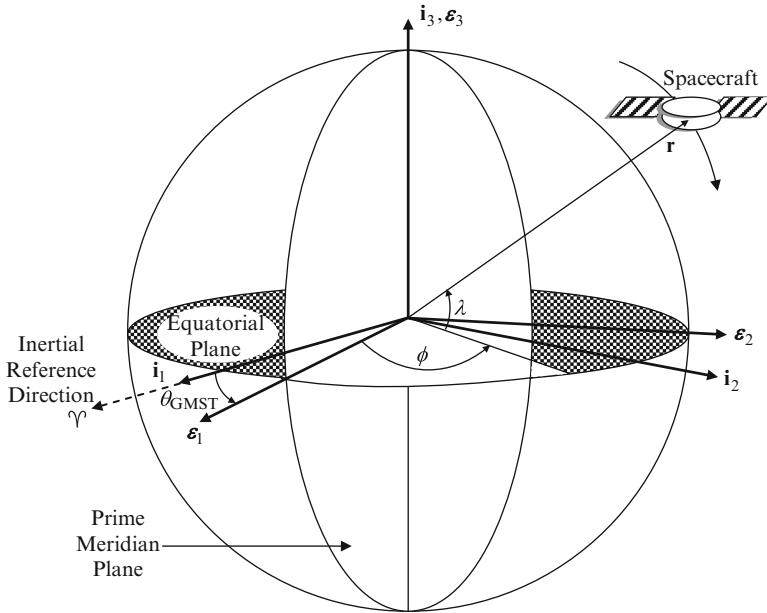
### 2.6.3 Earth-Centered/Earth-Fixed Frame

The Earth-Centered/Earth-Fixed (ECEF) Frame is denoted by  $\{\epsilon_1, \epsilon_2, \epsilon_3\}$  as shown in Fig. 2.3. This frame is similar to the GCI frame with  $\epsilon_3 = \mathbf{i}_3$ ; however, the  $\epsilon_1$  axis points in the direction of the Earth’s prime meridian, and the  $\epsilon_2$  axis completes the right-handed system. Unlike the GCI frame, the ECEF frame rotates with the Earth. The rotation angle is known as the Greenwich Mean Sidereal Time (GMST) angle and is denoted by  $\theta_{\text{GMST}}$  in Fig. 2.3.

The transformation of a position vector  $\mathbf{r}$  from its GCI representation  $\mathbf{r}_I$  to its ECEF representation  $\mathbf{r}_E$  follows

$$\mathbf{r}_E = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A_{EI} \mathbf{r}_I = \begin{bmatrix} \cos \theta_{\text{GMST}} & \sin \theta_{\text{GMST}} & 0 \\ -\sin \theta_{\text{GMST}} & \cos \theta_{\text{GMST}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{r}_I \quad (2.67)$$





**Fig. 2.3** Definitions of various reference frames

Determining the GMST angle requires the Julian date,  $JD$ . For a given year  $Y$  (between 1901 and 2099), month  $M$ , day  $D$ , hour  $h$ , minute  $m$ , and second  $s$ , the Julian date is calculated by [24]

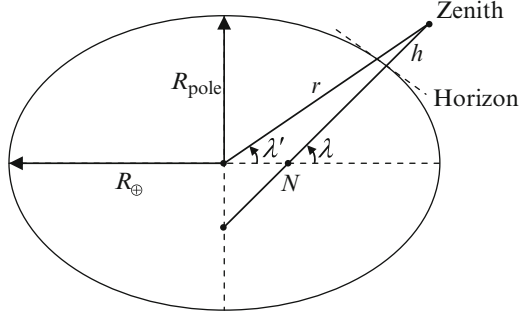
$$JD(Y, M, D, h, m, s) = 1,721,013.5 + 367Y - \text{INT} \left\{ \frac{7}{4} \left[ Y + \text{INT} \left( \frac{M+9}{12} \right) \right] \right\} \\ + \text{INT} \left( \frac{275M}{9} \right) + D + \frac{60h + m + s/60^*}{1440} \quad (2.68)$$

where INT denotes the integer part and  $60^*$  denotes using 61 s for days with a leap second. We need to compute  $T_0$ , the number of Julian centuries elapsed from the epoch J2000.0 to zero hours of the date in question:

$$T_0 = \frac{JD(Y, M, D, 0, 0, 0) - 2,451,545}{36,525} \quad (2.69)$$

The GCI coordinate system is fixed relative to the stars, not the Sun, so the GMST angle is the mean *sidereal* time at zero longitude. A *sidereal day* is the length of time that passes between successive crossings of a given projected meridian by a given fixed star in the sky. It is approximately 3 min and 56 s shorter than a *solar*

**Fig. 2.4** Geocentric and geodetic latitude



day of 86,400 s, which is the length of time that elapses between the Sun reaching its highest point in the sky two consecutive times [1]. Therefore  $\theta_{\text{GMST}}$  in units of seconds is calculated by

$$\begin{aligned} \theta_{\text{GMST}} = & 24,110.54841 + 8,640,184.812866 T_0 + 0.093104 T_0^2 \\ & - 6.2 \times 10^{-6} T_0^3 + 1.002737909350795(3600 h + 60 m + s) \end{aligned} \quad (2.70)$$

This quantity is next reduced to a range from 0 to 86,400 s by adding/subtracting multiples of 86,400. Then  $\theta_{\text{GMST}}$  in degrees is obtained by dividing by 240, because  $1 \text{ s} = 1/240^\circ$ .

The ECEF position vector can be specified by its magnitude  $r \equiv \|\mathbf{r}_E\| = \|\mathbf{r}_I\|$ , longitude  $\phi$ , and geocentric latitude  $\lambda'$ . An alternative description in terms of geodetic latitude,  $\lambda$ , is often employed. The Earth's geoid can be approximated by an ellipsoid of revolution about its minor axis, the Earth's rotation axis, as shown in Fig. 2.4 [11]. The geocentric latitude  $\lambda' = \sin^{-1}(z/r)$  is the angle between the equatorial plane and the radius vector from the center of the Earth. The geodetic latitude  $\lambda$  is the angle between the equatorial plane and the normal to the reference ellipsoid. The *flattening* of the ellipsoid is given by

$$f = \frac{R_\oplus - R_{\text{pole}}}{R_\oplus} \quad (2.71)$$

where  $R_\oplus$  is the equatorial radius of the Earth and  $R_{\text{pole}}$  is the distance from the center of the Earth to a pole. The eccentricity of the reference ellipsoid is given by

$$e = \sqrt{1 - (1 - f)^2} = \sqrt{f(2 - f)} \quad (2.72)$$

Many reference ellipsoid models exist, but for all of them the difference between the equatorial and polar radii is less than 22 km, so that  $f \approx 1/298.257$  is a valid approximation. A common ellipsoid model is given by the World Geodetic System 1984 model (WGS-84), with semimajor axis  $R_\oplus \equiv a = 6,378,137.0 \text{ m}$  and semiminor axis  $R_{\text{pole}} \equiv b = 6,356,752.3142 \text{ m}$ . The eccentricity of this ellipsoid is given by  $e = 0.0818$ .

To determine the ECEF position vector from the geodetic coordinates, latitude  $\lambda$ , longitude  $\phi$ , and height  $h$ , we first compute the distance between the  $z$  axis and the normal to the ellipsoid, finding [3]

$$N = \frac{R_{\oplus}}{\sqrt{1 - e^2 \sin^2 \lambda}} \quad (2.73)$$

Then the ECEF position coordinates are given by

$$x = (N + h) \cos \lambda \cos \phi \quad (2.74a)$$

$$y = (N + h) \cos \lambda \sin \phi \quad (2.74b)$$

$$z = [N(1 - e^2) + h] \sin \lambda \quad (2.74c)$$

This gives the following relationship between geocentric and geodetic latitudes

$$\tan \lambda = \frac{N + h}{N(1 - e^2) + h} \tan \lambda' \quad (2.75)$$

which has the first order approximation in the flattening  $f$

$$\lambda = \lambda' + \frac{f R_{\oplus}}{R_{\oplus} + h} \sin(2\lambda') \quad (2.76)$$

The difference between geodetic and geocentric latitudes amounts to 12 arcmin at most [11].

The conversion from ECEF to geodetic coordinates is not straightforward, but a closed-form solution is given in [20]. Given,  $x$ ,  $y$  and  $z$  in ECEF coordinates, the solution is given by

$$e^2 = 1 - b^2/a^2, \quad \epsilon^2 = a^2/b^2 - 1, \quad \rho = \sqrt{x^2 + y^2} \quad (2.77a)$$

$$p = |z|/\epsilon^2, \quad s = \rho^2/(e^2 \epsilon^2), \quad q = p^2 - b^2 + s \quad (2.77b)$$

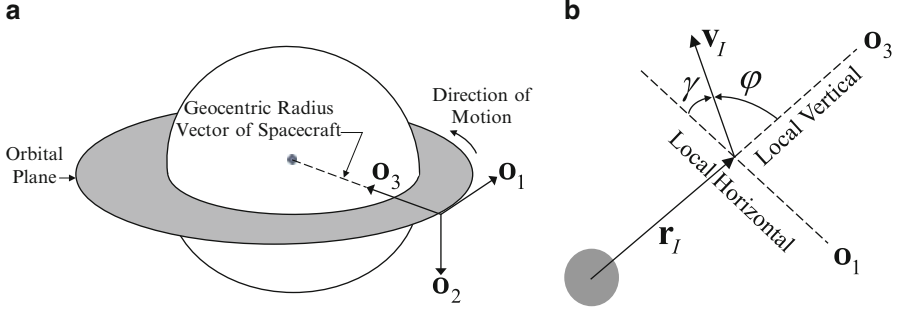
$$u = p/\sqrt{q}, \quad v = b^2 u^2/q, \quad P = 27vs/q, \quad Q = (\sqrt{P+1} + \sqrt{P})^{2/3} \quad (2.77c)$$

$$t = (1 + Q + 1/Q)/6, \quad c = \sqrt{u^2 - 1 + 2t}, \quad w = (c - u)/2 \quad (2.77d)$$

$$d = \text{sign}(z)\sqrt{q} \left[ w + (\sqrt{t^2 + v - uw - t/2 - 1/4})^{1/2} \right] \quad (2.77e)$$

$$N = a\sqrt{1 + \epsilon^2 d^2/b^2}, \quad \lambda = \sin^{-1}[(\epsilon^2 + 1)(d/N)] \quad (2.77f)$$

$$h = \rho \cos \lambda + z \sin \lambda - a^2/N, \quad \phi = \text{atan2}(y, x) \quad (2.77g)$$



**Fig. 2.5** Local-vertical/local-horizontal frame. (a) Frame definition. (b) Flight path angle

where  $\text{atan2}(y, x)$  is the standard function giving the argument of the complex number  $x + iy$ . It should be noted that longitude here is assumed to range between  $-180^\circ$  (West) to  $+180^\circ$  (East).

### 2.6.4 Local-Vertical/Local-Horizontal Frame

It is often convenient, especially for Earth-pointing spacecraft, to define a reference frame referenced to the spacecraft's orbit, which we will identify by the subscript  $O$ . The most common case is the Local-Vertical/Local-Horizontal (LVLH) orbit frame shown in Fig. 2.5a. It has its  $z$  axis  $\mathbf{o}_3$  pointing along the nadir vector, directly toward the center of the Earth from the spacecraft,<sup>4</sup> and its  $y$  axis  $\mathbf{o}_2$  pointing along the negative orbit normal, in the direction opposite to the spacecraft's orbital angular velocity. The  $x$  axis  $\mathbf{o}_1$  completes the right-handed triad. The representations of these vectors in an inertial frame  $I$  are

$$\mathbf{o}_{3I} = -\mathbf{r}_I / \|\mathbf{r}_I\| \equiv -g_3 \mathbf{r}_I \quad (2.78a)$$

$$\mathbf{o}_{2I} = -(\mathbf{r}_I \times \mathbf{v}_I) / \|\mathbf{r}_I \times \mathbf{v}_I\| \equiv -g_2 (\mathbf{r}_I \times \mathbf{v}_I) \quad (2.78b)$$

$$\mathbf{o}_{1I} = \mathbf{o}_{2I} \times \mathbf{o}_{3I} = g_2 g_3 (\mathbf{r}_I \times \mathbf{v}_I) \times \mathbf{r}_I = g_2 g_3 [\|\mathbf{r}_I\|^2 \mathbf{v}_I - (\mathbf{r}_I \cdot \mathbf{v}_I) \mathbf{r}_I] \quad (2.78c)$$

where  $\mathbf{r}_I$  and  $\mathbf{v}_I = \dot{\mathbf{r}}_I$  are the spacecraft position and velocity in the  $I$  frame. Note that the  $x$  axis is in the direction of the velocity for a circular orbit. The rotation matrix from the  $O$  frame to the  $I$  frame can be expressed by Eq. (2.45) as

$$A_{IO} = [\mathbf{o}_{1I} \ \mathbf{o}_{2I} \ \mathbf{o}_{3I}] \quad (2.79)$$

<sup>4</sup>This is the *geocentric* nadir vector. Some spacecraft use the *geodetic* nadir vector, which is normal to the surface of the reference ellipsoid, but we will not consider this complication.

The angle between the velocity vector and the local horizontal ( $\mathbf{o}_{1I} - \mathbf{o}_{2I}$ ) plane is called the *flight-path angle*, denoted by  $\gamma$  in Fig. 2.5b. The specific angular momentum is given by  $\mathbf{h}_I = \mathbf{r}_I \times \mathbf{v}_I$ , as discussed in Appendix C. From Eq. (2.58) we have  $h \equiv \|\mathbf{h}_I\| = rv \sin \varphi$ , where  $r \equiv \|\mathbf{r}_I\|$  and  $v \equiv \|\mathbf{v}_I\|$ . Since  $\gamma + \varphi = 90^\circ$ , then the flight path angle can be computed using

$$\cos \gamma = \frac{h}{rv} \quad (2.80)$$

The sign of  $\gamma$  is the same as the sign of  $\mathbf{r}_I \cdot \mathbf{v}_I$ . Notice that  $\mathbf{o}_{1I}$  is in the direction of the spacecraft velocity and the flight path angle is always zero for a circular orbit, because  $\mathbf{r}_I \cdot \mathbf{v}_I = 0$  in this case.

## 2.7 Quaternions

We consider a quaternion to be a four-component vector with some additional operations defined on it. A quaternion  $\mathbf{q}$  has a three-vector part  $\mathbf{q}_{1:3}$  and a scalar part  $q_4$

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_{1:3} \\ q_4 \end{bmatrix} \quad \text{where} \quad \mathbf{q}_{1:3} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (2.81)$$

The most important added quaternion operations are two different products of a pair of quaternions  $\bar{\mathbf{q}}$  and  $\mathbf{q}$

$$\bar{\mathbf{q}} \otimes \mathbf{q} = \begin{bmatrix} q_4 \bar{\mathbf{q}}_{1:3} + \bar{q}_4 \mathbf{q}_{1:3} - \bar{\mathbf{q}}_{1:3} \times \mathbf{q}_{1:3} \\ \bar{q}_4 q_4 - \bar{\mathbf{q}}_{1:3} \cdot \mathbf{q}_{1:3} \end{bmatrix} \quad (2.82a)$$

$$\bar{\mathbf{q}} \odot \mathbf{q} = \begin{bmatrix} q_4 \bar{\mathbf{q}}_{1:3} + \bar{q}_4 \mathbf{q}_{1:3} + \bar{\mathbf{q}}_{1:3} \times \mathbf{q}_{1:3} \\ \bar{q}_4 q_4 - \bar{\mathbf{q}}_{1:3} \cdot \mathbf{q}_{1:3} \end{bmatrix} \quad (2.82b)$$

Notice that these definitions differ only in the sign of the cross product in the vector part, from which it follows that<sup>5</sup>

$$\bar{\mathbf{q}} \otimes \mathbf{q} = \mathbf{q} \odot \bar{\mathbf{q}} \quad (2.83)$$

Our quaternions are conceptually different from those introduced by Hamilton in 1844, before the introduction of vector notation. Hamilton defined a quaternion as  $q = q_0 + iq_1 + jq_2 + kq_3$ , a hypercomplex extension of a complex number  $z = x + iy$ , with  $i$ ,  $j$ , and  $k$  obeying the relations  $i^2 = j^2 = k^2 = -1$ ,  $ij = -ji = k$ ,

<sup>5</sup>The notation  $\bar{\mathbf{q}} \otimes \mathbf{q}$  was introduced in [7], and the notation  $\bar{\mathbf{q}} \odot \mathbf{q}$  is a modification of notation introduced in [13].

$jk = -kj = i$ , and  $ki = -ik = j$ . Hamilton's product  $\bar{q}q$  corresponds to our product  $\bar{\mathbf{q}} \odot \mathbf{q}$ , but the product  $\bar{\mathbf{q}} \otimes \mathbf{q}$  has proven to be more useful in attitude analysis. Some authors' notations differ from ours in labeling the scalar part of a quaternion  $q_0$  and putting it at the top of the column vector. Care must be taken to thoroughly understand the conventions embodied in any quaternion equation that one chooses to reference.

Quaternion multiplication is associative,  $\mathbf{q} \otimes (\bar{\mathbf{q}} \otimes \bar{\bar{\mathbf{q}}}) = (\mathbf{q} \otimes \bar{\mathbf{q}}) \otimes \bar{\bar{\mathbf{q}}}$  and distributive,  $\mathbf{q} \otimes (\bar{\mathbf{q}} + \bar{\bar{\mathbf{q}}}) = \mathbf{q} \otimes \bar{\mathbf{q}} + \mathbf{q} \otimes \bar{\bar{\mathbf{q}}}$ , but not commutative in general,  $\mathbf{q} \otimes \bar{\mathbf{q}} \neq \bar{\mathbf{q}} \otimes \mathbf{q}$ . This parallels the situation for matrix multiplication. In those cases for which  $\mathbf{q} \otimes \bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes \mathbf{q}$ , the quaternions  $\mathbf{q}$  and  $\bar{\mathbf{q}}$  are said to commute. Analogous equations hold for the product  $\bar{\mathbf{q}} \odot \mathbf{q}$ .

Quaternion products can be represented by matrix multiplication, very much like the cross product:

$$\mathbf{q} \otimes \bar{\mathbf{q}} = [\mathbf{q} \otimes] \bar{\mathbf{q}} = \bar{\mathbf{q}} \odot \mathbf{q} \quad (2.84a)$$

$$\mathbf{q} \odot \bar{\mathbf{q}} = [\mathbf{q} \odot] \bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes \mathbf{q} \quad (2.84b)$$

where

$$[\mathbf{q} \otimes] \equiv \begin{bmatrix} q_4 I_3 - [\mathbf{q}_{1:3} \times] & \mathbf{q}_{1:3} \\ -\mathbf{q}_{1:3}^T & q_4 \end{bmatrix} = [\Psi(\mathbf{q}) \quad \mathbf{q}] \quad (2.85)$$

and

$$[\mathbf{q} \odot] \equiv \begin{bmatrix} q_4 I_3 + [\mathbf{q}_{1:3} \times] & \mathbf{q}_{1:3} \\ -\mathbf{q}_{1:3}^T & q_4 \end{bmatrix} = [\mathcal{E}(\mathbf{q}) \quad \mathbf{q}] \quad (2.86)$$

with  $\Psi(\mathbf{q})$  and  $\mathcal{E}(\mathbf{q})$  being the  $4 \times 3$  matrices

$$\Psi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_3 - [\mathbf{q}_{1:3} \times] \\ -\mathbf{q}_{1:3}^T \end{bmatrix} = \begin{bmatrix} q_4 & q_3 & -q_2 \\ -q_3 & q_4 & q_1 \\ q_2 & -q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix} \quad (2.87)$$

$$\mathcal{E}(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_3 + [\mathbf{q}_{1:3} \times] \\ -\mathbf{q}_{1:3}^T \end{bmatrix} = \begin{bmatrix} q_4 - q_3 & q_2 \\ q_3 & q_4 - q_1 \\ -q_2 & q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix} \quad (2.88)$$

It is easy to show that

$$\Psi^T(\mathbf{q})\Psi(\mathbf{q}) = \mathcal{E}^T(\mathbf{q})\mathcal{E}(\mathbf{q}) = \|\mathbf{q}\|^2 I_3 \quad (2.89a)$$

$$\Psi(\mathbf{q})\Psi^T(\mathbf{q}) = \mathcal{E}(\mathbf{q})\mathcal{E}^T(\mathbf{q}) = \|\mathbf{q}\|^2 I_4 - \mathbf{q}\mathbf{q}^T \quad (2.89b)$$

$$\Psi^T(\mathbf{q})\mathbf{q} = \mathcal{E}^T(\mathbf{q})\mathbf{q} = \mathbf{0}_3 \quad (2.89c)$$

from which it follows that  $\|\mathbf{q}\|^{-1}[\mathbf{q} \otimes]$  and  $\|\mathbf{q}\|^{-1}[\mathbf{q} \odot]$  are orthogonal matrices.

We define the identity quaternion

$$\mathbf{I}_q \equiv \begin{bmatrix} \mathbf{0}_3 \\ 1 \end{bmatrix} \quad (2.90)$$

which obeys  $\mathbf{I}_q \otimes \mathbf{q} = \mathbf{q} \otimes \mathbf{I}_q = \mathbf{I}_q \odot \mathbf{q} = \mathbf{q} \odot \mathbf{I}_q = \mathbf{q}$ , as required of the identity.

We also define the conjugate  $\mathbf{q}^*$  of a quaternion, obtained by changing the sign of the three-vector part:

$$\mathbf{q}^* = \begin{bmatrix} \mathbf{q}_{1:3} \\ q_4 \end{bmatrix}^* \equiv \begin{bmatrix} -\mathbf{q}_{1:3} \\ q_4 \end{bmatrix} \quad (2.91)$$

The product of a quaternion with its conjugate is equal to the square of its norm times the identity quaternion

$$\mathbf{q} \otimes \mathbf{q}^* = \mathbf{q}^* \otimes \mathbf{q} = \mathbf{q} \odot \mathbf{q}^* = \mathbf{q}^* \odot \mathbf{q} = \|\mathbf{q}\|^2 \mathbf{I}_q \quad (2.92)$$

The conjugate of the product of two quaternions  $\bar{\mathbf{q}}$  and  $\mathbf{q}$  is the product of the conjugates in the opposite order  $(\mathbf{p} \otimes \mathbf{q})^* = \mathbf{q}^* \otimes \mathbf{p}^*$ . This relation, Eq. (2.92), and the associativity of quaternion multiplication can be used to show that

$$\|\mathbf{p} \otimes \mathbf{q}\| = \|\mathbf{p} \odot \mathbf{q}\| = \|\mathbf{p}\| \|\mathbf{q}\| \quad (2.93)$$

It is not difficult to see that

$$[\mathbf{q}^* \otimes] = [\mathbf{q} \otimes]^T \quad \text{and} \quad [\mathbf{q}^* \odot] = [\mathbf{q} \odot]^T \quad (2.94)$$

The inverse of any quaternion having nonzero norm is defined by

$$\mathbf{q}^{-1} \equiv \mathbf{q}^* / \|\mathbf{q}\|^2 \quad (2.95)$$

so that  $\mathbf{q} \otimes \mathbf{q}^{-1} = \mathbf{q}^{-1} \otimes \mathbf{q} = \mathbf{q} \odot \mathbf{q}^{-1} = \mathbf{q}^{-1} \odot \mathbf{q} = \mathbf{I}_q$ , as required by the definition of an inverse. The inverse of the product of two quaternions is the product of the inverses in the opposite order  $(\mathbf{p} \otimes \mathbf{q})^{-1} = \mathbf{q}^{-1} \otimes \mathbf{p}^{-1}$ .

We will overload the quaternion product notation to allow us to multiply a three-component vector  $\mathbf{x}$  and a quaternion, using the definitions

$$\mathbf{x} \otimes \mathbf{q} \equiv \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} \otimes \mathbf{q} = [\mathbf{x} \otimes] \mathbf{q} \quad \text{and} \quad \mathbf{q} \otimes \mathbf{x} \equiv \mathbf{q} \otimes \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} \quad (2.96)$$

with analogous definitions for  $\mathbf{x} \odot \mathbf{q}$ ,  $[\mathbf{x} \odot]$ , and  $\mathbf{q} \odot \mathbf{x}$ . Note that the matrices

$$[\mathbf{x} \otimes] = \begin{bmatrix} -[\mathbf{x} \times] & \mathbf{x} \\ -\mathbf{x}^T & 0 \end{bmatrix} \equiv \Omega(\mathbf{x}) \quad \text{and} \quad (2.97a)$$

$$[\mathbf{x} \odot] = \begin{bmatrix} [\mathbf{x} \times] & \mathbf{x} \\ -\mathbf{x}^T & 0 \end{bmatrix} \equiv \Gamma(\mathbf{x}) \quad (2.97b)$$

are skew-symmetric. The relation

$$\mathbf{x} \otimes \mathbf{q} = \mathbf{q} \odot \mathbf{x} = [\mathcal{E}(\mathbf{q}) \ \mathbf{q}] \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = \mathcal{E}(\mathbf{q}) \mathbf{x} \quad (2.98)$$

is often very useful.

## 2.8 Rotations and Euler's Theorem

The discipline of spacecraft attitude determination is basically the study of methods for estimating the proper orthogonal matrix that transforms vectors from a reference frame fixed in space to a frame fixed in the spacecraft body. Thus it is the study of proper orthogonal  $3 \times 3$  matrices, or matrices in the group  $\text{SO}(3)$ . We will refer to these as rotation matrices or *attitude* matrices and denote them by the letter  $A$ .

*Euler's Theorem*<sup>6</sup> states one of the most important properties of attitude matrices, namely that any rotation is a rotation about a *fixed* axis. Recall that the transformation of a vector representation  $\mathbf{x}_F$  from reference frame  $F$  to reference frame  $G$  by the attitude matrix  $A_{GF}$  is given by

$$A_{GF} \mathbf{x}_F = \mathbf{x}_G \quad (2.99)$$

Euler's Theorem asserts the existence of a vector  $\mathbf{e}$  along the direction of the rotation axis that has the same representation in frame  $G$  as in frame  $F$ . This means that we can substitute  $\mathbf{x}_F = \mathbf{x}_G = \mathbf{e}$  in Eq. (2.99) and state Euler's theorem algebraically as

$$A \mathbf{e} = \mathbf{e} \quad (2.100)$$

This is a special case of an eigenvalue/eigenvector relationship. An *eigenvector* of a general square matrix  $M$  is a nonzero vector  $\mathbf{x}$  for which multiplication by  $M$  has the same effect as multiplication by a scalar, i.e.

$$M \mathbf{x} = \lambda \mathbf{x} \quad (2.101)$$

where the scalar  $\lambda$  is the *eigenvalue* corresponding to the eigenvector  $\mathbf{x}$ . The solution for  $\mathbf{x}$  is only determined up to a scale factor in general, so eigenvectors are almost invariably given as unit vectors. In order for Eq. (2.101) to have a nonzero solution for  $\mathbf{x}$ , the matrix  $(\lambda I - M)$  must be singular.<sup>7</sup> Therefore, from Eq. (2.16) we have

$$\det(\lambda I - M) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1} \lambda + \alpha_n = 0 \quad (2.102)$$

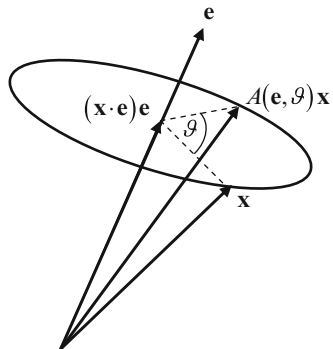
---

<sup>6</sup>Leonhard Euler (1707–1783) laid the foundations for the analysis of rotations, and his fingerprints are all over the subject. Thus, attaching his name to anything serves poorly for distinguishing it from other results also bearing his name.

<sup>7</sup>Otherwise, we would have  $\mathbf{x} = (\lambda I - M)^{-1} \mathbf{0} = \mathbf{0}$ .



**Fig. 2.6** Euler axis/angle rotation



This polynomial equation of degree  $n$  is known as the *characteristic equation* of  $M$ . The eigenvalues are the  $n$  roots of the characteristic equation, counting multiple roots by their multiplicity.

In the language of eigenvalues and eigenvectors, Euler's theorem says that one of the eigenvalues of an attitude matrix has the value  $\lambda = 1$ . The characteristic equation of a  $3 \times 3$  matrix  $M$  is easily found by explicit computation of the determinant to take the form

$$\lambda^3 - \lambda^2 \text{tr } M + \lambda \text{tr}(\text{adj } M) - \det M = 0 \quad (2.103)$$

For the special case of a proper orthogonal  $3 \times 3$  matrix, we find the characteristic equation to be

$$0 = \lambda^3 - \lambda^2 \text{tr } A + \lambda \text{tr } A - 1 = (\lambda - 1)[\lambda^2 + \lambda(1 - \text{tr } A) + 1] \quad (2.104)$$

This clearly has a root  $\lambda = 1$ , which proves Euler's theorem with the rotation axis  $\mathbf{e}$  being the eigenvector corresponding to this eigenvalue.

## 2.9 Attitude Representations

### 2.9.1 Euler Axis/Angle Representation

We have seen that every proper orthogonal  $3 \times 3$  matrix has a rotation axis specified by a unit vector  $\mathbf{e}$ . The only other parameter needed to completely specify the matrix is the angle of rotation  $\vartheta$  about this axis. This axis and angle are known as the *Euler axis* and *Euler angle* of the rotation. We will now show how to parameterize the attitude matrix in terms of these parameters.

Figure 2.6 depicts the rotation of an arbitrary vector  $\mathbf{x}$  through an angle  $\vartheta$  about an axis  $\mathbf{e}$ . This figure is actually more illustrative of an active (alibi) rotation of the

vector, but we want to represent the transformation as a passive (alias) rotation of the reference frame [5]. We make the connection between these two interpretations by referring to Fig. 2.2b,c, which show that we can move from one interpretation to the other by simply changing the direction of the rotation. The mapping of the vector  $\mathbf{x}$  into the rotated reference frame is denoted as  $A(\mathbf{e}, \vartheta)\mathbf{x}$ , as indicated in the figure.

We express  $\mathbf{x}$  as the sum of vectors parallel and perpendicular to the rotation axis

$$\mathbf{x} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp} \quad (2.105)$$

where

$$\mathbf{x}_{\parallel} \equiv (\mathbf{x} \cdot \mathbf{e}) \mathbf{e} = (\mathbf{e} \mathbf{e}^T) \mathbf{x} \quad \text{and} \quad \mathbf{x}_{\perp} \equiv \mathbf{x} - (\mathbf{x} \cdot \mathbf{e}) \mathbf{e} = (I_3 - \mathbf{e} \mathbf{e}^T) \mathbf{x} \quad (2.106)$$

The rotation leaves  $\mathbf{x}_{\parallel}$  alone, but rotates  $\mathbf{x}_{\perp}$  out of the plane defined by  $\mathbf{e}$  and  $\mathbf{x}$ . The cross product  $\mathbf{e} \times \mathbf{x}$  is perpendicular to that plane and has magnitude  $\|\mathbf{e} \times \mathbf{x}\| = \|\mathbf{x}_{\perp}\|$  by Eq. (2.57), so the result of the norm-preserving rotation by angle  $\vartheta$  is

$$A(\mathbf{e}, \vartheta) \mathbf{x} = \mathbf{x}_{\parallel} + (\cos \vartheta) \mathbf{x}_{\perp} - (\sin \vartheta) \mathbf{e} \times \mathbf{x} \quad (2.107)$$

The sign of the last term is chosen to agree with the sense of the rotation shown in Fig. 2.2b, which is regarded as a rotation in the positive sense about  $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$ , the outward normal from the plane of the figure.

Because  $\mathbf{x}$  is an arbitrary vector, Eq. (2.107) means, inserting the definitions of  $\mathbf{x}_{\parallel}$  and  $\mathbf{x}_{\perp}$ , that

$$A(\mathbf{e}, \vartheta) = (\cos \vartheta) I_3 - \sin \vartheta [\mathbf{e} \times] + (1 - \cos \vartheta) \mathbf{e} \mathbf{e}^T \quad (2.108)$$

This is the *Euler axis/angle* parameterization of an attitude matrix. We can also express this, using Eq. (2.56b) as

$$A(\mathbf{e}, \vartheta) = I_3 - \sin \vartheta [\mathbf{e} \times] + (1 - \cos \vartheta) [\mathbf{e} \times]^2 \quad (2.109)$$

The attitude matrix is expressed in explicit component form as

$$A(\mathbf{e}, \vartheta) = \begin{bmatrix} c + (1-c)e_1^2 & (1-c)e_1e_2 + se_3 & (1-c)e_1e_3 - se_2 \\ (1-c)e_2e_1 - se_3 & c + (1-c)e_2^2 & (1-c)e_2e_3 + se_1 \\ (1-c)e_3e_1 + se_2 & (1-c)e_3e_2 - se_1 & c + (1-c)e_3^2 \end{bmatrix} \quad (2.110)$$

where we have written  $c \equiv \cos \vartheta$  and  $s \equiv \sin \vartheta$  for conciseness. The attitude matrix appears to depend on four parameters, but there are only three independent parameters owing to the constraint  $\|\mathbf{e}\| = 1$ . The nine-component attitude matrix has only three independent parameters because of the orthogonality constraint  $AA^T = I$ . This matrix constraint is equivalent to six scalar constraints rather than nine because the product  $AA^T$  is symmetric.

Equations (2.108)–(2.110) show that the attitude matrix is a periodic function of the rotation angle over an unlimited range with period  $2\pi$ . Some useful identities satisfied by the Euler axis/angle representation are

$$A(\mathbf{e}, \vartheta) = A(-\mathbf{e}, -\vartheta) \quad (2.111a)$$

$$A^{-1}(\mathbf{e}, \vartheta) = A^T(\mathbf{e}, \vartheta) = A(-\mathbf{e}, \vartheta) = A(\mathbf{e}, -\vartheta) \quad (2.111b)$$

$$A(\mathbf{e}, \pi) = A(-\mathbf{e}, \pi) = 2\mathbf{e}\mathbf{e}^T - I_3 \quad (2.111c)$$

We now turn to the question of finding the rotation axis and angle corresponding to a given attitude matrix. Noting from Eq. (2.108) that

$$\text{tr}A(\mathbf{e}, \vartheta) = 1 + 2\cos\vartheta \quad (2.112)$$

we see that the rotation angle is given by

$$\vartheta = \cos^{-1} \left( \frac{\text{tr}A(\mathbf{e}, \vartheta) - 1}{2} \right) \quad (2.113)$$

If  $\cos\vartheta = 1$  the attitude matrix is  $A(\mathbf{e}, \vartheta) = I_3$ , and the rotation axis is clearly undefined. If  $-1 < \cos\vartheta < 1$ , the axis of rotation is given by

$$\mathbf{e} = \frac{1}{2\sin\vartheta} \begin{bmatrix} A_{23}(\mathbf{e}, \vartheta) - A_{32}(\mathbf{e}, \vartheta) \\ A_{31}(\mathbf{e}, \vartheta) - A_{13}(\mathbf{e}, \vartheta) \\ A_{12}(\mathbf{e}, \vartheta) - A_{21}(\mathbf{e}, \vartheta) \end{bmatrix} \quad (2.114)$$

If  $\cos\vartheta = -1$  the axis of rotation can be found by normalizing any nonzero column of

$$A(\mathbf{e}, \vartheta) + I_3 = 2\mathbf{e}\mathbf{e}^T \quad (2.115)$$

because all the columns of this matrix are parallel to  $\mathbf{e}$ . The overall sign of the rotation axis vector is undetermined in this case, but Eq. (2.111c) shows that this sign makes no difference.

The other two eigenvalues of the attitude matrix are the other two roots of Eq. (2.104). Inserting the value of  $\text{tr}A(\mathbf{e}, \vartheta)$  into this equation gives

$$0 = \lambda^2 - 2\lambda\cos\vartheta + 1 = \lambda^2 - \lambda(e^{i\vartheta} + e^{-i\vartheta}) + 1 = (\lambda - e^{i\vartheta})(\lambda - e^{-i\vartheta}) \quad (2.116)$$

These two eigenvalues form a complex conjugate pair on the unit circle in the complex plane, and the corresponding eigenvectors are complex as well. This result can be generalized to proper orthogonal matrices of higher dimensionality. A matrix in  $\text{SO}(2n+1)$  has one eigenvalue equal to  $+1$  and  $n$  conjugate pairs on the unit circle in the complex plane. A matrix in  $\text{SO}(2n)$  has only the  $n$  conjugate pairs

on the unit circle. Thus Euler's theorem holds in any space of odd dimensionality, but not in a space with an even number of dimensions. If we regard Fig. 2.2b as a rotation in two-dimensional space rather than a projection onto the plane of a rotation in a higher-dimensional space, for example, it is easy to see that there is no invariant vector in the plane. This does not preclude the possibility of a pair of complex conjugate eigenvectors accidentally having the common value  $+1$ .

Cross-product and trigonometric identities can be used to find the unsurprising result of successive rotations about the same axis,

$$A(\mathbf{e}, \vartheta)A(\mathbf{e}, \varphi) = A(\mathbf{e}, \varphi)A(\mathbf{e}, \vartheta) = A(\mathbf{e}, \vartheta + \varphi) \quad (2.117)$$

but this is more easily derived using the quaternion representation of rotations. The composition of rotations about non-parallel axes does not have a simple form in the angle/axis representation. Another useful result holds for two attitude matrices  $A_0$  and  $A(\mathbf{e}, \vartheta)$ . From Eq. (2.108) we have

$$A_0A(\mathbf{e}, \vartheta)A_0^T = (\cos \vartheta) I_3 - (\sin \vartheta)A_0[\mathbf{e} \times]A_0^T + (1 - \cos \vartheta)A_0\mathbf{e}\mathbf{e}^TA_0^T \quad (2.118)$$

This can be written, using Eqs. (2.7c) and (2.63), as

$$A_0A(\mathbf{e}, \vartheta)A_0^T = A(A_0\mathbf{e}, \vartheta) \quad (2.119)$$

which shows that  $A_0A(\mathbf{e}, \vartheta)A_0^T$  is a rotation by the same angle as  $A$ , but about a rotated axis.

### 2.9.2 Rotation Vector Representation

It is convenient for analysis, but not for computations, to combine the Euler axis and angle into the three-component *rotation vector*

$$\boldsymbol{\vartheta} \equiv \vartheta \mathbf{e} \quad (2.120)$$

To express the attitude matrix in terms of the rotation vector, we insert the Taylor series expansions of the sine and cosine into Eq. (2.109), giving<sup>8</sup>

$$A(\mathbf{e}, \vartheta) = I_3 - [\mathbf{e} \times] \sum_{i=0}^{\infty} \frac{(-1)^i \vartheta^{2i+1}}{(2i+1)!} - [\mathbf{e} \times]^2 \sum_{i=1}^{\infty} \frac{(-1)^i \vartheta^{2i}}{(2i)!} \quad (2.121)$$

Equation (2.56b) can be used to show that  $[\mathbf{e} \times]^3 = -[\mathbf{e} \times]$ , so we have

$$A(\mathbf{e}, \vartheta) = I_3 + \sum_{i=0}^{\infty} \frac{[(-\vartheta \mathbf{e}) \times]^{2i+1}}{(2i+1)!} + \sum_{i=1}^{\infty} \frac{[(-\vartheta \mathbf{e}) \times]^{2i}}{(2i)!} = \sum_{n=0}^{\infty} \frac{[(-\vartheta \mathbf{e}) \times]^n}{n!} \quad (2.122)$$

---

<sup>8</sup>Note that these power series expansions assume that we measure angles in radians.

A function of a square matrix is defined by its Taylor series, so this expression for the attitude matrix in terms of the rotation vector can finally be written as

$$A(\mathbf{e}, \vartheta) = \exp(-[\vartheta \times]) \quad (2.123)$$

where  $\exp$  is the matrix exponential.

This is the first three-parameter representation of rotations that we have encountered. It is a very useful representation for the analysis of small rotations; but it is not useful for large rotations, mainly because it obscures the periodicity of the attitude matrix as a function of  $\vartheta$ . In particular, it is not at all obvious in this representation that a rotation by an angle  $\vartheta = 2\pi$  is equivalent to the identity transformation. Equation (2.113) shows that we can restrict the rotation angle to the range  $0 \leq \vartheta \leq \pi$ , which avoids this problem and gives a 1:1 mapping of rotations with  $\vartheta < \pi$  to rotation vectors. The rotation vectors fill a ball of radius  $\pi$ , with the two vectors at the ends of a diameter of the ball representing the same attitude according to Eq. (2.111c). This causes the difficulty that the rotation vector parameterization of a smoothly-varying attitude can jump discontinuously from one side of the ball to the other. This discontinuity can be avoided by giving up the 1:1 mapping by allowing rotation angles greater than  $\pi$ , but this causes other difficulties. It is an unavoidable fact that the rotation group has no global three-component parameterization without singular points [22]. The expense of computing the matrix exponential also renders the rotation vector parameterization impractical for numerical computations.

### 2.9.3 Quaternion Representation

Substituting  $\sin \vartheta = 2 \sin(\vartheta/2) \cos(\vartheta/2)$  and  $\cos \vartheta = \cos^2(\vartheta/2) - \sin^2(\vartheta/2)$  into Eq. (2.108) and defining the quaternion

$$\mathbf{q}(\mathbf{e}, \vartheta) = \begin{bmatrix} \mathbf{e} \sin(\vartheta/2) \\ \cos(\vartheta/2) \end{bmatrix} \quad (2.124)$$

gives the quaternion representation of the attitude matrix

$$\begin{aligned} A(\mathbf{q}) &= (q_4^2 - \|\mathbf{q}_{1:3}\|^2) I_3 - 2q_4[\mathbf{q}_{1:3} \times] + 2\mathbf{q}_{1:3} \mathbf{q}_{1:3}^T \\ &= \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_2q_1 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_3q_1 + q_2q_4) & 2(q_3q_2 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \end{aligned} \quad (2.125)$$

We are abusing the notation by using both  $A(\mathbf{q})$  and  $A(\mathbf{e}, \vartheta)$  to denote the attitude matrix, and we will abuse it further when we define other representations. The meaning of the argument of  $A(\cdot)$  will always be clear in context, however.

The four parameters of the quaternion representation were first considered by Euler, but their full significance was revealed by Rodrigues.<sup>9</sup> For this reason, they are often referred to as the *Euler symmetric parameters* or the *Euler-Rodrigues parameters*. The beauty of the quaternion representation is that it expresses the attitude matrix as a homogenous quadratic function of the elements of the quaternion, requiring no trigonometric or other transcendental function evaluations. Quaternions are more efficient for specifying rotations than the attitude matrix itself, having only four components instead of nine, and obeying only one constraint, the norm constraint, instead of the six constraints imposed on the attitude matrix by orthogonality.

Quaternions used to parameterize rotations are *unit quaternions*, i.e. quaternions with unit norm, as defined by Eq. (2.124). A unit quaternion always has an inverse, which is identical with its conjugate. Also, the discussion in Sect. 2.7 says that the matrices  $[\mathbf{q} \otimes]$  and  $[\mathbf{q} \odot]$  for a unit quaternion are orthogonal. We will now derive a useful expression for  $[\mathbf{q} \otimes]$  in terms of the rotation vector. Because it is a linear function of  $\mathbf{q}$ , we have

$$[\mathbf{q}(\mathbf{e}, \vartheta) \otimes] = \cos(\vartheta/2)[\mathbf{I}_q \otimes] + \sin(\vartheta/2)[\mathbf{e} \otimes] = \cos(\vartheta/2)I_4 + \sin(\vartheta/2)[\mathbf{e} \otimes] \quad (2.126)$$

Now expand the sine and cosine in Taylor series and use  $[\mathbf{e} \otimes]^2 = -I_4$ , to get

$$[\mathbf{q}(\mathbf{e}, \vartheta) \otimes] = \sum_{i=0}^{\infty} \frac{[(\vartheta \mathbf{e}/2) \otimes]^{2i}}{(2n)!} + \sum_{i=0}^{\infty} \frac{[(\vartheta \mathbf{e}/2) \otimes]^{2i+1}}{(2i+1)!} = \exp[(\vartheta/2) \otimes] \quad (2.127)$$

This result and its derivation are very similar to Eq. (2.123).

We now want to show how a rotation of a three-component vector  $\mathbf{x}$  can be implemented by quaternion multiplication. This is accomplished by the quaternion product

$$\mathbf{q} \otimes \mathbf{x} \otimes \mathbf{q}^* = [\mathbf{q}^* \odot](\mathbf{q} \otimes \mathbf{x}) = [\mathbf{q} \odot]^T [\mathbf{q} \otimes] \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathcal{E}^T(\mathbf{q})\Psi(\mathbf{q})\mathbf{x} \\ 0 \end{bmatrix} \quad (2.128)$$

where we have used several relations from Sect. 2.7. Explicit multiplication and comparison with Eq. (2.125) gives

$$\mathcal{E}^T(\mathbf{q})\Psi(\mathbf{q}) = A(\mathbf{q}) \quad (2.129)$$

---

<sup>9</sup>Olinde Rodrigues (1795–1851) obtained a doctorate in mathematics in 1815, with a thesis containing his well-known formula for the Legendre polynomials. He published nothing mathematical for the next 21 years, devoting himself to banking, the development of the French railways, utopian socialism, writing several pamphlets on banking, and editing an anthology of workers' poetry. Then he published eight papers between 1838 and 1845, including his 1840 paper [14] greatly advancing the state of the art in attitude analysis.

so that

$$\mathbf{q} \otimes \mathbf{x} \otimes \mathbf{q}^* = \begin{bmatrix} A(\mathbf{q}) \mathbf{x} \\ 0 \end{bmatrix} \quad (2.130)$$

This can be used to derive the rule for performing successive transformations using quaternions. Applying a transformation by a second quaternion  $\bar{\mathbf{q}}$  gives

$$\bar{\mathbf{q}} \otimes (\mathbf{q} \otimes \mathbf{x} \otimes \mathbf{q}^*) \otimes \bar{\mathbf{q}}^* = \bar{\mathbf{q}} \otimes [A(\mathbf{q}) \mathbf{x}] \otimes \bar{\mathbf{q}}^* = \begin{bmatrix} A(\bar{\mathbf{q}})A(\mathbf{q}) \mathbf{x} \\ 0 \end{bmatrix} \quad (2.131)$$

This transformation can also be written as

$$(\bar{\mathbf{q}} \otimes \mathbf{q}) \otimes \mathbf{x} \otimes (\bar{\mathbf{q}} \otimes \mathbf{q})^* = \begin{bmatrix} A(\bar{\mathbf{q}} \otimes \mathbf{q}) \mathbf{x} \\ 0 \end{bmatrix} \quad (2.132)$$

Because this relation must hold for any  $\mathbf{x}$ , we have proved that

$$A(\bar{\mathbf{q}} \otimes \mathbf{q}) = A(\bar{\mathbf{q}})A(\mathbf{q}) \quad (2.133)$$

Thus the quaternion representation of successive transformations is just the product of the quaternions of the constituent transformations, in the same way that the attitude matrix of the combined transformation is the product of the individual attitude matrices. A simple bilinear composition rule of this type holds only for the attitude matrix and quaternion representations, which is one of the reasons for the popularity of quaternions. With our  $\otimes$  quaternion product definition, the order of quaternion multiplication is identical to the order of matrix multiplication. The order would have been reversed if we had used the classical  $\odot$  definition of quaternion multiplication. The quaternion equivalent of Eq. (2.117) for successive rotations about the same axis follows from straightforward quaternion multiplication:

$$\mathbf{q}(\mathbf{e}, \vartheta) \otimes \mathbf{q}(\mathbf{e}, \varphi) = \mathbf{q}(\mathbf{e}, \varphi) \otimes \mathbf{q}(\mathbf{e}, \vartheta) = \mathbf{q}(\mathbf{e}, \vartheta + \varphi) \quad (2.134)$$

Unit quaternions reside on the three-dimensional unit sphere  $S^3$  embedded in four-dimensional quaternion space. Equation (2.124) shows that a rotation by  $720^\circ$ , but not a rotation by  $360^\circ$ , is equivalent to the identity transformation in quaternion space, because  $\mathbf{q}(\mathbf{e}, \vartheta + 4\pi) = \mathbf{q}(\mathbf{e}, \vartheta)$  but  $\mathbf{q}(\mathbf{e}, \vartheta + 2\pi) = -\mathbf{q}(\mathbf{e}, \vartheta)$ . The attitude matrix  $A(\mathbf{q})$  is a homogenous quadratic function of the elements of the quaternion, though, so  $\mathbf{q}$  and  $-\mathbf{q}$  give the same attitude matrix. This 2:1 mapping of quaternions to rotations is a minor annoyance that cannot be removed without introducing discontinuities like those that plague all three-parameter attitude representations. Because the quaternions  $\mathbf{q}$  and  $-\mathbf{q}$  are on opposite hemispheres of  $S^3$ , we could get a 1:1 mapping of quaternions to rotations by restricting the quaternions to one hemisphere, which is usually taken to be the hemisphere with  $q_4 \geq 0$ . This gives rise to the same problem as restricting the rotation vector to  $\vartheta \leq \pi$ , namely that a

smoothly varying quaternion can jump discontinuously from one side to the other of the equator bounding the hemisphere. Restricting the representation to positive  $q_4$  effectively gives a three-parameter representation with  $q_4 \equiv \sqrt{1 - \|\mathbf{q}_{1:3}\|^2}$ , so it is not surprising that it leads to the same problems as other three-parameter representations.

We finally turn to the problem of extracting a quaternion from an attitude matrix [10]. We construct four four-component vectors from the components of  $A$ :

$$\begin{aligned} \begin{bmatrix} 1 + 2A_{11} - \text{tr}A \\ A_{12} + A_{21} \\ A_{13} + A_{31} \\ A_{23} - A_{32} \end{bmatrix} &= 4q_1\mathbf{q}, & \begin{bmatrix} A_{21} + A_{12} \\ 1 + 2A_{22} - \text{tr}A \\ A_{23} + A_{32} \\ A_{31} - A_{13} \end{bmatrix} &= 4q_2\mathbf{q} \\ \begin{bmatrix} A_{31} + A_{13} \\ A_{32} + A_{23} \\ 1 + 2A_{33} - \text{tr}A \\ A_{12} - A_{21} \end{bmatrix} &= 4q_3\mathbf{q}, & \begin{bmatrix} A_{23} - A_{32} \\ A_{31} - A_{13} \\ A_{12} - A_{21} \\ 1 + \text{tr}A \end{bmatrix} &= 4q_4\mathbf{q} \end{aligned} \quad (2.135)$$

The quaternion can be found by normalizing any one of these four vectors. Numerical errors are minimized by choosing the vector with the greatest norm, which is the vector with the largest value of  $|q_i|$  on the right side. This can be found by the following procedure. Find the largest of  $\text{tr}A$  and  $A_{ii}$  for  $i = 1, 2, 3$ . If  $\text{tr}A$  is the largest of these, then  $|q_4|$  is the largest of the  $|q_i|$ , otherwise the largest value of  $|q_i|$  is the one with the same index as the largest  $A_{ii}$ . The overall sign of the normalized vector is not determined, reflecting the twofold ambiguity of the quaternion representation.

### 2.9.4 Rodrigues Parameter Representation

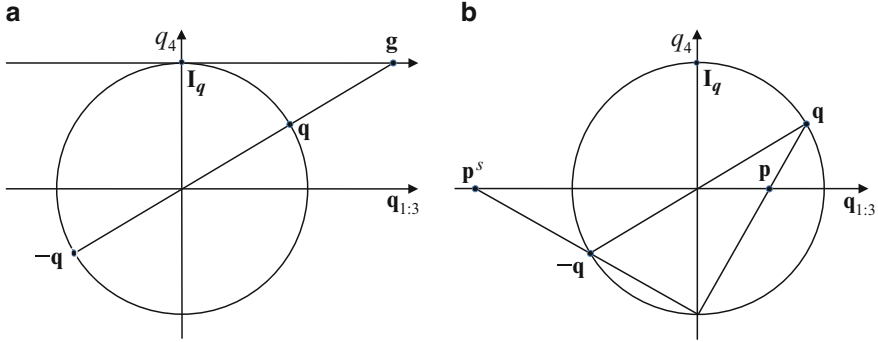
The three *Rodrigues parameters* made their appearance in Rodrigues' classic 1840 paper [14]. They were later represented as the “vector semitangent of version” by J. Willard Gibbs, who invented modern vector notation. For this reason, the vector of Rodrigues parameters is often called the *Gibbs vector* and denoted by  $\mathbf{g}$ . They are related to the quaternion by

$$\mathbf{g} = \frac{\mathbf{q}_{1:3}}{q_4} \quad (2.136)$$

which has the inverse

$$\mathbf{q} = \frac{\pm 1}{\sqrt{1 + \|\mathbf{g}\|^2}} \begin{bmatrix} \mathbf{g} \\ 1 \end{bmatrix} \quad (2.137)$$





**Fig. 2.7** Relationship of the Rodrigues parameters (a) and the modified Rodrigues parameters (b) to the quaternion

Using Eq. (2.124) to express the quaternion in terms of the Euler axis and angle gives

$$\mathbf{g}(\mathbf{e}, \vartheta) = \mathbf{e} \tan(\vartheta/2) \quad (2.138)$$

which explains Gibbs' peculiar terminology.

The mapping from quaternions to Rodrigues parameters is illustrated in Fig. 2.7a. The plane of the figure is the plane containing the origin,  $\mathbf{q}$ , and the identity quaternion  $\mathbf{I}_q$ . The circle is the cross-section of the quaternion sphere  $S^3$ , so it has unit radius. The vertical axis is the  $q_4$  axis, and the horizontal axis represents the three-dimensional  $\mathbf{q}_{1:3}$  hyperplane. The horizontal line passing through  $\mathbf{I}_q$  represents the three-dimensional Gibbs vector hyperplane, which is tangent to  $S^3$  at the point  $\mathbf{q} = \mathbf{I}_q$ . The Gibbs vector  $\mathbf{g}$  is the projection of the quaternion from the origin onto the Gibbs vector hyperplane. It is clear from the figure or from Eq. (2.136) that  $\mathbf{q}$  and  $-\mathbf{q}$  map to the same Gibbs vector, so the Rodrigues parameters provide a 1:1 mapping of rotations. The price paid for this is that the Gibbs vector is infinite for a  $180^\circ$  rotation. Thus this parameterization is not recommended as a global attitude representation, but it provides an excellent representation of small rotations.

Substituting Eq. (2.137) into Eq. (2.125) gives the Rodrigues parameter representation of the attitude matrix

$$\begin{aligned} A(\mathbf{g}) &= \frac{(1 - \|\mathbf{g}\|^2)I_3 - 2[\mathbf{g} \times] + 2\mathbf{g}\mathbf{g}^T}{1 + \|\mathbf{g}\|^2} = I_3 + 2\frac{[\mathbf{g} \times]^2 - [\mathbf{g} \times]}{1 + \|\mathbf{g}\|^2} \\ &= \frac{1}{1 + \|\mathbf{g}\|^2} \begin{bmatrix} 1 + g_1^2 - g_2^2 - g_3^2 & 2(g_1g_2 + g_3) & 2(g_1g_3 - g_2) \\ 2(g_2g_1 - g_3) & 1 - g_1^2 + g_2^2 - g_3^2 & 2(g_2g_3 + g_1) \\ 2(g_3g_1 + g_2) & 2(g_3g_2 - g_1) & 1 - g_1^2 - g_2^2 + g_3^2 \end{bmatrix} \end{aligned} \quad (2.139)$$

This is similar to the quaternion representation in requiring no transcendental function evaluations. However, it is a rational function of the Rodrigues parameters rather than a simple polynomial function.

The relationship between the Rodrigues parameters and the attitude matrix can also be expressed as a *Cayley transform*. The equations

$$A(\mathbf{g}) = (I_3 - [\mathbf{g} \times]) (I_3 + [\mathbf{g} \times])^{-1} = (I_3 + [\mathbf{g} \times])^{-1} (I_3 - [\mathbf{g} \times]) \quad (2.140)$$

can be verified by multiplying Eq. (2.139) by  $I_3 + [\mathbf{g} \times]$ . Because the matrices  $(I_3 - [\mathbf{g} \times])$  and  $(I_3 + [\mathbf{g} \times])^{-1}$  commute, the order of multiplication is irrelevant and the Cayley transform can be written as

$$A(\mathbf{g}) = \frac{I_3 - [\mathbf{g} \times]}{I_3 + [\mathbf{g} \times]} \quad (2.141)$$

This is not a useful form for computation, owing to the required matrix inversion; but it can be generalized to higher dimensions. Any  $n \times n$  proper orthogonal matrix  $M$  can be expressed in terms of an  $n \times n$  skew-symmetric matrix  $S$  by

$$D = \frac{I_n - S}{I_n + S} \quad (2.142)$$

A skew-symmetric  $n \times n$  matrix has  $n(n-1)/2$  free parameters, the correct number to parameterize an orthogonal  $n \times n$  matrix that must obey  $n(n+1)/2$  constraints on its  $n^2$  elements.

The rule for computing the Gibbs vector representing a composite of two rotations is easily derived from Eqs. (2.82), (2.136), and (2.137). The Gibbs vector corresponding to the quaternion product  $\tilde{\mathbf{q}} = \tilde{\mathbf{q}} \otimes \mathbf{q}$  is

$$\tilde{\mathbf{g}} = \frac{\tilde{\mathbf{g}} + \mathbf{g} - \tilde{\mathbf{g}} \times \mathbf{g}}{1 - \tilde{\mathbf{g}} \cdot \mathbf{g}} \quad (2.143)$$

This is not a bilinear function of the constituent Gibbs vectors, so it cannot be represented as a matrix product like quaternion composition.

Extracting the Rodrigues parameters from the attitude matrix is similar to extracting the quaternion using one of the four-component vectors of Eq. (2.135). Instead of normalizing one of those vectors, though, the Rodrigues parameters are found by dividing the first three components of the vector by the fourth component.

### 2.9.5 Modified Rodrigues Parameters

The modified Rodrigues parameters (MRPs) are the newest of the commonly-employed attitude representations. They were invented by T. F. Wiener in 1962 [26], rediscovered by Marandi and Modi in 1987 [9], and have been championed by Junkins and Schaub [16]. They are related to the quaternion by

$$\mathbf{p} = \frac{\mathbf{q}_{1:3}}{1 + q_4} \quad (2.144)$$

which has the inverse

$$\mathbf{q} = \frac{1}{1 + \|\mathbf{p}\|^2} \begin{bmatrix} 1 \\ 2\mathbf{p} \\ 1 - \|\mathbf{p}\|^2 \end{bmatrix} \quad (2.145)$$

Using Eq. (2.124) to express the quaternion in terms of the Euler axis and angle gives

$$\mathbf{p}(\mathbf{e}, \vartheta) = \mathbf{e} \tan(\vartheta/4) \quad (2.146)$$

It is easy to see that

$$\mathbf{p}(\mathbf{e}, \vartheta + 2\pi) = -\mathbf{e} \cot(\vartheta/4) = -\frac{\mathbf{p}(\mathbf{e}, \vartheta)}{\|\mathbf{p}(\mathbf{e}, \vartheta)\|^2} \quad (2.147)$$

but that  $\mathbf{p}(\mathbf{e}, \vartheta + 4\pi) = \mathbf{p}(\mathbf{e}, \vartheta)$ . Thus the MRP representation is 2:1 just like the quaternion representation.

The *shadow* set of MRPs

$$\mathbf{p}^S \equiv -\frac{\mathbf{p}}{\|\mathbf{p}\|^2} \quad (2.148)$$

represents the same attitude as  $\mathbf{p}$ , in the same way that  $\mathbf{q}$  and  $-\mathbf{q}$  represent the same attitude. These two MRP vectors are illustrated in Fig. 2.7b, which shows them as stereographic projections from the point  $\mathbf{q} = -\mathbf{I}_q$  onto the MRP hyperplane, which is coincident with the  $\mathbf{q}_{1:3}$  hyperplane. It is clear from the figure or from Eq. (2.148) that  $\|\mathbf{p}^S\| \geq 1$  if  $\|\mathbf{p}\| \leq 1$ . Thus, in following the variation of an attitude represented by MRPs, one can always keep the magnitude of the MRP vector from exceeding unity by switching to the shadow MRP when needed. The logic required for this is regarded by many practitioners to be less burdensome than carrying the fourth component of a quaternion and enforcing the quaternion norm constraint, leading them to prefer MRPs for numerical simulation of attitude motion. It is good practice to allow the MRP norm to exceed unity by some amount to avoid “chattering” between the MRP and its shadow in case the norm remains close to unity for an extended period.

The MRP representation of the attitude matrix can be found by substituting Eq. (2.145) into Eq. (2.125). It is easier to note that  $A(\mathbf{e}, \vartheta) = A^2(\mathbf{e}, \vartheta/2)$  from Eq. (2.117), so the MRP representation can be obtained by squaring the Gibbs vector representation for a half-angle rotation:

$$\begin{aligned} A(\mathbf{p}) &= \left( \frac{I_3 - [\mathbf{p} \times]}{I_3 + [\mathbf{p} \times]} \right)^2 = \left( I_3 + 2 \frac{[\mathbf{p} \times]^2 - [\mathbf{p} \times]}{1 + \|\mathbf{p}\|^2} \right)^2 \\ &= I_3 + \frac{8[\mathbf{p} \times]^2 - 4(1 - \|\mathbf{p}\|^2)[\mathbf{p} \times]}{(1 + \|\mathbf{p}\|^2)^2} \end{aligned} \quad (2.149)$$

The rule for computing the MRPs representing a composite of two rotations can be derived using Eqs. (2.82), (2.144), and (2.145). The MRP vector corresponding to the quaternion product  $\bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes \mathbf{q}$  is

$$\bar{\mathbf{p}} = \frac{(1 - \|\mathbf{p}\|^2)\bar{\mathbf{p}} + (1 - \|\bar{\mathbf{p}}\|^2)\mathbf{p} - 2\bar{\mathbf{p}} \times \mathbf{p}}{1 + \|\mathbf{p}\|^2\|\bar{\mathbf{p}}\|^2 - 2\bar{\mathbf{p}} \cdot \mathbf{p}} \quad (2.150)$$

This involves more computation than the composition of Gibbs vectors, but not an unreasonable burden.

The most convenient way to extract the MRPs from the attitude matrix is to first extract the quaternion and then compute the MRPs by Eq. (2.144).

Junkins and Schaub have developed a family of attitude representations intermediate between the Rodrigues parameters and the MRPs by choosing a projection point in quaternion space intermediate between the point  $\mathbf{q} = \mathbf{0}$  in Fig. 2.7a and the point  $\mathbf{q} = -\mathbf{I}_q$  in Fig. 2.7b [15]. Tsiotras, Junkins, and Schaub have investigated higher-order Cayley transforms

$$A(\mathbf{v}) = \left( \frac{I_3 - [\mathbf{v} \times]}{I_3 + [\mathbf{v} \times]} \right)^n \quad (2.151)$$

for  $n > 2$  [23]. Neither of these generalizations has found wide application, however.

### 2.9.6 Euler Angles

An *Euler angle* representation expresses a rotation from an initial frame  $I$  to a final frame  $F$  as the product of three rotations: a rotation first from  $I$  to an intermediate frame  $H$ , then to a second intermediate frame  $G$ , and finally to frame  $F$ . The frame indices are generally omitted to simplify the notation, but including them clarifies the form of the overall transformation

$$A_{FI}(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi) \equiv A_{FG}(\mathbf{e}_\psi, \psi) A_{GH}(\mathbf{e}_\theta, \theta) A_{HI}(\mathbf{e}_\phi, \phi) \quad (2.152)$$

The rotation axis vectors of the constituent rotations are constant column vectors, and their subscripts do not label frames explicitly. The column vector  $\mathbf{e}_\psi$  is a representation of a rotation axis in frames  $F$  and  $G$ ,  $\mathbf{e}_\theta$  is a representation in frames  $G$  and  $H$ , and  $\mathbf{e}_\phi$  is a representation in frames  $H$  and  $I$ . The rotation angles  $\phi$ ,  $\theta$ , and  $\psi$  are the variables used to specify the rotation. The possibility of employing Euler axis sequences with a wide choice of rotation axes was discovered by Davenport [2, 18], generalizing the classical applications of Euler angles that use a more restricted set of rotation axes. We will establish general results for Euler axis sequences using Davenport's general formulation, and then the classical axis sequences will follow as special cases.

We want to be able to represent any attitude matrix by this Euler angle sequence. In particular, we must be able to represent the attitude matrix that transforms  $\mathbf{e}_\phi$  into  $\mathbf{e}_\psi$ , which is to say

$$\mathbf{e}_\psi = A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi) \mathbf{e}_\phi = A(\mathbf{e}_\psi, \psi) A(\mathbf{e}_\theta, \theta) A(\mathbf{e}_\phi, \phi) \mathbf{e}_\phi \quad (2.153)$$

Multiplying on the left by  $A^T(\mathbf{e}_\psi, \psi)$  and noting that  $A^T(\mathbf{e}_\psi, \psi) \mathbf{e}_\psi = \mathbf{e}_\psi$  and  $A(\mathbf{e}_\phi, \phi) \mathbf{e}_\phi = \mathbf{e}_\phi$ , we see that there must be an angle  $\theta_0$  for which

$$\mathbf{e}_\psi = A(\mathbf{e}_\theta, \theta_0) \mathbf{e}_\phi \quad (2.154)$$

A little thought (or algebra) shows that this requires  $\mathbf{e}_\psi \cdot \mathbf{e}_\theta = \mathbf{e}_\phi \cdot \mathbf{e}_\theta$ . We must also be able to represent the attitude matrix that transforms  $\mathbf{e}_\phi$  into  $-\mathbf{e}_\psi$ , which means that there must be an angle  $\theta_1$  for which  $-\mathbf{e}_\psi = A(\mathbf{e}_\theta, \theta_1) \mathbf{e}_\phi$ . It follows from this that  $-\mathbf{e}_\psi \cdot \mathbf{e}_\theta = \mathbf{e}_\phi \cdot \mathbf{e}_\theta$ . These conditions can be satisfied simultaneously only if the rotation axis  $\mathbf{e}_\theta$  is perpendicular to both  $\mathbf{e}_\phi$  and  $\mathbf{e}_\psi$ , so that

$$\mathbf{e}_\psi \cdot \mathbf{e}_\theta = \mathbf{e}_\phi \cdot \mathbf{e}_\theta = 0 \quad (2.155)$$

This relation leads to a more complete description of the orientation of the physical rotation axis vectors. Axis  $\mathbf{e}_\phi$  is fixed in the initial reference frame,  $\mathbf{e}_\psi$  is fixed in the final reference frame, and  $\mathbf{e}_\theta$  is perpendicular to both  $\mathbf{e}_\phi$  and  $\mathbf{e}_\psi$ .

We have shown that Eq. (2.155) is a necessary condition for Eq. (2.152) to represent a general attitude. We will now show that it is a sufficient condition. Note that the angle  $\theta_0$  is not a variable, but is defined by the choice of rotation axes. Then using Eqs. (2.111b), (2.117) and (2.119) gives

$$\begin{aligned} A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi) &= A(A(\mathbf{e}_\theta, \theta_0) \mathbf{e}_\phi, \psi) A(\mathbf{e}_\theta, \theta) A(\mathbf{e}_\phi, \phi) \\ &= A(\mathbf{e}_\theta, \theta_0) A(\mathbf{e}_\phi, \psi) A(\mathbf{e}_\theta, \theta') A(\mathbf{e}_\phi, \phi) \\ &= A(\mathbf{e}_\theta, \theta_0) A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\phi; \phi, \theta', \psi) \end{aligned} \quad (2.156)$$

where  $\theta' \equiv \theta - \theta_0$ . As  $A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi)$  covers all of  $\text{SO}(3)$ , the product  $A^T(\mathbf{e}_\theta, \theta_0) A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi)$  also covers  $\text{SO}(3)$ , so we only need to show that  $A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\phi; \phi, \theta', \psi)$  can represent any attitude, or equivalently to show that this rotation can transform the orthonormal basis  $\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\phi \times \mathbf{e}_\theta$  into any other orthonormal basis. It is sufficient to show that  $\mathbf{e}'_\phi \equiv A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\phi; \phi, \theta', \psi) \mathbf{e}_\phi$  can be any unit vector and  $\mathbf{e}'_\theta \equiv A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\phi; \phi, \theta', \psi) \mathbf{e}_\theta$  can be any unit vector perpendicular to  $\mathbf{e}'_\phi$ , because Eq. (2.65) then implies that  $\mathbf{e}_\phi \times \mathbf{e}_\theta$  will transform into  $\mathbf{e}'_\phi \times \mathbf{e}'_\theta$ , completing the orthonormal triad. Now

$$\begin{aligned} \mathbf{e}'_\phi &= A(\mathbf{e}_\phi, \psi) A(\mathbf{e}_\theta, \theta') A(\mathbf{e}_\phi, \phi) \mathbf{e}_\phi = A(\mathbf{e}_\phi, \psi) A(\mathbf{e}_\theta, \theta') \mathbf{e}_\phi \\ &= \cos \theta' \mathbf{e}_\phi + \sin \theta' \sin \psi \mathbf{e}_\theta + \sin \theta' \cos \psi (\mathbf{e}_\phi \times \mathbf{e}_\theta) \end{aligned} \quad (2.157)$$

and we can find values of  $\theta'$  and  $\psi$  to make this equal any desired unit vector. Furthermore, because  $A(\mathbf{e}_\phi, \phi)\mathbf{e}_\theta = \cos \phi \mathbf{e}_\theta - \sin \phi (\mathbf{e}_\phi \times \mathbf{e}_\theta)$ , we have

$$\mathbf{e}'_\theta = \cos \phi A(\mathbf{e}_\phi, \psi)A(\mathbf{e}_\theta, \theta')\mathbf{e}_\theta - \sin \phi A(\mathbf{e}_\phi, \psi)A(\mathbf{e}_\theta, \theta')(\mathbf{e}_\phi \times \mathbf{e}_\theta) \quad (2.158)$$

Rotations preserve orthogonality, so this superposition can represent any unit vector in the plane perpendicular to  $\mathbf{e}'_\phi$ , completing the proof that any Euler angle sequence obeying Eq. (2.155) is sufficient to represent any attitude matrix.

An Euler angle parameterization has a twofold ambiguity in addition to the usual  $2\pi$  ambiguity in specifying any angle. To see this ambiguity, insert the product  $A(\mathbf{e}_\phi, -\pi)A(\mathbf{e}_\phi, \pi)$  before and after  $A(\mathbf{e}_\theta, \theta')$  in the second line of Eq. (2.156) and use some axis/angle representation identities to get

$$\begin{aligned} A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi) &= A(\mathbf{e}_\theta, \theta_0)A(\mathbf{e}_\phi, \psi - \pi)A(A(\mathbf{e}_\phi, \pi)\mathbf{e}_\theta, \theta - \theta_0)A(\mathbf{e}_\phi, \phi + \pi) \\ &= A(\mathbf{e}_\theta, \theta_0)A(\mathbf{e}_\phi, \psi - \pi)A(-\mathbf{e}_\theta, \theta - \theta_0)A(\mathbf{e}_\phi, \phi + \pi) \\ &= A(\mathbf{e}_\theta, \theta_0)A(\mathbf{e}_\phi, \psi - \pi)A(\mathbf{e}_\theta, -\theta_0)A(\mathbf{e}_\theta, 2\theta_0 - \theta)A(\mathbf{e}_\phi, \phi + \pi) \\ &= A(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi + \pi, 2\theta_0 - \theta, \psi - \pi) \end{aligned} \quad (2.159)$$

The rotation axes of the classical Euler angle representation are selected from the set

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (2.160)$$

and a more compact notation is used for these representations

$$A_{ijk}(\phi, \theta, \psi) = A(\mathbf{e}_k, \psi)A(\mathbf{e}_j, \theta)A(\mathbf{e}_i, \phi) \quad (2.161)$$

The possible choice of axes is constrained by the requirements  $i \neq j$  and  $j \neq k$ , as required by Eq. (2.155). This leaves us with six symmetric sets of Euler parameters, with  $i-j-k$  equal to:

$$1-2-1, 1-3-1, 2-3-2, 2-1-2, 3-1-3, \text{ and } 3-2-3$$

and six asymmetric sets:

$$1-2-3, 1-3-2, 2-3-1, 2-1-3, 3-1-2, \text{ and } 3-2-1$$

The explicit forms of the attitude matrices for all 12 sets are collected in Appendix B.

As a specific example of the symmetric sequences, we consider the 3-1-3 sequence, which is often used for analytical treatments of rigid body motion and for representing the attitude of spinning spacecraft:

$$\begin{aligned}
A_{313}(\phi, \theta, \psi) &= A(\mathbf{e}_3, \psi)A(\mathbf{e}_1, \theta)A(\mathbf{e}_3, \phi) \\
&= \begin{bmatrix} c\psi & s\psi & 0 \\ -s\psi & c\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\theta & s\theta \\ 0 & -s\theta & c\theta \end{bmatrix} \begin{bmatrix} c\phi & s\phi & 0 \\ -s\phi & c\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} c\psi c\phi - s\psi c\theta s\phi & c\psi s\phi + s\psi c\theta c\phi & s\psi s\theta \\ -s\psi c\phi - c\psi c\theta s\phi & -s\psi s\phi + c\psi c\theta c\phi & c\psi s\theta \\ s\theta s\phi & -s\theta c\phi & c\theta \end{bmatrix} \quad (2.162)
\end{aligned}$$

We have written  $c\psi \equiv \cos \psi$ ,  $s\psi \equiv \sin \psi$ , and analogous equations for  $\theta$  and  $\phi$ . The symmetric Euler angle sets all have  $\theta_0 = 0$ , so the angle ambiguity relation is

$$A_{313}(\phi, \theta, \psi) = A_{313}(\phi + \pi, -\theta, \psi - \pi) \quad (2.163)$$

The asymmetric sets are often called the *Tait-Bryan* angles, although this terminology has been called into question [4]. The three angles in an asymmetric Euler angle sequence are often referred to as *roll*, *pitch*, and *yaw*. This terminology originally described the motions of ships and then was carried over into aircraft and spacecraft. Roll is a rotation about the vehicle body axis that is closest to the vehicle's usual direction of motion, and hence would be perceived as a screwing motion. The roll axis is conventionally assigned index 1. Yaw is a rotation about the vehicle body axis that is usually closest to the direction of local gravity, and hence would be often be perceived as a motion that points the vehicle left or right. The yaw axis is conventionally assigned index 3. Pitch is a rotation about the remaining vehicle body axis, and hence would often be perceived as a motion that points the vehicle up or down. The pitch axis is conventionally assigned index 2. Note that this associates the terms roll, pitch, and yaw with the vehicle axes, while Eq. (2.161) assigns the variables  $\phi$ ,  $\theta$ , and  $\psi$  based on the order of rotations in the sequence rather than on the axis indices. Thus there is no definite association between the variables  $\phi$ ,  $\theta$ , and  $\psi$  and the axis labels 1, 2, and 3 or the names roll, pitch and yaw. A different convention is followed by many authors who denote roll by  $\phi$ , pitch by  $\theta$ , and yaw by  $\psi$ . As always, the reader consulting any source should be careful to understand the conventions that it follows.

The 3–2–1 sequence, which is often used to describe the attitude of an Earth-pointing spacecraft, is a specific example of an asymmetric sequence.

$$\begin{aligned}
A_{321}(\phi, \theta, \psi) &= A(\mathbf{e}_1, \psi)A(\mathbf{e}_2, \theta)A(\mathbf{e}_3, \phi) \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\psi & s\psi \\ 0 & -s\psi & c\psi \end{bmatrix} \begin{bmatrix} c\theta & 0 & -s\theta \\ 0 & 1 & 0 \\ s\theta & 0 & c\theta \end{bmatrix} \begin{bmatrix} c\phi & s\phi & 0 \\ -s\phi & c\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} c\theta c\phi & c\theta s\phi & -s\theta \\ -c\psi s\phi + s\psi s\theta c\phi & c\psi c\phi + s\psi s\theta s\phi & s\psi c\theta \\ s\psi s\phi + c\psi s\theta c\phi & -s\psi c\phi + c\psi s\theta s\phi & c\psi c\theta \end{bmatrix} \quad (2.164)$$

Equation (2.154) shows that the angle  $\theta_0$  for this sequence is equal to  $-\pi/2$ , so the angle ambiguity relation is

$$A_{321}(\phi, \theta, \psi) = A_{321}(\phi + \pi, \pi - \theta, \psi - \pi) \quad (2.165)$$

In fact, all the Tait-Bryan angle axis sets have  $\theta_0 = \pm\pi/2$  so they all obey Eq. (2.165), taking into account the  $2\pi$  ambiguity in the definition of  $\theta$ . The Tait-Bryan representations have useful small-angle approximations, as will be shown in Sect. 2.10, but the small angle limits of the symmetric Euler angles are not as useful.

Separate procedures for the different axis sequences are generally used to find the Euler angles representing a given attitude matrix. We will consider the 3–1–3 and 3–2–1 sequences as specific examples, and will follow these examples with a general algorithm.

For the symmetric 3–1–3 sequence, the angle  $\theta$  is computed from the 33 element of  $A_{313}$ :

$$\theta = \cos^{-1}([A_{313}]_{33}) \quad (2.166)$$

Unless  $[A_{313}]_{33}$  has magnitude unity, two distinct values of  $\theta$  have the same cosine, corresponding to the two possible signs for  $\sin \theta$  and to the twofold ambiguity shown in Eq. (2.163). We are free to select either of these values, but we can avoid the twofold ambiguity by computing  $\theta$  as the principal value of the inverse cosine, which restricts its range to  $0 \leq \theta \leq \pi$  and gives  $\sin \theta \geq 0$ . If  $\sin \theta \neq 0$ , the other two Euler angles can be determined, modulo  $2\pi$ , by

$$\phi = \text{atan2}(\sigma[A_{313}]_{31}, -\sigma[A_{313}]_{32}) \quad (2.167a)$$

$$\psi = \text{atan2}(\sigma[A_{313}]_{13}, \sigma[A_{313}]_{23}) \quad (2.167b)$$

where  $\sigma = \pm 1$  is the sign of  $\sin \theta$  and  $\text{atan2}(y, x)$  is the standard function giving the argument of the complex number  $x + iy$ .

It is obvious that  $\phi$  and  $\psi$  cannot be determined from Eqs. (2.167) if  $\theta$  is equal to 0 or  $\pi$ , since these values give  $\sin \theta = 0$ . The usual twofold ambiguity is absent in these cases, but the attitude matrix takes the form

$$A_{313}(\phi, (\pi \mp \pi)/2, \psi) = \begin{bmatrix} \cos(\phi \pm \psi) & \sin(\phi \pm \psi) & 0 \\ \mp \sin(\phi \pm \psi) & \pm \cos(\phi \pm \psi) & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \quad (2.168)$$

It can be seen that only the sum or difference  $\phi \pm \psi$  is determined, and not the angles individually. This is known as *gimbal lock*, for reasons that will become



apparent when we discuss the kinematics of rotations. Gimbal lock is caused by collinearity of the *physical* rotation axis vectors of the first and third rotations in the sequence. Note that the column vector representations of the rotation axes are always parallel for the symmetric Euler angle sequences, but that does not cause gimbal lock. Gimbal lock occurs for all the symmetric Euler angle sequences when  $\sin \theta = 0$ .

We could employ a special algorithm for the  $\sin \theta = 0$  case, but it is preferable to develop a general algorithm for all values of  $\theta$ , because we could encounter a loss of precision for small but nonzero values of  $\sin \theta$ . To this end, we note that

$$[A_{313}]_{11} \pm [A_{313}]_{22} = (1 \pm c\theta) \cos(\phi \pm \psi) \quad (2.169a)$$

$$[A_{313}]_{12} \mp [A_{313}]_{21} = (1 \pm c\theta) \sin(\phi \pm \psi) \quad (2.169b)$$

and that  $(1 \pm c\theta)$  is positive if  $\cos \theta \neq \mp 1$ . Therefore, we find either  $\phi$  or  $\psi$ , but not both, from Eq. (2.167). If  $\sin \theta = 0$  we can set one of these angles to any convenient value. Then we find the other angle from their sum or difference by

$$\phi + \psi = \text{atan2}([A_{313}]_{12} - [A_{313}]_{21}, [A_{313}]_{11} + [A_{313}]_{22}) \quad \text{if } [A_{313}]_{33} \geq 0 \quad (2.170a)$$

$$\phi - \psi = \text{atan2}([A_{313}]_{12} + [A_{313}]_{21}, [A_{313}]_{11} - [A_{313}]_{22}) \quad \text{if } [A_{313}]_{33} < 0 \quad (2.170b)$$

Extraction of the asymmetric Euler angles proceeds in a similar manner. For the 3–2–1 sequence, the angle  $\theta$  is computed from the 13 element of  $A_{321}$ :

$$\theta = \sin^{-1}(-[A_{321}]_{13}) \quad (2.171)$$

Unless  $[A_{321}]_{13}$  has magnitude unity, two distinct values of  $\theta$  have the same sine, corresponding to the two possible signs for  $\cos \theta$  and to the twofold ambiguity shown in Eq. (2.165). We are free to select either value, but can avoid the ambiguity by computing  $\theta$  as the principal value of the inverse sine, restricting its range to  $|\theta| \leq \pi/2$  and giving  $\cos \theta \geq 0$ . If  $\cos \theta \neq 0$ , the other two angles can be determined by

$$\phi = \text{atan2}(\sigma'[A_{321}]_{12}, \sigma'[A_{321}]_{11}) \quad (2.172a)$$

$$\psi = \text{atan2}(\sigma'[A_{321}]_{23}, \sigma'[A_{321}]_{33}) \quad (2.172b)$$

where  $\sigma' = \pm 1$  is the sign of  $\cos \theta$ .

Gimbal lock for the asymmetric Euler, or Tait-Bryan, sequences occurs when  $\cos \theta = 0$ , i.e. when  $\theta = \mp\pi/2$ . Then the usual twofold ambiguity is absent, and the attitude matrix for the 3–2–1 example has the form

$$A_{321}(\phi, \mp\pi/2, \psi) = \begin{bmatrix} 0 & 0 & \pm 1 \\ -\sin(\phi \pm \psi) & \cos(\phi \pm \psi) & 0 \\ \mp \cos(\phi \pm \psi) & \mp \sin(\phi \pm \psi) & 0 \end{bmatrix} \quad (2.173)$$

Only the sum or difference  $\phi \pm \psi$  is determined, just as for the symmetric Euler axis sequence, again due to collinearity of the physical rotation axis vectors of the first and third rotations. The column vector representations of the rotation axes are always perpendicular for the asymmetric sequences, but that does not prevent gimbal lock from occurring.

In parallel with the 3–1–3 case, we deal with gimbal lock by finding either  $\phi$  or  $\psi$  from Eq. (2.172) and then the other from their sum or difference by

$$\phi + \psi = \text{atan2}(-[A_{321}]_{32} - [A_{321}]_{21}, [A_{321}]_{22} - [A_{321}]_{31}) \quad \text{if } [A_{321}]_{13} \geq 0 \quad (2.174a)$$

$$\phi - \psi = \text{atan2}([A_{321}]_{32} - [A_{321}]_{21}, [A_{321}]_{22} + [A_{321}]_{31}) \quad \text{if } [A_{321}]_{13} < 0 \quad (2.174b)$$

We finally discuss the extraction of the angles for Davenport's general axis sequences. We will accomplish this in a way that uses the results obtained above for extracting the 3–1–3 Euler angles. This technique can be applied to any sequence of conventional Euler or Tait-Bryan angles, as well as to the general Davenport angles, by selecting the rotation axes from the set of coordinate axes  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  [19].

We note that the proper orthogonal matrix

$$C = \begin{bmatrix} \mathbf{e}_\theta^T \\ (\mathbf{e}_\phi \times \mathbf{e}_\theta)^T \\ \mathbf{e}_\phi^T \end{bmatrix} \quad (2.175)$$

has the property that  $C\mathbf{e}_\theta = \mathbf{e}_1$  and  $C\mathbf{e}_\phi = \mathbf{e}_3$ . Then we have from Eq. (2.156)

$$\begin{aligned} CA(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi)C^T &= CA(\mathbf{e}_\theta, \theta_0)C^T CA(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta', \psi)C^T \\ &= A(\mathbf{e}_1, \theta_0)CA(\mathbf{e}_\phi, \psi)C^T C(\mathbf{e}_\theta, \theta')C^T CA(\mathbf{e}_\phi, \phi)C^T \\ &= A(\mathbf{e}_1, \theta_0)A_{313}(\phi, \theta', \psi) \end{aligned} \quad (2.176)$$

Thus we can extract  $\phi$ ,  $\theta'$ , and  $\psi$  from  $A^T(\mathbf{e}_1, \theta_0)CA(\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_\psi; \phi, \theta, \psi)C^T$  by the standard technique for the 3–1–3 sequence, and then compute  $\theta = \theta' + \theta_0$ . An easily-verified special case of Eq. (2.176) is

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} A_{321}(\phi, \theta, \psi) \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A_{313}(\phi, \theta + \pi/2, \psi) \quad (2.177)$$

Explicit expressions for the Euler angles resulting from successive transformations have been found in some special cases [17], but they have not been widely applied in practice.

## 2.10 Attitude Error Representations

The attitude matrix represents the rotation  $A_{BR}$  from some reference frame  $R$  to the spacecraft body frame  $B$ . Attitude estimation errors can be represented either as a small rotation  $A_{\hat{R}R}$  between  $R$  and an estimated reference frame  $\hat{R}$

$$A_{BR} = A_{B\hat{R}}A_{\hat{R}R} \quad (2.178)$$

or more commonly as a small rotation  $A_{B\hat{B}}$  between  $B$  and an estimated body frame  $\hat{B}$

$$A_{BR} = A_{B\hat{B}}A_{\hat{B}R} \quad (2.179)$$

The estimated attitude is represented by  $A_{B\hat{R}}$  in the former case and by  $A_{\hat{B}R}$  in the latter. In either case the matrix representing the errors,  $A_{\hat{R}R}$  or  $A_{B\hat{B}}$ , is expected to be close to the identity matrix.

The most natural representation of attitude errors is in terms of the rotation vector, Eq. (2.123), and its small-angle approximation

$$A(\delta\boldsymbol{\vartheta}) = \exp(-[\delta\boldsymbol{\vartheta} \times]) \approx I_3 - [\delta\boldsymbol{\vartheta} \times] + \frac{1}{2}[\delta\boldsymbol{\vartheta} \times]^2 \quad (2.180)$$

Other attitude parameterizations can be used to represent attitude errors, such as the quaternion, Eq. (2.125),

$$A(\delta\mathbf{q}) \approx I_3 - 2[\delta\mathbf{q}_{1:3} \times] + 2[\delta\mathbf{q}_{1:3} \times]^2 \quad (2.181)$$

the Gibbs vector, Eq. (2.139),

$$A(\delta\mathbf{g}) \approx I_3 - 2[\delta\mathbf{g} \times] + 2[\delta\mathbf{g} \times]^2 \quad (2.182)$$

or the MRPs, Eq. (2.149),

$$A(\delta\mathbf{p}) \approx I_3 - 4[\delta\mathbf{p} \times] + 8[\delta\mathbf{p} \times]^2 \quad (2.183)$$

It is notable that these representations are all equivalent through second order in the errors with the identification

$$\delta\boldsymbol{\vartheta} = 2\delta\mathbf{q}_{1:3} = 2\delta\mathbf{g} = 4\delta\mathbf{p} \quad (2.184)$$

In fact, only the first-order approximation is required for most applications. It must be emphasized that Eqs. (2.180)–(2.184) are only true in the (very useful) approximation of small error angles. Attitude error representations are often used

for errors that are not especially small. In that case, the quaternion, Gibbs vector, or MRP representation is often preferred to the rotation vector for computational convenience.

The small-angle approximation of the Tait-Bryan or asymmetric Euler angle representation, Eq. (2.161), is to second order

$$\begin{aligned}
 A_{ijk}(\delta\phi, \delta\theta, \delta\psi) &\approx \left( I_3 - \delta\psi [\mathbf{e}_k \times] + \frac{1}{2} \delta\psi^2 [\mathbf{e}_k \times]^2 \right) \\
 &\quad \times \left( I_3 - \delta\theta [\mathbf{e}_j \times] + \frac{1}{2} \delta\theta^2 [\mathbf{e}_j \times]^2 \right) \left( I_3 - \delta\phi [\mathbf{e}_i \times] + \frac{1}{2} \delta\phi^2 [\mathbf{e}_i \times]^2 \right) \\
 &\approx I_3 - [\delta\boldsymbol{\vartheta}_1 \times] + \frac{1}{2} [\delta\boldsymbol{\vartheta}_1 \times]^2 \mp \frac{1}{2} [(\delta\theta \delta\psi \mathbf{e}_i - \delta\psi \delta\phi \mathbf{e}_j + \delta\phi \delta\theta \mathbf{e}_k) \times]
 \end{aligned} \tag{2.185}$$

where

$$\delta\boldsymbol{\vartheta}_1 \equiv \delta\phi \mathbf{e}_i + \delta\theta \mathbf{e}_j + \delta\psi \mathbf{e}_k \tag{2.186}$$

The upper sign in the last term in Eq. (2.185), which was derived with the use of Eq. (2.56c), holds if  $\{i, j, k\}$  is an even permutation of  $\{1, 2, 3\}$ , and the lower sign applies if it is an odd permutation. With the identification  $\delta\boldsymbol{\vartheta} = \delta\boldsymbol{\vartheta}_1$  Eq. (2.185) agrees with Eq. (2.180) to first order, but not to second order. Agreement to second order would require

$$\delta\boldsymbol{\vartheta} = \left( \delta\phi \pm \frac{\delta\theta \delta\psi}{2} \right) \mathbf{e}_i + \left( \delta\theta \mp \frac{\delta\psi \delta\phi}{2} \right) \mathbf{e}_j + \left( \delta\psi \pm \frac{\delta\phi \delta\theta}{2} \right) \mathbf{e}_k \tag{2.187}$$

This identification has never been used to our knowledge, however, because first-order approximations in the errors are generally adequate.

## Problems

**2.1.** Prove that any real  $n \times n$  matrix  $A$  can be decomposed into the sum of a symmetric and skew symmetric matrix. Hint: the matrix  $A - A^T$  is clearly skew symmetric.

**2.2.** It is known that a general  $n \times n$  symmetric matrix  $A$  must possess  $n$  mutually orthogonal eigenvectors, even if some of the eigenvalues are repeated. Here, you will prove that this is true when the eigenvalues are not repeated. Let  $A \mathbf{x}_i = \lambda_i \mathbf{x}_i$  for  $i = 1, 2$ , where  $\lambda_i$  is the  $i$ th eigenvalue and  $\mathbf{x}_i$  is the  $i$ th eigenvector. Assume that  $\lambda_1 \neq \lambda_2$ . Start by taking the transpose of  $A \mathbf{x}_1 = \lambda_1 \mathbf{x}_1$  and right multiplying both sides by  $\mathbf{x}_2$ , and prove that when  $A = A^T$  the vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  must be orthogonal.

**2.3.** Suppose that the characteristic equation of an  $n \times n$  matrix  $A$  is given by

$$\Delta(\lambda) = \det(\lambda I_n - A) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n$$

The Cayley-Hamilton theorem states that the matrix  $A$  obeys its characteristic equation so that

$$A^n + a_1 A^{n-1} + \cdots + A + a_n I_n = 0$$

Multiplying by  $A$  gives

$$A^{n+1} + a_1 A^n + \cdots + A^2 + a_n A = 0$$

This implies that  $A^{n+1}$  can be written as a linear combination of  $A, A^2, \dots, A^n$ , which in turn can be written as a linear combination of  $I_n, A, \dots, A^{n-1}$ . In fact, any  $A^m$  with  $m > n - 1$  can be written using the same linear combination. Use this fact to compute a closed-form expression for the following:

$$\begin{bmatrix} 1 & 0 & 4 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}^{100}$$

Note that  $f(A)$  is the same linear combination of powers of  $A$  as  $f(J)$  is of powers of  $J$ , where  $J$  is a diagonal matrix of the eigenvalues of  $A$ .

**2.4.** Consider the following  $2 \times 2$  matrix for real  $a, b$  and  $d$ :

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

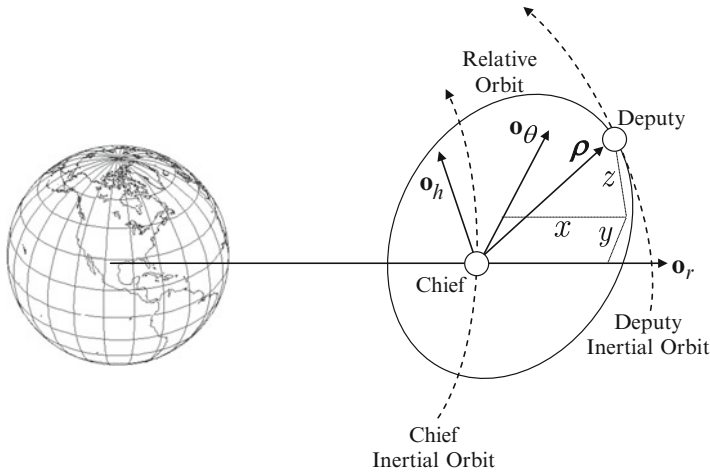
Determine the eigenvalues  $\lambda_1$  and  $\lambda_2$  of  $A$  in terms of  $a, b$  and  $d$ . Since  $A$  is symmetric then the eigenvalue/eigenvector decomposition gives

$$V^T A V = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

where  $V$  is an orthogonal matrix. Suppose that  $V$  is given by the form

$$V = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

where  $s \equiv \sin \theta$  and  $c \equiv \cos \theta$  for some angle  $\theta$ . Find  $c$  and  $s$  in terms of  $a, b$  and  $d$ .



**Fig. 2.8** Hill frame

**2.5.** A matrix that is used to reflect an object over a line or plane is called a reflection matrix. Consider the following natural basis, given by Eq. (2.35), for  $n = 2$ :

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Let  $\phi$  be the counterclockwise angle between some line  $\ell$  through the origin and the  $x$ -axis. Consider the following matrix that rotates the  $x$ -axis onto  $\ell$ :

$$A = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$$

Noting that  $A\mathbf{e}_1$  lies on  $\ell$  and  $A\mathbf{e}_1$  is perpendicular to  $\ell$ , determine the reflection matrix that sends any vector not on  $\ell$  to its mirror image about  $\ell$ . Provide simulation plots for various angles  $\phi$  and any chosen  $2 \times 1$  vector to reflect.

**2.6.** Write a computer program that takes some latitude, longitude and height, and converts these quantities to ECEF position using Eq. (2.74). Also, write a computer program that takes ECEF position and converts it to latitude, longitude and height using Eq. (2.77). Pick some latitude, longitude and height. Then compute the ECEF position and convert this position to latitude, longitude and height to ensure that the original quantities are obtained.

**2.7.** A useful frame for formation flying applications is the “Hill frame” shown in Fig. 2.8 [6]. The frame is given by  $\{\mathbf{o}_r, \mathbf{o}_\theta, \mathbf{o}_h\}$ , where  $\mathbf{o}_r$  points in the chief spacecraft’s radial direction,  $\mathbf{o}_h$  is along the chief orbit momentum vector and  $\mathbf{o}_\theta$

completes the right-handed coordinate system, so that

$$\mathbf{o}_r = \frac{\mathbf{r}_I}{\|\mathbf{r}_I\|}, \quad \mathbf{o}_\theta = \mathbf{o}_h \times \mathbf{o}_r, \quad \mathbf{o}_h = \frac{\mathbf{h}_I}{\|\mathbf{h}_I\|}$$

where  $\mathbf{h}_I = \mathbf{r}_I \times \mathbf{v}_I$ . Using the dot product approach similar to Eq. (2.43), fully derive the attitude matrix that rotates vectors from the Hill frame to an inertial frame. Also, determine the relationship between the Hill frame and the LVLH frame.

**2.8.** This is an alternative proof of Euler's Theorem that avoids eigenvalues [12].

- a) First assume that the attitude matrix is not symmetric, i.e. that  $A \neq A^T$ . Define the skew symmetric matrix  $S \equiv \frac{1}{2}(A - A^T) = [\mathbf{s} \times]$ . Prove that  $A S A^T = S$ . Next show that  $\mathbf{d} \equiv \mathbf{s}/\|\mathbf{s}\|$  is equivalent to the Euler axis  $\mathbf{e}$ , to within a sign ambiguity,
- b) Now assume that  $A = A^T$ . Why does the above argument fail in this case? If  $A$  is symmetric the orthogonality relation is  $A^2 = I_3$ . Show that this gives  $A(A + I_3) = A + I_3$ , so all the columns of  $A + I_3$  are unchanged by  $A$ . Complete the proof by showing that at least one of these columns is not a column of zeros, and is therefore equivalent to the Euler axis  $\mathbf{e}$ , to within a sign ambiguity. What Euler angles of rotation,  $\vartheta$ , correspond to the case of  $A = A^T$ ?

**2.9.** The generalized Rodrigues parameters [15] (GRPs) can be written as

$$\boldsymbol{\rho} = \frac{f \mathbf{q}_{1:3}}{a + q_4}$$

where  $a$  is a parameter from 0 to 1 and  $f$  is a scale factor.

- a) Draw a plot of the rotation angle  $\vartheta$  versus  $a$  that causes the GPRs to become singular.
- b) Determine the inverse transformation for  $\mathbf{q}_{1:3}$  and  $q_4$  in terms of  $\boldsymbol{\rho}$ ,  $a$  and  $f$ . Note that your answer for  $q_4$  may seem to provide two solutions, but only one of them is correct.
- c) Determine  $f$  in terms of  $a$  so that the small angle approximation gives  $\|\boldsymbol{\rho}\| \approx \vartheta$ .

**2.10.** Show that the square of the quaternion elements can be extracted from the attitude matrix by using the following equations:

$$q_1^2 = \frac{1}{4}(1 + a_{11} - a_{22} - a_{33})$$

$$q_2^2 = \frac{1}{4}(1 + a_{22} - a_{11} - a_{33})$$

$$q_3^2 = \frac{1}{4}(1 + a_{33} - a_{11} - a_{22})$$

$$q_4^2 = \frac{1}{4}(1 + a_{11} + a_{22} + a_{33})$$

**2.11.** Derive the attitude matrix for a  $1-2-3$  sequence. Also, explicitly compute the determinant of this matrix to show that it is  $+1$ .

**2.12.** Derive the direct relationship from Euler angles to quaternion for a  $1-2-3$  sequence and for a  $1-2-1$  sequence. Compare your results to the ones shown in Table B.5 to make sure they are equivalent.

## References

1. Bate, R.R., Mueller, D.D., White, J.E.: *Fundamentals of Astrodynamics*. Dover Publications, New York (1971)
2. Davenport, P.B.: Rotations about nonorthogonal axes. *AIAA J.* **11**(6), 853–857 (1973)
3. Farrell, J., Barth, M.: *The Global Positioning System & Inertial Navigation*. McGraw-Hill, New York (1998)
4. Fraiture, L.: A history of the description of the three-dimensional finite rotation. *J. Astronaut. Sci.* **57**(1/2), 207–232 (2009)
5. Goldstein, H.: *Classical Mechanics*, 2nd edn. Addison-Wesley, Reading (1980)
6. Hill, G.W.: Researches in lunar theory. *Am. J. Math.* **1**(1), 5–26 (1878)
7. Lefferts, E.J., Markley, F.L., Shuster, M.D.: Kalman filtering for spacecraft attitude estimation. *J. Guid. Contr. Dynam.* **5**(5), 417–429 (1982)
8. Ma, C., Arias, E.F., Eubanks, T.M., Frey, A.L., Gontier, A.M., Jacobs, C.S., Sovers, O.J., Archinal, B.A., Charlot, P.: The International Celestial Reference Frame as realized by very long baseline interferometry. *Astron. J.* **116**, 516–546 (1998)
9. Marandi, S.R., Modi, V.J.: A preferred coordinate system and the associated orientation representation in attitude dynamics. *Acta Astronautica* **15**(11), 833–843 (1987)
10. Markley, F.L.: Unit quaternion from rotation matrix. *J. Guid. Contr. Dynam.* **31**(2), 440–442 (2008)
11. Montenbruck, O., Gill, E.: *Satellite Orbits: Models, Methods, and Applications*. Springer, Berlin Heidelberg New York (2000)
12. Palais, B., Palais, R.: Euler's fixed point theorem: The axis of a rotation. *J. Fixed point Theory Appl.* **2**(2), 215–220 (2007)
13. Pittelkau, M.E.: An analysis of the quaternion attitude determination filter. *J. Astronaut. Sci.* **51**(1) (2003)
14. Rodrigues, O.: Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire. *J. de Mathématiques Pures et Appliquées* **5**, 380–440 (1840)
15. Schaub, H., Junkins, J.L.: Stereographic orientation parameters for attitude dynamics: A generalization of the Rodrigues parameters. *J. Astronaut. Sci.* **44**(1), 1–19 (1996)
16. Schaub, H., Junkins, J.L.: *Analytical Mechanics of Aerospace Systems*, 2nd edn. American Institute of Aeronautics and Astronautics, New York (2009)
17. Shuster, M.D.: A survey of attitude representations. *J. Astronaut. Sci.* **41**(4), 439–517 (1993)
18. Shuster, M.D., Markley, F.L.: Generalization of the Euler angles. *J. Astronaut. Sci.* **51**(2), 123–132 (2003)
19. Shuster, M.D., Markley, F.L.: General formula for extracting the Euler angles. *J. Guid. Contr. Dynam.* **29**(1), 215–217 (2006)
20. Sofair, I.: Improved method for calculating exact geodetic latitude and altitude revisited. *J. Guid. Contr. Dynam.* **23**(2), 369 (2000)
21. Stewart, G.W.: *Introduction to Matrix Computations*. Academic Press, New York (1973)



22. Stuelpnagel, J.: On the parametrization of the three-dimensional rotation group. *SIAM Rev.* **6**(4), 422–430 (1964)
23. Tsiotras, P., Junkins, J.L., Schaub, H.: Higher-order Cayley transforms with applications to attitude representations. *J. Guid. Contr. Dynam.* **20**(3), 528–534 (1997)
24. Vallado, D.A.: *Fundamentals of Astrodynamics and Applications*, 3rd edn. Microcosm Press, Hawthorne and Springer, New York (2007)
25. Weinberg, S.: *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley, Chichester (1972)
26. Wiener, T.F.: Theoretical analysis of gimballess inertial reference equipment using delta-modulated instruments. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge (1962)

Fundamentals of Spacecraft Attitude Determination  
and Control

Markley, F.L.; Crassidis, J.L.

2014, XV, 486 p. 96 illus., 4 illus. in color., Hardcover

ISBN: 978-1-4939-0801-1