1

Kinematics

Spacecraft are free bodies, possessing both translational and rotational motion. The translational component is the subject of *orbital dynamics*, the rotational component is the subject of *attitude dynamics*. It will be seen that the two classes of motion are essentially uncoupled, and can be treated separately.

To be able to study the motion of a spacecraft mathematically, we need a framework for describing it. For this purpose, we need to have a solid understanding of vectors and reference frames, and the associated calculus.

1.1 Physical Vectors

A *physical vector* is a three-dimensional quantity that possesses a *magnitude* and a *direction*. A physical vector will be denoted as $\vec{\mathbf{r}}$, for example. It can be represented graphically by an arrow. Vector addition is defined head-to-tail as shown in Figure 1.1. Multiplication of a vector $\vec{\mathbf{r}}$ by a scalar a scales the magnitude by $|a|$. If a is positive, the direction is unchanged, and if a is negative, the direction is reversed. It is also useful to define a zero-vector denoted by $\vec{\mathbf{0}}$, which has magnitude 0, but no specified direction.

Under these definitions, physical vectors satisfy the following rules for addition:

$$
\begin{aligned}
(\vec{\mathbf{a}} + \vec{\mathbf{b}}) + \vec{\mathbf{c}} &= \vec{\mathbf{a}} + (\vec{\mathbf{b}} + \vec{\mathbf{c}}),\\
\vec{\mathbf{a}} + \vec{\mathbf{b}} &= \vec{\mathbf{b}} + \vec{\mathbf{a}},\\
\vec{\mathbf{a}} + \vec{\mathbf{0}} &= \vec{\mathbf{a}},\\
\vec{\mathbf{a}} + (-\vec{\mathbf{a}}) &= \vec{\mathbf{0}},
\end{aligned}
$$

Spacecraft Dynamics and Control: An Introduction, First Edition.
Anton H.J. de Ruiter, Christopher J. Damaren and James R. Forbes.

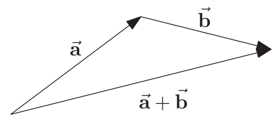

Figure 1.1 Physical vector addition

and the following rules for scalar multiplication:

$$a(b\vec{\mathbf{c}}) = (ab)\vec{\mathbf{c}},$$
$$(a+b)\vec{\mathbf{c}} = a\vec{\mathbf{c}} + b\vec{\mathbf{c}},$$
$$a(\vec{\mathbf{b}}+\vec{\mathbf{c}}) = a\vec{\mathbf{b}} + a\vec{\mathbf{c}},$$
$$1\vec{\mathbf{a}} = \vec{\mathbf{a}},$$
$$0\vec{\mathbf{a}} = \vec{\mathbf{0}}.$$

It is very important to note that the concept of a physical vector is *independent* **of a coordinate system.**

1.1.1 Scalar Product

Given vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, the scalar (or dot) product between the two vectors is defined as

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} \triangleq |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}|\cos\theta,$$

where $0 \leq \theta \leq 180^\circ$ is the small angle between the two vectors, as shown in Figure 1.2. By this definition, the scalar product is commutative, that is

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = \vec{\mathbf{b}} \cdot \vec{\mathbf{a}}.$$

As demonstrated in Figure 1.2, the scalar product $\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}$ is just the projection of $\vec{\mathbf{a}}$ onto $\vec{\mathbf{b}}$ multiplied by $|\vec{\mathbf{b}}|$. Projections are additive, as shown in Figure 1.3, therefore, the scalar product is also distributive, that is

$$(\vec{\mathbf{a}} + \vec{\mathbf{b}}) \cdot \vec{\mathbf{c}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{c}} + \vec{\mathbf{b}} \cdot \vec{\mathbf{c}}. \tag{1.1}$$

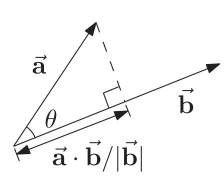

Figure 1.2 Scalar product geometry

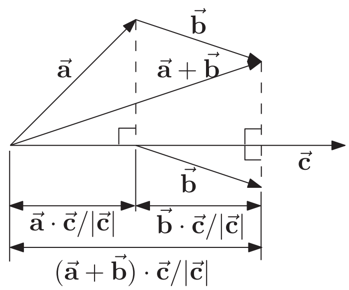

Figure 1.3 Distributivity of scalar product

The following properties are also readily verified from the definition

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{a}} = |\vec{\mathbf{a}}|^2 \geq 0, \tag{1.2}$$

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{a}} = 0 \Leftrightarrow \vec{\mathbf{a}} = \vec{\mathbf{0}}, \tag{1.3}$$

$$\vec{\mathbf{a}} \cdot (c\vec{\mathbf{b}}) = c\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}, \tag{1.4}$$

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = 0 \Leftrightarrow \vec{\mathbf{a}} \perp \vec{\mathbf{b}} \text{ or } \vec{\mathbf{a}} = \vec{\mathbf{0}} \text{ or } \vec{\mathbf{b}} = \vec{\mathbf{0}}. \tag{1.5}$$

1.1.2 Vector Cross Product

Given vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, the cross-product is defined as a vector $\vec{\mathbf{c}}$, denoted by $\vec{\mathbf{c}} = \vec{\mathbf{a}} \times \vec{\mathbf{b}}$ with magnitude

$$|\vec{\mathbf{c}}| = |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}| \sin\theta,$$

with a direction perpendicular to both $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, chosen according to the right-hand rule, as shown in Figure 1.4. Note that $0 \leq \theta \leq 180^\circ$ is again the small angle between the two vectors.

From the definition of the cross-product, it is clear that changing the order simply reverses the direction of the cross-product, that is

$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} = -\vec{\mathbf{b}} \times \vec{\mathbf{a}}.$$

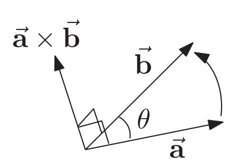

Figure 1.4 Vector cross product

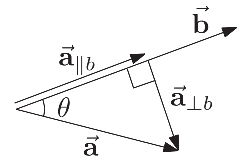

Figure 1.5 Parallel and perpendicular vector components

Now, as shown in Figure 1.5, the vector $\vec{\mathbf{a}}$ can be decomposed into two mutually perpendicular vectors $\vec{\mathbf{a}} = \vec{\mathbf{a}}_{\perp b} + \vec{\mathbf{a}}_{\parallel b}$, where $\vec{\mathbf{a}}_{\perp b}$ is perpendicular to $\vec{\mathbf{b}}$, and $\vec{\mathbf{a}}_{\parallel b}$ is parallel to $\vec{\mathbf{b}}$. These components are given by

$$\vec{\mathbf{a}}_{\parallel b} = \frac{(\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})}{|\vec{\mathbf{b}}|^2}\vec{\mathbf{b}},$$

which is the projection of $\vec{\mathbf{a}}$ onto the direction of $\vec{\mathbf{b}}$, and

$$\vec{\mathbf{a}}_{\perp b} = \vec{\mathbf{a}} - \vec{\mathbf{a}}_{\parallel b} = \vec{\mathbf{a}} - \frac{(\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})}{|\vec{\mathbf{b}}|^2}\vec{\mathbf{b}}.$$

Since $|\vec{\mathbf{a}}_{\perp b}| = |\vec{\mathbf{a}}| \sin\theta$ (see Figure 1.5), and $\vec{\mathbf{a}}_{\perp b}$ is perpendicular to $\vec{\mathbf{b}}$, $|\vec{\mathbf{a}}_{\perp b} \times \vec{\mathbf{b}}| = |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}| \sin\theta$. Since $\vec{\mathbf{a}}_{\perp b}$ lies in the plane defined by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, and points to the same side of $\vec{\mathbf{b}}$ as $\vec{\mathbf{a}}$, $\vec{\mathbf{a}}_{\perp b} \times \vec{\mathbf{b}}$ has the same direction as $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$. Therefore,

$$\vec{\mathbf{a}}_{\perp b} \times \vec{\mathbf{b}} = \vec{\mathbf{a}} \times \vec{\mathbf{b}}. \tag{1.6}$$

Now, we are in a position to show a distributive property of the cross-product. Consider three vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$. First of all, note that

$$\begin{aligned}
(\vec{\mathbf{a}} + \vec{\mathbf{b}})_{\perp c} &= (\vec{\mathbf{a}} + \vec{\mathbf{b}}) - \frac{\left((\vec{\mathbf{a}} + \vec{\mathbf{b}}) \cdot \vec{\mathbf{c}}\right)}{|\vec{\mathbf{c}}|^2}\vec{\mathbf{c}} \\
&= (\vec{\mathbf{a}} + \vec{\mathbf{b}}) - \frac{\left(\vec{\mathbf{a}} \cdot \vec{\mathbf{c}} + \vec{\mathbf{b}} \cdot \vec{\mathbf{c}}\right)}{|\vec{\mathbf{c}}|^2}\vec{\mathbf{c}} \\
&= \left(\vec{\mathbf{a}} - \frac{(\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})}{|\vec{\mathbf{c}}|^2}\vec{\mathbf{c}}\right) + \left(\vec{\mathbf{b}} - \frac{(\vec{\mathbf{b}} \cdot \vec{\mathbf{c}})}{|\vec{\mathbf{c}}|^2}\vec{\mathbf{c}}\right) \\
&= \vec{\mathbf{a}}_{\perp c} + \vec{\mathbf{b}}_{\perp c}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\left(\vec{\mathbf{a}} + \vec{\mathbf{b}}\right) \times \vec{\mathbf{c}} &= \left(\vec{\mathbf{a}} + \vec{\mathbf{b}}\right)_{\perp c} \times \vec{\mathbf{c}} \\
&= \left(\vec{\mathbf{a}}_{\perp c} + \vec{\mathbf{b}}_{\perp c}\right) \times \vec{\mathbf{c}}.
\end{aligned}$$

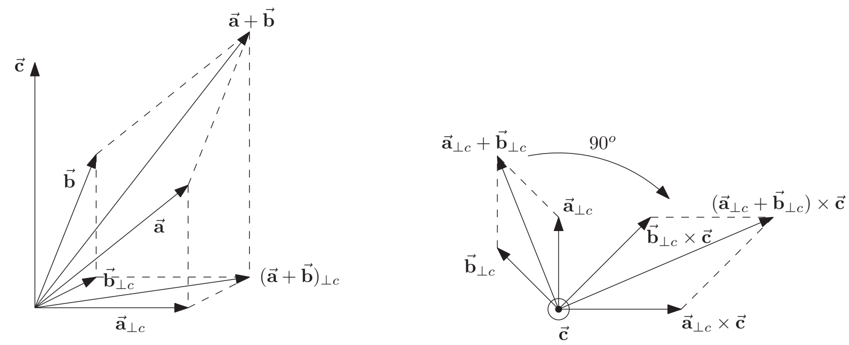

Figure 1.6 Distributivity of vector cross product

Now, the vectors $\vec{\mathbf{a}}_{\perp c}$, $\vec{\mathbf{b}}_{\perp c}$ and $\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}$ all are perpendicular to $\vec{\mathbf{c}}$. Therefore,

$$\begin{aligned}
|\vec{\mathbf{a}}_{\perp c}\times\vec{\mathbf{c}}| &= |\vec{\mathbf{a}}_{\perp c}|\,|\vec{\mathbf{c}}|,\\
|\vec{\mathbf{b}}_{\perp c}\times\vec{\mathbf{c}}| &= |\vec{\mathbf{b}}_{\perp c}|\,|\vec{\mathbf{c}}|,\\
\left|\left(\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}\right)\times\vec{\mathbf{c}}\right| &= \left|\left(\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}\right)\right|\left|\vec{\mathbf{c}}\right|.
\end{aligned}$$

Since the vectors $\vec{\mathbf{a}}_{\perp c}$, $\vec{\mathbf{b}}_{\perp c}$ and $\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}$ are all perpendicular to $\vec{\mathbf{c}}$, the cross-products $\vec{\mathbf{a}}_{\perp c}\times\vec{\mathbf{c}}$, $\vec{\mathbf{b}}_{\perp c}\times\vec{\mathbf{c}}$ and $\left(\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}\right)\times\vec{\mathbf{c}}$ are all simply the vectors $\vec{\mathbf{a}}_{\perp c}$, $\vec{\mathbf{b}}_{\perp c}$ and $\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}$ rotated by 90° about the vector $\vec{\mathbf{c}}$, and then scaled by the factor $|\vec{\mathbf{c}}|$, as shown in Figure 1.6. What this shows is that

$$\left(\vec{\mathbf{a}}_{\perp c}+\vec{\mathbf{b}}_{\perp c}\right)\times\vec{\mathbf{c}}=\vec{\mathbf{a}}_{\perp c}\times\vec{\mathbf{c}}+\vec{\mathbf{b}}_{\perp c}\times\vec{\mathbf{c}},$$

and therefore by (1.6),

$$\left(\vec{\mathbf{a}}+\vec{\mathbf{b}}\right)\times\vec{\mathbf{c}}=\vec{\mathbf{a}}\times\vec{\mathbf{c}}+\vec{\mathbf{b}}\times\vec{\mathbf{c}}, \tag{1.7}$$

which is the distributive property we wanted to show. Finally, the following results are also readily derived from the definition:

$$\vec{\mathbf{a}}\times\vec{\mathbf{a}}=\vec{\mathbf{0}}, \tag{1.8}$$

$$(a\vec{\mathbf{b}})\times\vec{\mathbf{c}}=a(\vec{\mathbf{b}}\times\vec{\mathbf{c}}). \tag{1.9}$$

1.1.3 *Other Useful Vector Identities*

Some other useful vector identities are:

$$\begin{aligned}
&\vec{\mathbf{a}} \times \left(\vec{\mathbf{b}} \times \vec{\mathbf{c}}\right) = (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})\vec{\mathbf{b}} - (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})\vec{\mathbf{c}},\\
&\vec{\mathbf{a}} \cdot \left(\vec{\mathbf{b}} \times \vec{\mathbf{c}}\right) = \vec{\mathbf{b}} \cdot (\vec{\mathbf{c}} \times \vec{\mathbf{a}}) = \vec{\mathbf{c}} \cdot \left(\vec{\mathbf{a}} \times \vec{\mathbf{b}}\right),\\
&\vec{\mathbf{a}} \times \left(\vec{\mathbf{b}} \times \vec{\mathbf{c}}\right) + \vec{\mathbf{b}} \times (\vec{\mathbf{c}} \times \vec{\mathbf{a}}) + \vec{\mathbf{c}} \times \left(\vec{\mathbf{a}} \times \vec{\mathbf{b}}\right) = \vec{\mathbf{0}},\\
&\left(\vec{\mathbf{a}} \times \vec{\mathbf{b}}\right) \cdot \left(\vec{\mathbf{c}} \times \vec{\mathbf{d}}\right) = (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})\left(\vec{\mathbf{b}} \cdot \vec{\mathbf{d}}\right) - \left(\vec{\mathbf{a}} \cdot \vec{\mathbf{d}}\right)\left(\vec{\mathbf{b}} \cdot \vec{\mathbf{c}}\right).
\end{aligned}$$

Note that the definitions of scalar- and cross-product and all of the associated properties and identities above are ***independent*** **of a coordinate system.**

1.2 Reference Frames and Physical Vector Coordinates

Up to this point, we have only considered physical vectors, without any mention of a frame of reference. For computational purposes we need to introduce the concept of a reference frame. Reference frames are also needed to describe the orientation of an object, and are needed for the formulation of kinematics and dynamics.

To define a reference frame, say reference frame "1" (which we will label $\mathcal{F}_1$), it is customary to identify three mutually perpendicular unit length (length of one) physical vectors, labeled as $\vec{\mathbf{x}}_1$, $\vec{\mathbf{y}}_1$ and $\vec{\mathbf{z}}_1$ respectively. The notation used here corresponds to the usual *x*-*y*-*z* axes defined for a Cartesian three-dimensional coordinate system. These three vectors then define the reference frame. The unit vectors are chosen according to the right-handed rule, as shown in Figure 1.7. Under the right-handed rule, the unit vectors satisfy

$$\begin{aligned}
\vec{\mathbf{x}}_1 \times \vec{\mathbf{y}}_1 &= \vec{\mathbf{z}}_1,\\
\vec{\mathbf{y}}_1 \times \vec{\mathbf{z}}_1 &= \vec{\mathbf{x}}_1,\\
\vec{\mathbf{z}}_1 \times \vec{\mathbf{x}}_1 &= \vec{\mathbf{y}}_1.
\end{aligned}$$

Since they are perpendicular, they also satisfy

$$\begin{aligned}
\vec{\mathbf{x}}_1 \cdot \vec{\mathbf{x}}_1 = \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{y}}_1 = \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{z}}_1 = 1,\\
\vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_1 = \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_1 = \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_1 = 0.
\end{aligned} \tag{1.10}$$

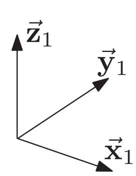

Figure 1.7 Reference frame basis vectors

Now, since the three unit vectors form a basis for physical three-dimensional space, any physical vector $\vec{\mathbf{r}}$ can be written as a linear combination of the unit vectors, that is

$$\begin{aligned}\vec{\mathbf{r}} &= r_{x,1}\vec{\mathbf{x}}_1 + r_{y,1}\vec{\mathbf{y}}_1 + r_{z,1}\vec{\mathbf{z}}_1 \\ &= \begin{bmatrix}\vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1\end{bmatrix}\begin{bmatrix}r_{x,1}\\ r_{y,1}\\ r_{z,1}\end{bmatrix} \\ &= \vec{\mathcal{F}}_1^T\mathbf{r}_1.\end{aligned} \tag{1.11}$$

where

$$\mathbf{r}_1 = \begin{bmatrix}r_{x,1}\\ r_{y,1}\\ r_{z,1}\end{bmatrix} \tag{1.12}$$

is a column matrix containing the coordinates of the physical vector $\vec{\mathbf{r}}$ in reference frame $\mathcal{F}_1$, and

$$\vec{\mathcal{F}}_1 = \begin{bmatrix}\vec{\mathbf{x}}_1\\ \vec{\mathbf{y}}_1\\ \vec{\mathbf{z}}_1\end{bmatrix} \tag{1.13}$$

is a column matrix containing the unit physical vectors defining reference frame $\mathcal{F}_1$. We shall refer to $\vec{\mathcal{F}}_1$ as a *vectrix* (that is, a matrix of physical vectors).

To determine the coordinates of the vector $\vec{\mathbf{r}}$ in frame $\mathcal{F}_1$, we can simply take the dot product of the physical vector (1.11) with each of the unit vectors. For example,

$$\begin{aligned}\vec{\mathbf{r}} \cdot \vec{\mathbf{x}}_1 &= \left(r_{x,1}\vec{\mathbf{x}}_1 + r_{y,1}\vec{\mathbf{y}}_1 + r_{z,1}\vec{\mathbf{z}}_1\right) \cdot \vec{\mathbf{x}}_1, \\ &= r_{x,1}\vec{\mathbf{x}}_1 \cdot \vec{\mathbf{x}}_1 + r_{y,1}\vec{\mathbf{y}}_1 \cdot \vec{\mathbf{x}}_1 + r_{z,1}\vec{\mathbf{z}}_1 \cdot \vec{\mathbf{x}}_1, \\ &= r_{x,1}.\end{aligned}$$

Here, we have made use of properties (1.1) and (1.4) of the dot product of physical vectors. In fact, these properties allow us to treat the dot product of physical vectors in the same manner as scalar multiplication. Using the vectrix notation, we can take advantage of this fact to concisely determine $\mathbf{r}_1$ by taking the dot product of the vector (1.11) with the vectrix (1.13) as follows

$$\begin{aligned}
\vec{\mathcal{F}}_1 \cdot \vec{\mathbf{r}} = \vec{\mathcal{F}}_1 \cdot \left(\vec{\mathcal{F}}_1^T \mathbf{r}_1\right) &= \left(\begin{bmatrix} \vec{\mathbf{x}}_1 \\ \vec{\mathbf{y}}_1 \\ \vec{\mathbf{z}}_1 \end{bmatrix} \cdot \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \right) \mathbf{r}_1 \\
&= \begin{bmatrix} \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{x}}_1 & \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_1 & \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{z}}_1 \end{bmatrix} \mathbf{r}_1 \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{r}_1
\end{aligned}$$

Note that properties (1.1) and (1.4) allowed us to treat the dot product in the same manner as scalar multiplication, and apply the associativity rule for matrix multiplication as we did above. Finally, we have

$$\begin{aligned}
r_{x,1} &= \vec{\mathbf{r}} \cdot \vec{\mathbf{x}}_1, \\
r_{y,1} &= \vec{\mathbf{r}} \cdot \vec{\mathbf{y}}_1, \\
r_{z,1} &= \vec{\mathbf{r}} \cdot \vec{\mathbf{z}}_1.
\end{aligned}$$

1.2.1 Vector Addition and Scalar Multiplication

We can now determine how to perform vector addition and scalar multiplication operations in terms of the coordinates of a vector in a given reference frame. To this end, let us consider two physical vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ expressed in the same reference frame $\mathcal{F}_1$, and a scalar, c:

$$\vec{\mathbf{a}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} a_{x,1} \\ a_{y,1} \\ a_{z,1} \end{bmatrix} = \vec{\mathcal{F}}_1^T \mathbf{a}, \quad \vec{\mathbf{b}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \vec{\mathcal{F}}_1^T \mathbf{b},$$

It is obvious from the rules for physical vector addition and scalar multiplication that

$$\vec{\mathbf{a}} + \vec{\mathbf{b}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} a_{x,1} + b_{x,1} \\ a_{y,1} + b_{y,1} \\ a_{z,1} + b_{z,1} \end{bmatrix} = \vec{\mathcal{F}}_1^T (\mathbf{a} + \mathbf{b}),$$

and

$$c\vec{\mathbf{a}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} c\, a_{x,1} \\ c\, a_{y,1} \\ c\, a_{z,1} \end{bmatrix} = \vec{\mathcal{F}}_1^T \,(c\mathbf{a})\,.$$

That is, vector addition and scalar multiplication operations can be directly applied to the coordinates of the vectors.

1.2.2 Scalar Product

Let us now examine how to compute the scalar (or dot) product in terms of the coordinates of the vectors in a given reference frame. To this end, let us consider two physical vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ expressed in the same reference frame $\mathcal{F}_1$:

$$\vec{\mathbf{a}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} a_{x,1} \\ a_{y,1} \\ a_{z,1} \end{bmatrix}, \quad \vec{\mathbf{b}} = \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix}$$

The dot product is now given by

$$\begin{aligned}
\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} &= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{x}}_1 \\ \vec{\mathbf{y}}_1 \\ \vec{\mathbf{z}}_1 \end{bmatrix} \cdot \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{x}}_1 & \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_1 & \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_1 & \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \mathbf{a}_1^T \mathbf{b}_1
\end{aligned}$$

Again, properties (1.1) and (1.4) allowed us to treat the dot product in the same manner as scalar multiplication, and apply the associativity rule for matrix multiplication as we did above. Making use of identity (1.2), we can relate the length of the physical vector to the length of its coordinate representation, that is:

$$||\mathbf{a}_1|| = \sqrt{\mathbf{a}_1^T \mathbf{a}_1} = |\vec{\mathbf{a}}|,$$

where $||\mathbf{a}_1||$ is the standard Euclidean length of a column matrix.

1.2.3 Vector Cross Product

We can also determine the cross-product of two vectors in terms of the coordinates with respect to a given reference frame. Consider again the same two vectors as in Section 1.2.2. Since the vector cross product satisfies the same distributive and scalar multiplication properties (1.7) and (1.9) as the vector dot product (compare to (1.1) and (1.4)), we can concisely determine the vector cross-product in terms of the coordinates in the same manner as we determined the dot product in Section 1.2.2 (provided we respect the order in which each individual vector cross-product is taken). We have

$$
\begin{aligned}
\vec{\mathbf{a}} \times \vec{\mathbf{b}} &= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{x}}_1 \\ \vec{\mathbf{y}}_1 \\ \vec{\mathbf{z}}_1 \end{bmatrix} \times \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{x}}_1 \times \vec{\mathbf{x}}_1 & \vec{\mathbf{x}}_1 \times \vec{\mathbf{y}}_1 & \vec{\mathbf{x}}_1 \times \vec{\mathbf{z}}_1 \\ \vec{\mathbf{y}}_1 \times \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 \times \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_1 \times \vec{\mathbf{z}}_1 \\ \vec{\mathbf{z}}_1 \times \vec{\mathbf{x}}_1 & \vec{\mathbf{z}}_1 \times \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \times \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \begin{bmatrix} a_{x,1} & a_{y,1} & a_{z,1} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{0}} & \vec{\mathbf{z}}_1 & -\vec{\mathbf{y}}_1 \\ -\vec{\mathbf{z}}_1 & \vec{\mathbf{0}} & \vec{\mathbf{x}}_1 \\ \vec{\mathbf{y}}_1 & -\vec{\mathbf{x}}_1 & \vec{\mathbf{0}} \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \begin{bmatrix} 0 & -a_{z,1} & a_{y,1} \\ a_{z,1} & 0 & -a_{x,1} \\ -a_{y,1} & a_{x,1} & 0 \end{bmatrix} \begin{bmatrix} b_{x,1} \\ b_{y,1} \\ b_{z,1} \end{bmatrix} \\
&= \vec{\mathcal{F}}_1^T \mathbf{a}_1^{\times} \mathbf{b}_1,
\end{aligned}
$$

where the 3 × 3 matrix

$$
\mathbf{a}_1^{\times} \triangleq \begin{bmatrix} 0 & -a_{z,1} & a_{y,1} \\ a_{z,1} & 0 & -a_{x,1} \\ -a_{y,1} & a_{x,1} & 0 \end{bmatrix}
$$

is the cross-product operator matrix corresponding to the vector $\vec{\mathbf{a}}$ in reference frame $\mathcal{F}_1$ coordinates.

1.2.4 Column Matrix Identities

The vector identities presented in Sections 1.1.2 and 1.1.3 can all be rewritten in terms of column matrices. To this end, let us consider four physical vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$, $\vec{\mathbf{c}}$ and $\vec{\mathbf{d}}$ expressed in the same reference frame $\mathcal{F}$, with corresponding coordinates $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ respectively.

Making use of the above results for the computations for the scalar- and cross-products, the following column matrix identities are automatically obtained

$$\begin{aligned}
&\mathbf{a}^{\times}\mathbf{a} = \mathbf{0}\\
&\mathbf{a}^{\times}\mathbf{b} = -\mathbf{b}^{\times}\mathbf{a}\\
&\mathbf{a}^{\times}\mathbf{b}^{\times}\mathbf{c} = (\mathbf{a}^T\mathbf{c})\mathbf{b} - (\mathbf{a}^T\mathbf{b})\mathbf{c},\\
&\mathbf{a}^T\mathbf{b}^{\times}\mathbf{c} = \mathbf{b}^T\mathbf{c}^{\times}\mathbf{a} = \mathbf{c}^T\mathbf{a}^{\times}\mathbf{b},\\
&\mathbf{a}^{\times}\mathbf{b}^{\times}\mathbf{c} + \mathbf{b}^{\times}\mathbf{c}^{\times}\mathbf{a} + \mathbf{c}^{\times}\mathbf{a}^{\times}\mathbf{b} = \mathbf{0},\\
&\left(\mathbf{a}^{\times}\mathbf{b}\right)^T\left(\mathbf{c}^{\times}\mathbf{d}\right) = \left(\mathbf{a}^T\mathbf{c}\right)\left(\mathbf{b}^T\mathbf{d}\right) - \left(\mathbf{a}^T\mathbf{d}\right)\left(\mathbf{b}^T\mathbf{c}\right).
\end{aligned}$$

The third identity leads to a very useful result. That is,

$$\begin{aligned}
\mathbf{a}^{\times}\mathbf{b}^{\times}\mathbf{c} &= (\mathbf{a}^T\mathbf{c})\mathbf{b} - (\mathbf{a}^T\mathbf{b})\mathbf{c}\\
&= \left(\mathbf{b}\mathbf{a}^T - (\mathbf{a}^T\mathbf{b})\mathbf{1}\right)\mathbf{c}
\end{aligned}$$

where we denote the identity matrix by **1**. Since the above result must hold for any column matrix $\mathbf{c} \in R^3$, we must have

$$\mathbf{a}^{\times}\mathbf{b}^{\times} = \mathbf{b}\mathbf{a}^T - (\mathbf{a}^T\mathbf{b})\mathbf{1}. \tag{1.14}$$

1.3 Rotation Matrices

Spacecraft dynamics problems generally involve the use of several reference frames. For example, to describe the orientation of the spacecraft with respect to the Earth, it makes sense to fix one reference frame (say $\mathcal{F}_1$) to the Earth, and the other (say $\mathcal{F}_2$) to the spacecraft body. The orientation of the spacecraft with respect to the Earth can then be described by the orientation of reference frame $\mathcal{F}_2$ with respect to $\mathcal{F}_1$. In spacecraft terminology, we call the orientation the *attitude*.

Since multiple reference frames will be used, it is necessary to know how to:

1. Describe the orientation of one reference frame with respect to another.
2. Transform the coordinates of a vector from one reference frame to another.

As we shall see, these two issues are directly related. Let us now consider two reference frames, $\mathcal{F}_1$ defined by the unit vectors $\vec{\mathbf{x}}_1$, $\vec{\mathbf{y}}_1$ and $\vec{\mathbf{z}}_1$, and $\mathcal{F}_2$ defined by the unit vectors $\vec{\mathbf{x}}_2$, $\vec{\mathbf{y}}_2$ and $\vec{\mathbf{z}}_2$. Let us also consider an arbitrary vector $\vec{\mathbf{r}}$, as shown in Figure 1.8. Let us now express the vector $\vec{\mathbf{r}}$ in each frame as

$$\vec{\mathbf{r}} = \vec{\mathcal{F}}_1^T\mathbf{r}_1 = \vec{\mathcal{F}}_2^T\mathbf{r}_2.$$

As we have shown before, the column matrices $\mathbf{r}_1$ and $\mathbf{r}_2$ contain the coordinates of the vector $\vec{\mathbf{r}}$, in the reference frames $\mathcal{F}_1$ and $\mathcal{F}_2$ respectively. To address point 1, we can find the

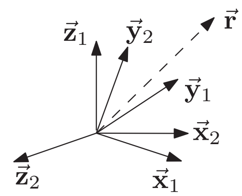

Figure 1.8 Multiple reference frames

coordinates in reference frame $\mathcal{F}_2$ of the unit vectors defining reference frame $\mathcal{F}_1$. As we have seen before, we can write

$$\vec{\mathbf{x}}_1 = \vec{\mathcal{F}}_2^T \mathbf{x}_{1,2}, \ \ \vec{\mathbf{y}}_1 = \vec{\mathcal{F}}_2^T \mathbf{y}_{1,2}, \ \text{and}\ \vec{\mathbf{z}}_1 = \vec{\mathcal{F}}_2^T \mathbf{z}_{1,2},$$

where

$$\mathbf{x}_{1,2} = \begin{bmatrix} \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{x}}_2 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{y}}_2 \\ \vec{\mathbf{x}}_1 \cdot \vec{\mathbf{z}}_2 \end{bmatrix}, \ \ \mathbf{y}_{1,2} = \begin{bmatrix} \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{x}}_2 \\ \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{y}}_2 \\ \vec{\mathbf{y}}_1 \cdot \vec{\mathbf{z}}_2 \end{bmatrix}, \ \text{and}\ \mathbf{z}_{1,2} = \begin{bmatrix} \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{x}}_2 \\ \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{y}}_2 \\ \vec{\mathbf{z}}_1 \cdot \vec{\mathbf{z}}_2 \end{bmatrix}. \tag{1.15}$$

To address point 2, we now obtain a relationship between coordinates of $\vec{\mathbf{r}}$ in the different reference frames as follows:

$$\begin{aligned} \vec{\mathcal{F}}_2^T \mathbf{r}_2 &= \vec{\mathcal{F}}_1^T \mathbf{r}_1 \quad \Rightarrow \\ \vec{\mathcal{F}}_2 \cdot \vec{\mathcal{F}}_2^T \mathbf{r}_2 &= \vec{\mathcal{F}}_2 \cdot \vec{\mathcal{F}}_1^T \mathbf{r}_1 \Rightarrow \\ \mathbf{r}_2 &= \mathbf{C}_{21} \mathbf{r}_1 \end{aligned} \tag{1.16}$$

where

$$\begin{aligned} \mathbf{C_{21}} &= \vec{\mathcal{F}}_2 \cdot \vec{\mathcal{F}}_1^T \\ &= \begin{bmatrix} \vec{\mathbf{x}}_2 \\ \vec{\mathbf{y}}_2 \\ \vec{\mathbf{z}}_2 \end{bmatrix} \cdot \begin{bmatrix} \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_1 \end{bmatrix} \\ &= \begin{bmatrix} \vec{\mathbf{x}}_2 \cdot \vec{\mathbf{x}}_1 & \vec{\mathbf{x}}_2 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{x}}_2 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{y}}_2 \cdot \vec{\mathbf{x}}_1 & \vec{\mathbf{y}}_2 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{y}}_2 \cdot \vec{\mathbf{z}}_1 \\ \vec{\mathbf{z}}_2 \cdot \vec{\mathbf{x}}_1 & \vec{\mathbf{z}}_2 \cdot \vec{\mathbf{y}}_1 & \vec{\mathbf{z}}_2 \cdot \vec{\mathbf{z}}_1 \end{bmatrix} \end{aligned} \tag{1.17}$$

The matrix $\mathbf{C}_{21}$ is called a *rotation matrix*. It is sometimes referred to as the "direction cosine matrix", since each entry is a scalar product between two unit vectors, giving the cosine of the angle between them (by definition of the scalar product). Comparing the matrix $\mathbf{C}_{21}$ in (1.17) with (1.15), it can be seen that the columns of the matrix $\mathbf{C}_{21}$ are just the unit vectors defining $\mathcal{F}_1$ expressed in $\mathcal{F}_2$. That is,

$$\mathbf{C}_{21} = \begin{bmatrix} \mathbf{x}_{1,2} & \mathbf{y}_{1,2} & \mathbf{z}_{1,2} \end{bmatrix}. \tag{1.18}$$

The rotation matrix $\mathbf{C}_{21}$ therefore contains all of the information needed to address both points 1 and 2. In particular, the rotation matrix fully describes the attitude of $\mathcal{F}_2$ with respect to $\mathcal{F}_1$.

Rotation matrices have a number of special properties. First of all, from equation (1.18), the rotation matrix is invertible, since its columns are the coordinates of mutually perpendicular unit vectors. Furthermore, making use of the relationship between the unit vectors in (1.10), we have

$$\begin{aligned} \mathbf{C}_{21}^T\mathbf{C}_{21} &= \begin{bmatrix} \mathbf{x}_{1,2}^T \\ \mathbf{y}_{1,2}^T \\ \mathbf{z}_{1,2}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1,2} & \mathbf{y}_{1,2} & \mathbf{z}_{1,2} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x}_{1,2}^T\mathbf{x}_{1,2} & \mathbf{x}_{1,2}^T\mathbf{y}_{1,2} & \mathbf{x}_{1,2}^T\mathbf{z}_{1,2} \\ \mathbf{y}_{1,2}^T\mathbf{x}_{1,2} & \mathbf{y}_{1,2}^T\mathbf{y}_{1,2} & \mathbf{y}_{1,2}^T\mathbf{z}_{1,2} \\ \mathbf{z}_{1,2}^T\mathbf{x}_{1,2} & \mathbf{z}_{1,2}^T\mathbf{y}_{1,2} & \mathbf{z}_{1,2}^T\mathbf{z}_{1,2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{1}. \end{aligned}$$

This shows that $\mathbf{C}_{21}^{-1} = \mathbf{C}_{21}^T$, and that the rotation matrix is an *orthonormal matrix*, since its inverse is equal to its transpose. Making use of this result, inverting the transformation between coordinates in equation (1.16) gives

$$\mathbf{r}_1 = \mathbf{C}_{21}^T\mathbf{r}_2 = \mathbf{C}_{12}\mathbf{r}_2.$$

That is, $\mathbf{C}_{12} = \mathbf{C}_{21}^T$.

Let us now consider three reference frames, $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$. The vector $\vec{\mathbf{r}}$ is then given by

$$\vec{\mathbf{r}} = \vec{\mathcal{F}}_1^T\mathbf{r}_1 = \vec{\mathcal{F}}_2^T\mathbf{r}_2 = \vec{\mathcal{F}}_3^T\mathbf{r}_3.$$

From our results obtained so far, we have the following transformations:

$$\mathbf{r}_3 = \mathbf{C}_{31}\mathbf{r}_1,$$

and

$$\mathbf{r}_3 = \mathbf{C}_{32}\mathbf{r}_2 = \mathbf{C}_{32}\mathbf{C}_{21}\mathbf{r}_1.$$

Combining these two results leads to (since they must hold for any column matrix $\mathbf{r}_1$)

$$\mathbf{C}_{31} = \mathbf{C}_{32}\mathbf{C}_{21}. \tag{1.19}$$

That is, products of rotation matrices are also rotation matrices, and successive rotations can be combined by multiplying the rotation matrices in the reverse order of the rotations (from right to left).

An important relationship arises when considering the cross-product of two vectors, expressed in two different reference frames. Consider the vectors $\vec{\mathbf{a}} = \vec{\mathcal{F}}_1^T \mathbf{a}_1 = \vec{\mathcal{F}}_2^T \mathbf{a}_2$ and $\vec{\mathbf{b}} = \vec{\mathcal{F}}_1^T \mathbf{b}_1 = \vec{\mathcal{F}}_2^T \mathbf{b}_2$. We have the relationships $\mathbf{a}_2 = \mathbf{C}_{21}\mathbf{a}_1$ and $\mathbf{b}_2 = \mathbf{C}_{21}\mathbf{b}_1$. Let us now examine the cross product of the two vectors. We have

$$\begin{aligned}\vec{\mathbf{a}} \times \vec{\mathbf{b}} = \vec{\mathcal{F}}_2^T \mathbf{a}_2^\times \mathbf{b}_2 &= \vec{\mathcal{F}}_1^T \mathbf{a}_1^\times \mathbf{b}_1 \\ &= \vec{\mathcal{F}}_2^T \mathbf{C}_{21}\mathbf{a}_1^\times \mathbf{b}_1 \\ &= \vec{\mathcal{F}}_2^T \mathbf{C}_{21}\mathbf{a}_1^\times \mathbf{C}_{21}^T \mathbf{b}_2.\end{aligned}$$

From this we obtain

$$\mathbf{a}_2^\times \mathbf{b}_2 = \mathbf{C}_{21}\mathbf{a}_1^\times \mathbf{C}_{21}^T \mathbf{b}_2.$$

Since this must hold for any $\mathbf{b}_2 \in R^3$, we have

$$\mathbf{a}_2^\times = \mathbf{C}_{21}\mathbf{a}_1^\times \mathbf{C}_{21}^T. \tag{1.20}$$

Finally, since $\mathbf{a}_2 = \mathbf{C}_{21}\mathbf{a}_1$, we have the result

$$(\mathbf{C}_{21}\mathbf{a}_1)^\times = \mathbf{C}_{21}\mathbf{a}_1^\times \mathbf{C}_{21}^T. \tag{1.21}$$

1.3.1 Principal Rotations

An important class of rotations are those about one of the coordinate axes, as shown in Figure 1.9. When reference frame $\mathcal{F}_2$ is obtained from frame $\mathcal{F}_1$ by a rotation about the z-axis, the associated rotation matrix is

$$\mathbf{C}_z(\theta_z) = \begin{bmatrix} \cos\theta_z & \sin\theta_z & 0 \\ -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The subscript "z" is used here instead of the subscript "21" to emphasize the fact that the rotation has occurred about the z-axis.

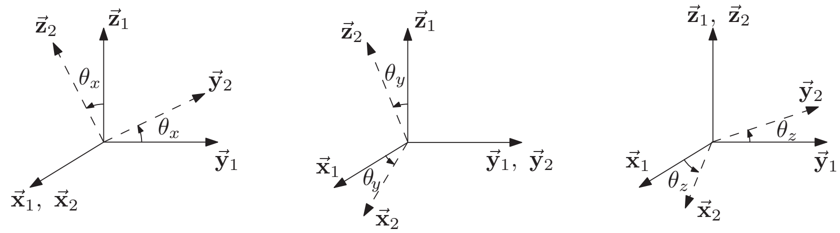

Figure 1.9 Principal rotations

For a rotation about the *y*-axis,

$$\mathbf{C}_y(\theta_y) = \begin{bmatrix} \cos\theta_y & 0 & -\sin\theta_y \\ 0 & 1 & 0 \\ \sin\theta_y & 0 & \cos\theta_y \end{bmatrix}.$$

For a rotation about the *x*-axis,

$$\mathbf{C}_x(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{bmatrix}.$$

1.3.2 General Rotations

Up to this point, we have given the definition of the transformation between reference frames (1.17), but we have not justified why it is called a rotation matrix. The answer is found in Euler's Theorem, which was obtained by Leonhard Euler in 1775.

Euler's Theorem: *The most general motion of a rigid body with one point fixed is a rotation about an axis through that point.*

As illustrated in Figure 1.10, what Euler's theorem means is that given any two reference frames $\mathcal{F}_1$ and $\mathcal{F}_2$, frame $\mathcal{F}_2$ can be obtained from frame $\mathcal{F}_1$ by a single rotation about some unit vector which we will denote $\vec{\mathbf{a}} = \vec{\mathcal{F}}_1^T\mathbf{a}$. We shall now demonstrate this. Consider the rotation matrix $\mathbf{C}_{12}$, which transforms coordinates of vectors in frame $\mathcal{F}_2$ to coordinates in $\mathcal{F}_1$. As we have seen, the rotation matrix is given by

$$\mathbf{C}_{12} = \begin{bmatrix} \mathbf{x}_{2,1} & \mathbf{y}_{2,1} & \mathbf{z}_{2,1} \end{bmatrix}, \tag{1.22}$$

where $\mathbf{x}_{2,1}$, $\mathbf{y}_{2,1}$ and $\mathbf{z}_{2,1}$ are the coordinates in $\mathcal{F}_1$ of the basis vectors defining $\mathcal{F}_2$, that is,

$$\vec{\mathbf{x}}_2 = \vec{\mathcal{F}}_1^T\mathbf{x}_{2,1},\ \ \vec{\mathbf{y}}_2 = \vec{\mathcal{F}}_1^T\mathbf{y}_{2,1},\ \ \vec{\mathbf{z}}_2 = \vec{\mathcal{F}}_1^T\mathbf{z}_{2,1}.$$

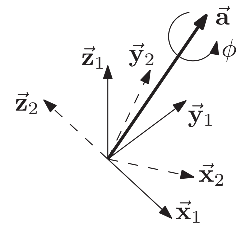

Figure 1.10 General rotation

First of all, we note that the determinant of $\mathbf{C}_{12}$ is given by

$$\det[\mathbf{C}_{12}] = \mathbf{x}_{2,1}^T \mathbf{y}_{2,1}^{\times} \mathbf{z}_{2,1}.$$

This can readily be shown by expansion of the determinant. Since $\mathcal{F}_1$ is a right-handed frame, we have seen that this means that

$$\det[\mathbf{C}_{12}] = \vec{\mathbf{x}}_2 \cdot \left(\vec{\mathbf{y}}_2 \times \vec{\mathbf{z}}_2\right).$$

Since $\mathcal{F}_2$ is a right-handed frame, we have $\vec{\mathbf{x}}_2 = \vec{\mathbf{y}}_2 \times \vec{\mathbf{z}}_2$, and therefore

$$\det[\mathbf{C}_{12}] = 1. \tag{1.23}$$

Therefore, a rotation matrix is a 3 by 3 orthonormal matrix with determinant equal to $+1$. Next, we shall show that $+1$ is an eigenvalue of the rotation matrix. We shall make use of the following facts for matrix determinants.

1. The determinant of a square matrix is equal to the determinant of its transpose,

$$\det[\mathbf{A}] = \det\left[\mathbf{A}^T\right].$$

2. The determinant of a product of square matrices is equal to the product of the individual determinants,

$$\det[\mathbf{AB}] = \det[\mathbf{A}]\det[\mathbf{B}].$$

3. Given an n by n matrix, $\mathbf{A}$ and a scalar a,

$$\det[a\mathbf{A}] = a^n \det[\mathbf{A}].$$

We shall now show that +1 is an eigenvalue of the rotation matrix $\mathbf{C}_{12}$. Consider $\det[\mathbf{C}_{12} - \mathbf{1}]$, where $\mathbf{1}$ is the 3 by 3 identity matrix. Then, since $\det[\mathbf{C}_{12}] = 1$, we have

$$\det[\mathbf{C}_{12} - \mathbf{1}] = \det\left[\mathbf{C}_{12}^T\right]\det[\mathbf{C}_{12} - \mathbf{1}],$$

by fact 1. By fact 2 and the orthonormality of $\mathbf{C}_{12}$, this becomes

$$\begin{aligned}\det\left[\mathbf{C}_{12}-\mathbf{1}\right] &= \det\left[\mathbf{C}_{12}^T\left(\mathbf{C}_{12}-\mathbf{1}\right)\right]\\ &= \det\left[\mathbf{1}-\mathbf{C}_{12}^T\right].\end{aligned}$$

By fact 1 and fact 3, this becomes

$$\begin{aligned}\det\left[\mathbf{C}_{12}-\mathbf{1}\right] &= \det\left[\mathbf{1}-\mathbf{C}_{12}\right]\\ &= (-1)^3\det\left[\mathbf{C}_{12}-\mathbf{1}\right].\end{aligned}$$

The only way that this final equality holds is if

$$\det\left[\mathbf{C}_{12}-\mathbf{1}\right] = 0,$$

which shows that +1 is indeed an eigenvalue of $\mathbf{C}_{12}$.

Now, let $\mathbf{a}$ be a unit eigenvector of $\mathbf{C}_{12}$, corresponding to the eigenvalue +1. Then, we have

$$\mathbf{C}_{12}\mathbf{a} = \mathbf{a},$$

and the rotation matrix $\mathbf{C}_{12}$ leaves the unit eigenvector $\mathbf{a}$ unchanged. We now define the unit physical vector $\vec{\mathbf{a}} = \vec{\mathcal{F}}_1^T\mathbf{a}$. Let us now choose another unit vector $\vec{\mathbf{b}} = \vec{\mathcal{F}}_1^T\mathbf{b}$ perpendicular to $\vec{\mathbf{a}}$, that is $\mathbf{b}^T\mathbf{b} = 1$ and $\mathbf{a}^T\mathbf{b} = 0$. Transforming the coordinates of the vector $\vec{\mathbf{b}}$ through the rotation matrix yields a new vector given by

$$\vec{\mathbf{b}}' = \vec{\mathcal{F}}_1^T\mathbf{C}_{12}\mathbf{b}.$$

We see immediately that

$$|\vec{\mathbf{b}}'| = \left((\mathbf{C}_{12}\mathbf{b})^T\,\mathbf{C}_{12}\mathbf{b}\right)^{\frac{1}{2}} = 1,$$

and

$$\begin{aligned}\vec{\mathbf{a}}\cdot\vec{\mathbf{b}}' &= \mathbf{a}^T\mathbf{C}_{12}\mathbf{b}\\ &= (\mathbf{C}_{12}\mathbf{a})^T\,\mathbf{C}_{12}\mathbf{b}\\ &= 0.\end{aligned}$$

Therefore, as shown in Figure 1.11, the transformed vector $\vec{\mathbf{b}}'$ is also a unit vector perpendicular to $\vec{\mathbf{a}}$. This means that the transformation of the coordinates $\mathbf{b}$ by the matrix $\mathbf{C}_{12}$ is equivalent to a rotation of the vector $\vec{\mathbf{b}}$ about the vector $\vec{\mathbf{a}}$ through some angle ϕ. Note that we define the angle of rotation ϕ to be positive in the right hand sense about $\vec{\mathbf{a}}$. Next, let us define a third unit vector $\vec{\mathbf{c}} = \vec{\mathcal{F}}_1^T\mathbf{c} = \vec{\mathbf{a}}\times\vec{\mathbf{b}}$. In frame $\mathcal{F}_1$, the coordinates are $\mathbf{c} = \mathbf{a}^\times\mathbf{b}$. The vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ therefore constitute an orthogonal triad, and any other vector can be written in terms of these. Transforming the coordinates of the vector $\vec{\mathbf{c}}$ through the rotation matrix yields a new vector given by

$$\vec{\mathbf{c}}' = \vec{\mathcal{F}}_1^T\mathbf{C}_{12}\mathbf{c}.$$

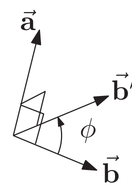

Figure 1.11 Rotation about $\vec{\mathbf{a}}$

We now see that

$$\begin{aligned}\vec{\mathbf{c}}' &= \vec{\mathcal{F}}_1^T \mathbf{C}_{12}\mathbf{a}^{\times}\mathbf{b},\\ &= \vec{\mathcal{F}}_1^T \mathbf{C}_{12}\mathbf{a}^{\times}\mathbf{C}_{12}^T\mathbf{C}_{12}\mathbf{b},\\ &= \vec{\mathcal{F}}_1^T \left(\mathbf{C}_{12}\mathbf{a}\right)^{\times} \mathbf{C}_{12}\mathbf{b},\\ &= \vec{\mathcal{F}}_1^T \mathbf{a}^{\times}\mathbf{C}_{12}\mathbf{b},\\ &= \vec{\mathbf{a}} \times \vec{\mathbf{b}}'.\end{aligned}$$

This means that the vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}'$ and $\vec{\mathbf{c}}'$ are also an orthogonal triad with the same sense as $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$. The only way that this is possible is if the transformation of the coordinates $\mathbf{c}$ by the matrix $\mathbf{C}_{12}$ is equivalent to a rotation of the vector $\vec{\mathbf{c}}$ about the vector $\vec{\mathbf{a}}$ through the same angle ϕ as the vector $\vec{\mathbf{b}}$, as shown in Figure 1.12. Now, since any vector $\vec{\mathbf{g}} = \vec{\mathcal{F}}_1^T \mathbf{g}$ can be written as a linear combination of the vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$, the transformation of the coordinates $\mathbf{g}$ by the matrix $\mathbf{C}_{12}$ is equivalent to a rotation of the vector $\vec{\mathbf{g}}$ about the vector $\vec{\mathbf{a}}$ through the same angle ϕ.

Let us now see what this means in terms of the relationship between frames $\mathcal{F}_1$ and $\mathcal{F}_2$. It is easy to from the expression for the rotation matrix in (1.22) that

$$\mathbf{x}_{2,1} = \mathbf{C}_{12}\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \mathbf{y}_{2,1} = \mathbf{C}_{12}\begin{bmatrix}0\\1\\0\end{bmatrix},\quad \mathbf{z}_{2,1} = \mathbf{C}_{12}\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Since the basis vectors defining frame $\mathcal{F}_1$ satisfy

$$\vec{\mathbf{x}}_1 = \vec{\mathcal{F}}_1^T \mathbf{x}_{1,1},\quad \vec{\mathbf{y}}_1 = \vec{\mathcal{F}}_1^T \mathbf{y}_{1,1},\quad \vec{\mathbf{z}}_1 = \vec{\mathcal{F}}_1^T \mathbf{y}_{1,1},$$

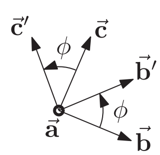

Figure 1.12 Rotation about $\vec{\mathbf{a}}$

with

$$\mathbf{x}_{1,1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{y}_{1,1} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{z}_{1,1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

it is clear now that the basis vectors defining frame $\mathcal{F}_2$ can all be obtained by rotating the basis vectors defining frame $\mathcal{F}_1$ about the unit vector $\vec{\mathbf{a}}$ through the angle ϕ. This is precisely Euler's Theorem.

1.3.2.1 Rotation Matrix in Terms of the Principal Axis and Angle of Rotation

We shall now find an expression for the rotation matrix in terms of the principal axis and angle of rotation. Let $\vec{\mathbf{a}} = \vec{\mathcal{F}}_1^T \mathbf{a}$ and ϕ be the principal axis and angle of rotation corresponding to the rotation matrix $\mathbf{C}_{12}$, respectively.

To this end, let us examine a rotation of an arbitrary vector $\vec{\mathbf{v}} = \vec{\mathcal{F}}_1^T \mathbf{v}$ about $\vec{\mathbf{a}}$ through angle ϕ in the right-hand sense, as shown in Figure 1.13. We can decompose the vector $\vec{\mathbf{v}}$ into a part parallel to $\vec{\mathbf{a}}$ (denoted $\vec{\mathbf{v}}_{\parallel}$), and a part perpendicular to $\vec{\mathbf{a}}$ (denoted $\vec{\mathbf{v}}_{\perp}$). That is,

$$\vec{\mathbf{v}} = \vec{\mathbf{v}}_{\parallel} + \vec{\mathbf{v}}_{\perp}.$$

These components are given by

$$\begin{aligned} \vec{\mathbf{v}}_{\parallel} &= (\vec{\mathbf{a}} \cdot \vec{\mathbf{v}})\vec{\mathbf{a}}, \\ \vec{\mathbf{v}}_{\perp} &= \vec{\mathbf{v}} - (\vec{\mathbf{a}} \cdot \vec{\mathbf{v}})\vec{\mathbf{a}}. \end{aligned} \tag{1.24}$$

Now, as seen in Figure 1.13, to rotate the vector $\vec{\mathbf{v}}$ about $\vec{\mathbf{a}}$ (which we will denote $\vec{\mathbf{v}}_{rot}$), we only need to rotate the perpendicular part (which we will denote $\vec{\mathbf{v}}_{\perp rot}$), i.e.

$$\vec{\mathbf{v}}_{rot} = \vec{\mathbf{v}}_{\parallel} + \vec{\mathbf{v}}_{\perp rot}.$$

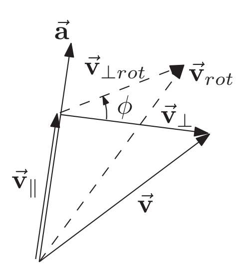

Figure 1.13 Rotation of a vector

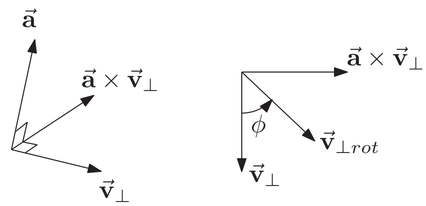

Figure 1.14 Rotation of the perpendicular component of a vector

To this end, we can set up a vector perpendicular to $\vec{\mathbf{a}}$ and $\vec{\mathbf{v}}_\perp$, given by $\vec{\mathbf{a}} \times \vec{\mathbf{v}}_\perp$ as shown in Figure 1.14. Note that $|\vec{\mathbf{a}} \times \vec{\mathbf{v}}_\perp| = |\vec{\mathbf{v}}_\perp|$. As seen in Figure 1.14, the rotated vector $\vec{\mathbf{v}}_{\perp rot}$ can now be expressed as

$$\vec{\mathbf{v}}_{\perp rot} = \vec{\mathbf{v}}_\perp \cos\phi + \vec{\mathbf{a}} \times \vec{\mathbf{v}}_\perp \sin\phi.$$

Substituting (1.24) and making use of the fact that $\vec{\mathbf{a}} \times \vec{\mathbf{a}} = \vec{\mathbf{0}}$ gives

$$\vec{\mathbf{v}}_{rot} = \vec{\mathbf{v}} \cos\phi + (\vec{\mathbf{a}} \cdot \vec{\mathbf{v}})\vec{\mathbf{a}}(1 - \cos\phi) + \vec{\mathbf{a}} \times \vec{\mathbf{v}} \sin\phi.$$

Expressing this in $\mathcal{F}_1$ coordinates, we have

$$\vec{\mathbf{v}}_{rot} = \vec{\mathcal{F}}_1^T \left[\cos\phi\mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T + \sin\phi\mathbf{a}^\times\right]\mathbf{v} = \vec{\mathcal{F}}_1^T \mathbf{C}_{12}\mathbf{v}. \tag{1.25}$$

Note that the second equality is due to the fact that the axis $\vec{\mathbf{a}}$ and angle ϕ correspond to the rotation matrix $\mathbf{C}_{12}$. Since the above relationship must hold for any $\mathbf{v} \in \mathbf{R}^3$, it must be that

$$\mathbf{C}_{12} = \cos\phi\mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T + \sin\phi\mathbf{a}^\times.$$

Taking the transpose, we finally obtain

$$\mathbf{C}_{21} = \cos\phi\mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi\mathbf{a}^\times. \tag{1.26}$$

Note that while we have used the coordinates of $\vec{\mathbf{a}}$ in reference frame $\mathcal{F}_1$, we could have used the coordinates in $\mathcal{F}_2$, since they are the same. A simple calculation shows that

$$\mathbf{C}_{21}\mathbf{a} = \mathbf{a},$$

as it should be. Another consequence of Euler's Theorem is that while the rotation matrix $\mathbf{C}_{21}$ contains nine entries, it can be fully described by four parameters, that is the three components of the axis of rotation $\mathbf{a}$, and that angle of rotation ϕ. In fact, the three components of the axis of rotation are not independent, since $\mathbf{a}^T\mathbf{a} = 1$ ($\vec{\mathbf{a}}$ is a unit vector), so in fact, three parameters could be used to fully describe the rotation matrix.

1.3.2.2 Finding the Principal Axis and Angle from the Rotation Matrix

Let

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \text{ and } \mathbf{C}_{21} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}.$$

Then, taking the trace of $\mathbf{C}_{21}$ in (1.26), we obtain

$$\text{trace}\,[\mathbf{C}_{21}] = 3\cos\phi + (1 - \cos\phi)(a_1^2 + a_2^2 + a_3^2) = 1 + 2\cos\phi.$$

Note that the last equality follows since $\mathbf{a}$ is a unit column matrix. Therefore, we can find the principal angle of rotation as

$$\phi = \cos^{-1}\left(\frac{\text{trace}\,[\mathbf{C}_{21}] - 1}{2}\right). \tag{1.27}$$

An important point to note is that a rotation about the axis $\vec{\mathbf{a}}$ through the angle ϕ is equivalent to a rotation about the axis $-\vec{\mathbf{a}}$ through the angle $-\phi$. Therefore, when solving 1.27, we shall restrict ourselves to the range $0 \leq \phi \leq \pi$, since a rotation angle in the range $-\pi < \phi < 0$ about $\mathbf{a}$ is equivalent to a rotation angle in the range $0 \leq \phi \leq \pi$ about the negative axis of rotation $-\mathbf{a}$.

Having found the angle ϕ, we can now find the corresponding axis of rotation $\mathbf{a}$. From (1.26), we find that

$$\mathbf{C}_{21}^T - \mathbf{C}_{21} = 2\sin\phi\mathbf{a}^\times,$$

from which we can readily find that if $0 < \phi < \pi$, the axis of rotation can be obtained from

$$\begin{aligned} a_1 &= \frac{c_{23} - c_{32}}{2\sin\phi}, \\ a_2 &= \frac{c_{31} - c_{13}}{2\sin\phi}, \\ a_3 &= \frac{c_{12} - c_{21}}{2\sin\phi}. \end{aligned} \tag{1.28}$$

When $\phi = \pm\pi$, the rotation matrix becomes

$$\mathbf{C}_{21} = -\mathbf{1} + 2\mathbf{a}\mathbf{a}^T.$$

From this, we find that if $\phi = \pm\pi$,

$$|a_1| = \left(\frac{c_{11} + 1}{2}\right)^{\frac{1}{2}}, \quad |a_2| = \left(\frac{c_{22} + 1}{2}\right)^{\frac{1}{2}}, \quad |a_3| = \left(\frac{c_{33} + 1}{2}\right)^{\frac{1}{2}}. \tag{1.29}$$

Since a rotation of either π or $-\pi$ about the axes **a** and $-$**a** are equivalent, we can arbitrarily choose the sign of one of a_1, a_2 or a_3. Some possible solutions are

- If $\phi = \pm\pi$ and $|a_1| > 0$,

$$\begin{aligned} a_1 &= |a_1|, \\ a_2 &= \text{sign}\,(c_{12})\,|a_2|, \\ a_3 &= \text{sign}\,(c_{13})\,|a_3|. \end{aligned}$$

- If $\phi = \pm\pi$ and $|a_2| > 0$,

$$\begin{aligned} a_1 &= \text{sign}\,(c_{12})\,|a_1|, \\ a_2 &= |a_2|, \\ a_3 &= \text{sign}\,(c_{23})\,|a_3|. \end{aligned}$$

- If $\phi = \pm\pi$ and $|a_3| > 0$,

$$\begin{aligned} a_1 &= \text{sign}\,(c_{13})\,|a_1|, \\ a_2 &= \text{sign}\,(c_{23})\,|a_2|, \\ a_3 &= |a_3|. \end{aligned}$$

When $\phi = 0$, the axis **a** cannot be determined uniquely. Physically this makes sense, since a zero rotation about any axis results in the same orientation.

1.3.3 Euler Angles

In the preceeding section we have found that the rotation matrix can be described by three parameters. There are in fact many parameterizations that could be used. A common set of parameters is known as the Euler Angles. Euler Angles describe three successive principal rotations. For example, a possible sequence is:

1. A rotation ψ about the original z-axis (called a "yaw" rotation).
2. A rotation θ about the intermediate y-axis (called a "pitch" rotation).
3. A rotation ϕ about the transformed x-axis (called a "roll" rotation).

This is a very common choice in aerospace applications, and is called a 3-2-1 attitude sequence, and is depicted in Figure 1.15. The terminology relates to the order of rotations. A principal z-axis (labeled 3) rotation is first, followed by a principal y-axis (labeled 2) rotation, followed

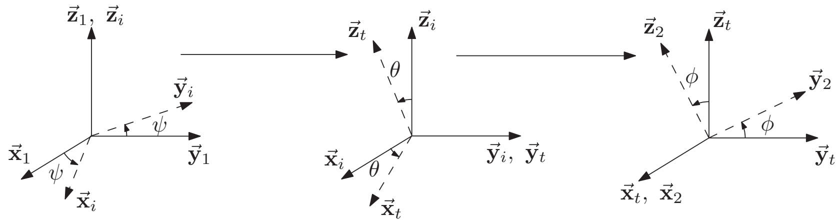

Figure 1.15 3-2-1 Euler rotation sequence

by a principal x-axis (labeled 1) rotation. In this case, the rotation matrix from frame $\mathcal{F}_1$ to frame $\mathcal{F}_2$ is given by

$$\begin{aligned}\mathbf{C}_{21}(\phi, \theta, \psi) &= \mathbf{C}_x(\phi)\mathbf{C}_y(\theta)\mathbf{C}_z(\psi)\\ &= \begin{bmatrix} c_\theta c_\psi & c_\theta s_\psi & -s_\theta \\ s_\phi s_\theta c_\psi - c_\phi s_\psi & s_\phi s_\theta s_\psi + c_\phi c_\psi & s_\phi c_\theta \\ c_\phi s_\theta c_\psi + s_\phi s_\psi & c_\phi s_\theta s_\psi - s_\phi c_\psi & c_\phi c_\theta \end{bmatrix}\end{aligned} \tag{1.30}$$

where $s_b = \sin b$ and $c_b = \cos b$.

An unfortunate consequence of using three parameters to describe the rotation matrix is that a singularity occurs. It can be shown that this will occur for any three-dimensional parameterization of the rotation matrix. For the 3-2-1 sequence above, the singularity occurs when the pitch angle is $\theta = \pm 90^\circ$. For example, when $\theta = 90^\circ$, the rotation matrix becomes

$$\mathbf{C}_{21}(\phi, 90^\circ, \psi) = \begin{bmatrix} 0 & 0 & -1 \\ \sin(\phi - \psi) & \cos(\phi - \psi) & 0 \\ \cos(\phi - \psi) & -\sin(\phi - \psi) & 0 \end{bmatrix}.$$

Physically, at the singularity the first and third rotations in the sequence occur about the same axis. In this case, the roll and yaw angles (ϕ and ψ) are associated with the same rotation, and cannot be determined uniquely. Outside of the singularity however, we can uniquely determine the angles from the rotation matrix. Denoting the rotation matrix by

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix},$$

from equation (1.30), we see that

$$\begin{aligned}\phi &= \tan^{-1}(c_{23}/c_{33}),\\ \theta &= -\sin^{-1}(c_{13}),\\ \psi &= \tan^{-1}(c_{12}/c_{11}).\end{aligned}$$

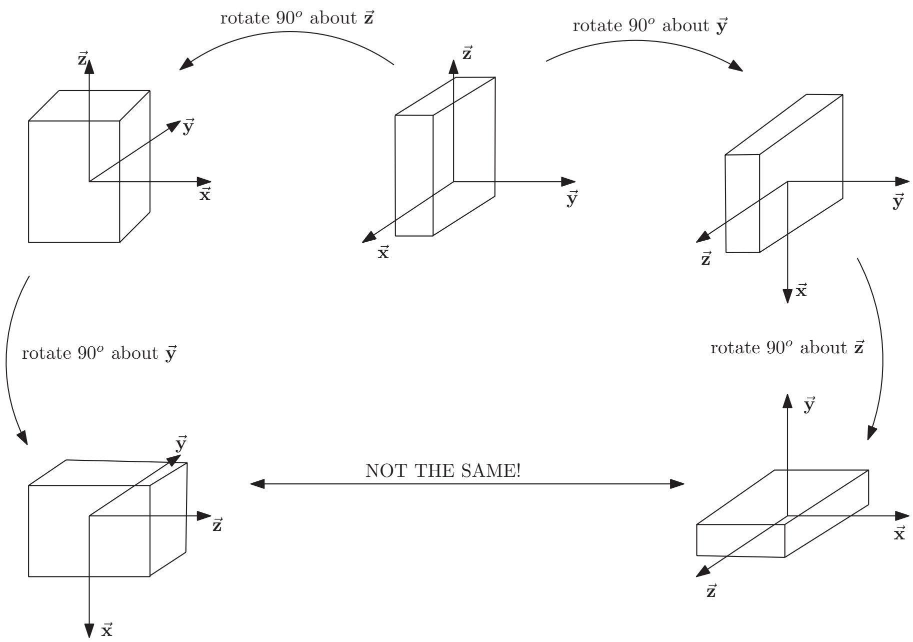

Figure 1.16 Non commuting rotations

The signs of c_{23}, c_{33} and c_{11}, c_{12} determine the quadrants of ϕ and ψ respectively.

Note that it is very important to specify the order of rotations (e.g. 3-2-1), since as seen in Figure 1.16, **rotations do NOT commute!**

1.3.4 Quaternions

As seen in the previous section, an unfortunate consequence of using a three-dimensional parameterization of the rotation matrix is the existence of a singularity, that is, a rotation for which the three parameters cannot be uniquely determined. To overcome this problem, it is necessary to add a redundant parameter to the parameterization. One possibility is the principal axis and angle of rotation. A more common and very useful choice of parameterization is the quaternion (also known as Euler parameters). The advantage of using quaternions, unlike axis-angle parameters is that the expression for the rotation matrix, and kinematics are purely algebraic (they contain no trigonometric functions). This makes them very useful and efficient for computational purposes.

To define the quaternion, we first need to reexamine the rotation matrix in terms of the principal axis $\mathbf{a}$ and angle ϕ of rotation. From (1.26), we have

$$\mathbf{C} = \cos\phi\mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi\mathbf{a}^{\times}.$$

Let us now make use of the trigonometric identities

$$\sin\phi = 2\sin\frac{\phi}{2}\cos\frac{\phi}{2}, \quad \cos\phi = 2\cos^2\frac{\phi}{2} - 1 = 1 - 2\sin^2\frac{\phi}{2}.$$

We can now rewrite the rotation matrix as

$$\mathbf{C} = \left(2\cos^2\frac{\phi}{2} - 1\right) + 2\sin^2\frac{\phi}{2}\mathbf{a}\mathbf{a}^T - 2\sin\frac{\phi}{2}\cos\frac{\phi}{2}\mathbf{a}^\times.$$

Based on this, we now define a vector and scalar part of the quaternion as

$$\boldsymbol{\epsilon} = \mathbf{a}\sin\frac{\phi}{2}, \quad \eta = \cos\frac{\phi}{2}, \tag{1.31}$$

respectively. We immediately see that the quaternion satisfies a unit magnitude constraint

$$\boldsymbol{\epsilon}^T\boldsymbol{\epsilon} + \eta^2 = 1. \tag{1.32}$$

The rotation matrix in terms of the quaternion is given by

$$\mathbf{C} = \left(2\eta^2 - 1\right)\mathbf{1} + 2\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T - 2\eta\boldsymbol{\epsilon}^\times. \tag{1.33}$$

1.3.4.1 Quaternion from Rotation Matrix

Let

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}, \text{ and } \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}.$$

Then, taking the trace of $\mathbf{C}$ in (1.33), we obtain

$$\text{trace}\,[\mathbf{C}] = 3\left(2\eta^2 - 1\right) + 2(\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2) = 4\eta^2 - 1, \tag{1.34}$$

Note that we have made use of (1.32) to achieve the last equality. Therefore, we obtain

$$\eta = \pm\frac{(\text{trace}\,[\mathbf{C}] + 1)^{\frac{1}{2}}}{2}. \tag{1.35}$$

An important point to note is that we may choose either sign for η. It can easily be seen from (1.33) that the quaternions $(\boldsymbol{\epsilon}, \eta)$ and $(-\boldsymbol{\epsilon}, -\eta)$ correspond to the same rotation matrix. Physically, a positive value for η corresponds to a principal angle in the range $-\pi < \phi < \pi$. A negative value for η corresponds to a principal angle in the range $-2\pi < \phi < -\pi$ or $\pi < \phi < 2\pi$. As shown in Figure 1.17, by appropriately choosing the direction of the principal

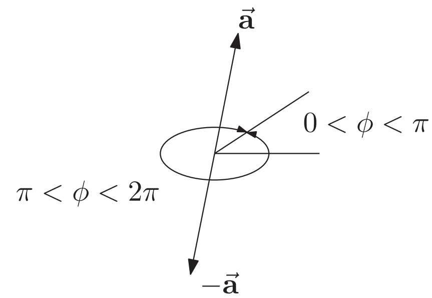

Figure 1.17 Equivalent rotations

axis of rotation, these rotations are equivalent. Having found η, we can now obtain $\boldsymbol{\epsilon}$. From (1.33),

$$\mathbf{C}^T - \mathbf{C} = 4\eta\boldsymbol{\epsilon}^{\times}, \tag{1.36}$$

and we see that if $\eta \neq 0$, then the corresponding $\boldsymbol{\epsilon}$ is given by

$$\begin{aligned} \epsilon_1 &= \frac{c_{23} - c_{32}}{4\eta}, \\ \epsilon_2 &= \frac{c_{31} - c_{13}}{4\eta}, \\ \epsilon_3 &= \frac{c_{12} - c_{21}}{4\eta}. \end{aligned} \tag{1.37}$$

When $\eta = 0$, the rotation matrix becomes

$$\mathbf{C} = -\mathbf{1} + 2\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T.$$

Physically, $\eta = 0$ corresponds to a principal rotation angle $\phi = \pm\pi$. We now find that if $\eta = 0$,

$$\begin{aligned} |\epsilon_1| &= \left(\frac{c_{11}+1}{2}\right)^{\frac{1}{2}} = |a_1|, \quad |\epsilon_2| = \left(\frac{c_{22}+1}{2}\right)^{\frac{1}{2}} = |a_2|, \quad |\epsilon_3| \\ &= \left(\frac{c_{33}+1}{2}\right)^{\frac{1}{2}} = |a_3|. \end{aligned} \tag{1.38}$$

Since a rotation of either π or $-\pi$ about the axes $\mathbf{a}$ and $-\mathbf{a}$ are equivalent, we can arbitrarily choose the sign of one of ϵ_1, ϵ_2 or ϵ_3. Some possible solutions are

- If $\eta = 0$ and $|\epsilon_1| > 0$,

$$\begin{aligned}\epsilon_1 &= |\epsilon_1|,\\ \epsilon_2 &= \text{sign}\,(c_{12})\,|\epsilon_2|,\\ \epsilon_3 &= \text{sign}\,(c_{13})\,|\epsilon_3|.\end{aligned}$$

- If $\eta = 0$ and $|\epsilon_2| > 0$,

$$\begin{aligned}\epsilon_1 &= \text{sign}\,(c_{12})\,|\epsilon_1|,\\ \epsilon_2 &= |\epsilon_2|,\\ \epsilon_3 &= \text{sign}\,(c_{23})\,|\epsilon_3|.\end{aligned}$$

- If $\eta = 0$ and $|\epsilon_3| > 0$,

$$\begin{aligned}\epsilon_1 &= \text{sign}\,(c_{13})\,|\epsilon_1|,\\ \epsilon_2 &= \text{sign}\,(c_{23})\,|\epsilon_2|,\\ \epsilon_3 &= |\epsilon_3|.\end{aligned}$$

We now note some useful points. First of all, the quaternion parameterization of the zero rotation ($\phi = 0$ or $\mathbf{C} = \mathbf{1}$) is $(\mathbf{0}, \pm 1)$. Next, from (1.33) is immediately obvious that if $(\boldsymbol{\epsilon}, \eta)$ is the quaternion parameterization of $\mathbf{C}$, then $(-\boldsymbol{\epsilon}, \eta)$ is a quaternion parameterization of the inverse rotation, $\mathbf{C}^T$.

1.3.4.2 Successive Rotations

Let us now consider three reference frames $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$, with associated rotation matrices $\mathbf{C}_{21}$, $\mathbf{C}_{32}$ and $\mathbf{C}_{31}$, such that

$$\mathbf{C}_{31} = \mathbf{C}_{32}\mathbf{C}_{21}.$$

We parameterize each rotation matrix with a quaternion, $\mathbf{C}_{21} = \mathbf{C}(\boldsymbol{\epsilon}_1, \eta_1)$, $\mathbf{C}_{32} = \mathbf{C}(\boldsymbol{\epsilon}_2, \eta_2)$ and $\mathbf{C}_{31} = \mathbf{C}(\boldsymbol{\epsilon}_3, \eta_3)$, such that

$$\begin{aligned}\mathbf{C}_{21} &= \left(2\eta_1^2 - 1\right)\mathbf{1} + 2\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T - 2\eta_1\boldsymbol{\epsilon}_1^\times,\\ \mathbf{C}_{32} &= \left(2\eta_2^2 - 1\right)\mathbf{1} + 2\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T - 2\eta_2\boldsymbol{\epsilon}_2^\times,\\ \mathbf{C}_{31} &= \left(2\eta_3^2 - 1\right)\mathbf{1} + 2\boldsymbol{\epsilon}_3\boldsymbol{\epsilon}_3^T - 2\eta_3\boldsymbol{\epsilon}_3^\times.\end{aligned}$$

We shall now find $(\boldsymbol{\epsilon}_3, \eta_3)$ as a function of $(\boldsymbol{\epsilon}_1, \eta_1)$ and $(\boldsymbol{\epsilon}_2, \eta_2)$. The resulting expressions are very useful, since they allow the computation of the new quaternion without having to first form the rotation matrix and then extracting it.

To address this problem, we note that the scalar and vector parts of the quaternion satisfy (1.34) and (1.36) respectively. Let us now compute $\mathbf{C}_{31}$. We have,

$$\begin{aligned}\mathbf{C}_{31} &= \left[\left(2\eta_2^2-1\right)\mathbf{1}+2\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T-2\eta_2\boldsymbol{\epsilon}_2^\times\right]\left[\left(2\eta_1^2-1\right)\mathbf{1}+2\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T-2\eta_1\boldsymbol{\epsilon}_1^\times\right]\\ &= \left(2\eta_2^2-1\right)\left(2\eta_1^2-1\right)\mathbf{1}+2\left(2\eta_2^2-1\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T-2\eta_1\left(2\eta_2^2-1\right)\boldsymbol{\epsilon}_1^\times\\ &\quad+2\left(2\eta_1^2-1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T+4\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T-4\eta_1\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times\\ &\quad-2\eta_2\left(2\eta_1^2-1\right)\boldsymbol{\epsilon}_2^\times-4\eta_2\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T+4\eta_1\eta_2\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1^\times.\end{aligned}$$

Noting that

$$\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T=\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T$$

and

$$\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1^\times=\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T-\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\mathbf{1},$$

we can rewrite the rotation matrix as

$$\begin{aligned}\mathbf{C}_{31} &= \left(2\eta_2^2-1\right)\left(2\eta_1^2-1\right)\mathbf{1}+2\left(2\eta_2^2-1\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T-2\eta_1\left(2\eta_2^2-1\right)\boldsymbol{\epsilon}_1^\times\\ &\quad+2\left(2\eta_1^2-1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T+4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T-4\eta_1\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times\\ &\quad-2\eta_2\left(2\eta_1^2-1\right)\boldsymbol{\epsilon}_2^\times-4\eta_2\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T+4\eta_1\eta_2\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T-4\eta_1\eta_2\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\mathbf{1}.\end{aligned}\tag{1.39}$$

First, we shall find η_3. Taking the trace of $\mathbf{C}_{31}$ in (1.39), we obtain (note that we make use of the fact that for any two matrices $\mathbf{A}$ and $\mathbf{B}$, the trace satisfies trace $[\mathbf{AB}]=$ trace $[\mathbf{BA}]$, provided the multiplication on the right makes sense),

$$\begin{aligned}\text{trace}\left[\mathbf{C}_{31}\right] &= 3\left(2\eta_2^2-1\right)\left(2\eta_1^2-1\right)+2\left(2\eta_2^2-1\right)\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1+2\left(2\eta_1^2-1\right)\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\\ &\quad+4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)^2-4\eta_1\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2-4\eta_2\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\\ &\quad+4\eta_1\eta_2\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)-12\eta_1\eta_2\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right).\end{aligned}$$

Making use of the facts $\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1=1-\eta_1^2$, $\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2=1-\eta_2^2$, and noting that $\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1=\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2=0$, we can reduce this to

$$\begin{aligned}\text{trace}\left[\mathbf{C}_{31}\right] &= 4\left(\eta_1\eta_2\right)^2-8\left(\eta_1\eta_2\right)\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)+4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)^2-1,\\ &= 4\left[\eta_1\eta_2-\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right]^2-1.\end{aligned}$$

Comparing this to (1.34), and recalling that we are free to choose the sign of η_3, we see that we can choose

$$\eta_3=\eta_1\eta_2-\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1.\tag{1.40}$$

Now, let us find the vector part of the quaternion, $\boldsymbol{\epsilon}_3$. From (1.36), we see that we must compute $\mathbf{C}_{31}^T - \mathbf{C}_{31}$. From (1.39) we have,

$$\begin{aligned}\mathbf{C}_{31}^T - \mathbf{C}_{31} &= 4\eta_1\left(2\eta_2^2 - 1\right)\boldsymbol{\epsilon}_1^\times - 4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right] \\ &\quad + 4\eta_1\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\right] + 4\eta_2\left(2\eta_1^2 - 1\right)\boldsymbol{\epsilon}_2^\times \\ &\quad + 4\eta_2\left[\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times + \boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\right] + 4\eta_1\eta_2\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right].\end{aligned} \tag{1.41}$$

We now derive a new column matrix identity, which shall be very useful. We have

$$\left(\mathbf{a}^\times\mathbf{b}\right)^\times\mathbf{c} = -\mathbf{c}^\times\mathbf{a}^\times\mathbf{b} = -\left[\left(\mathbf{b}^T\mathbf{c}\right)\mathbf{a} - \left(\mathbf{a}^T\mathbf{c}\right)\mathbf{b}\right] = \left[\mathbf{b}\mathbf{a}^T - \mathbf{a}\mathbf{b}^T\right]\mathbf{c}.$$

Since this must hold for all $\mathbf{c} \in \mathbf{R}^3$, we have

$$\left(\mathbf{a}^\times\mathbf{b}\right)^\times = \mathbf{b}\mathbf{a}^T - \mathbf{a}\mathbf{b}^T. \tag{1.42}$$

Applying this identity, we have

$$\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T = \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times.$$

Making use of this and noting that $2\eta^2 - 1 = \eta^2 - \boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$, we can rewrite (1.41) as

$$\begin{aligned}\mathbf{C}_{31}^T - \mathbf{C}_{31} &= 4\eta_1\eta_2\left(\eta_2\boldsymbol{\epsilon}_1^\times + \eta_1\boldsymbol{\epsilon}_2^\times\right) + 4\left(\eta_1\eta_2 - \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times \\ &\quad - 4\eta_1\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1^\times - 4\eta_2\left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2^\times \\ &\quad + 4\eta_1\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\right] + 4\eta_2\left[\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times + \boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\right].\end{aligned} \tag{1.43}$$

Now, making use of the identity (1.42), we shall reduce the square bracket terms. We have

$$\begin{aligned}\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T &= \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_2^T - \boldsymbol{\epsilon}_2\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T \\ &= \left(\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times,\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times + \boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T &= \left(\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\left(\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\right)^T \\ &= \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\right)^\times.\end{aligned}$$

Making use of the identity

$$\mathbf{a}^\times\mathbf{b}^\times\mathbf{c} = (\mathbf{a}^T\mathbf{c})\mathbf{b} - (\mathbf{a}^T\mathbf{b})\mathbf{c},$$

we can reduce these to

$$\begin{aligned}\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T &= \left(\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times \\ &= \left[\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1 - \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_2\right]^\times \\ &= \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1^\times - \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_2^\times,\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times + \boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T &= \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\right)^\times \\ &= \left[\left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2 - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_1\right]^\times \\ &= \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2^\times - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_1^\times.\end{aligned}$$

Making use of these in (1.43), we finally obtain

$$\mathbf{C}_{31}^T - \mathbf{C}_{31} = 4\left(\eta_1\eta_2 - \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right]^\times.$$

Recognizing η_3 from before, we have

$$\mathbf{C}_{31}^T - \mathbf{C}_{31} = 4\eta_3\left[\eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right]^\times.$$

Comparing this to (1.36), we see that if $\eta_3 \neq 0$, then we can find the corresponding vector part of the quaternion as

$$\boldsymbol{\epsilon}_3 = \eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2. \tag{1.44}$$

Now all that remains is to check the vector part of the quaternion when $\eta_3 = 0$. We note that when $\eta_3 = 0$, $\mathbf{C}_{31}$ becomes

$$\mathbf{C}_{31} = -\mathbf{1} + 2\boldsymbol{\epsilon}_3\boldsymbol{\epsilon}_3^T,$$

which is symmetric. In particular, we have

$$\mathbf{C}_{31} + \mathbf{C}_{31}^T = -2\mathbf{1} + 4\boldsymbol{\epsilon}_3\boldsymbol{\epsilon}_3^T. \tag{1.45}$$

From (1.39), we have

$$\begin{aligned}\mathbf{C}_{31} + \mathbf{C}_{31}^T = {} & 2\left(2\eta_2^2 - 1\right)\left(2\eta_1^2 - 1\right)\mathbf{1} + 4\left(2\eta_2^2 - 1\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T \\ & + 4\left(2\eta_1^2 - 1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T + 4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T + \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right] \\ & - 4\eta_1\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times - \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\right] - 4\eta_2\left[\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times\right] \\ & 4\eta_1\eta_2\left[\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T + \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T\right] - 8\eta_1\eta_2\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\mathbf{1}\end{aligned} \tag{1.46}$$

Let us now examine this term by term. Noting from (1.40) that $\eta_1\eta_2 = \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1$ when $\eta_3 = 0$, we have

$$2\left(2\eta_2^2 - 1\right)\left(2\eta_1^2 - 1\right) - 8\eta_1\eta_2\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right) = -2 + \left[4 - 4\left(\eta_1^2 + \eta_2^2\right)\right].$$

Next, since $\eta^2 + \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = 1$, we have

$$\begin{aligned} 4\left(2\eta_2^2 - 1\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T &= 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T - 4\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T, \\ 4\left(2\eta_1^2 - 1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T &= 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T - 4\left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T. \end{aligned}$$

We can also rewrite the following factors

$$\begin{aligned} -4\eta_1\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times - \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\right] &= 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T + 4\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T, \\ -4\eta_2\left[\boldsymbol{\epsilon}_2^\times\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2^\times\right] &= 4\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T + 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T, \\ 4\eta_1\eta_2\left[\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T + \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T\right] &= 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T + 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T. \end{aligned}$$

Making use of these expansions, we can rewrite (1.46) as

$$\begin{aligned} \mathbf{C}_{31} + \mathbf{C}_{31}^T = -2\mathbf{1} &+ 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T + 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T \\ &+ 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T + 4\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T \\ &+ 4\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T + 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T \\ &+ 4\left(\eta_2\boldsymbol{\epsilon}_1\right)\left(\eta_1\boldsymbol{\epsilon}_2\right)^T + 4\left(\eta_1\boldsymbol{\epsilon}_2\right)\left(\eta_2\boldsymbol{\epsilon}_1\right)^T \\ &+ 4\left[\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T + \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right] - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T - \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T\right. \\ &\left.+ \left[1 - \left(\eta_1^2 + \eta_2^2\right)\right]\mathbf{1}\right] \end{aligned} \tag{1.47}$$

Now we just need to take care of the final bracketed term. Making use of the identity $\mathbf{a}\mathbf{a}^T = \mathbf{a}^\times\mathbf{a}^\times + \mathbf{a}^T\mathbf{a}\mathbf{1}$, we have

$$\begin{aligned} \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T &= \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^\times + \left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\mathbf{1} \\ &= \left(\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right)\left(\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right) - \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\mathbf{1} \\ &= \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T - \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T + \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T \\ &\quad - \boldsymbol{\epsilon}_2^T\left(\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T - \boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\mathbf{1}\right)\boldsymbol{\epsilon}_2\mathbf{1} \\ &= \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T + \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right] - \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T \\ &\quad + \left[\left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right) - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)^2\right]\mathbf{1}. \end{aligned}$$

Since $\eta_3 = 0$ means that $\eta_1\eta_2 = \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1$, and making use of the unit magnitude constraint $\eta^2 + \boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = 1$, the last term in brackets becomes

$$\left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right) - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)^2 = 1 - \left(\eta_1^2 + \eta_2^2\right).$$

Substituting this back in, we have

$$\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)\left(\boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right)^T = \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1\right)\left[\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_1^T + \boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2^T\right] - \left(\boldsymbol{\epsilon}_1^T\boldsymbol{\epsilon}_1\right)\boldsymbol{\epsilon}_2\boldsymbol{\epsilon}_2^T - \left(\boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_2\right)\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T + \left[1 - \left(\eta_1^2 + \eta_2^2\right)\right]\mathbf{1}.$$

We can now see that this is precisely the last bracketed term in (1.47). Therefore, we can finally write

$$\mathbf{C}_{31} + \mathbf{C}_{31}^T = -2\mathbf{1} + 4\left[\eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right]\left[\eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2\right]^T. \tag{1.48}$$

We shall now make use of the fact that for two column matrices $\mathbf{a}$ and $\mathbf{b}$, if $\mathbf{a}\mathbf{a}^T = \mathbf{b}\mathbf{b}^T$, then $\mathbf{b} = \pm\mathbf{a}$. When $\eta_3 = 0$, the sign of the vector part of the quaternion does not matter. Comparing (1.48) with (1.45), we see therefore that we may take

$$\boldsymbol{\epsilon}_3 = \eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2,$$

which is the same as for the case $\eta_3 \neq 0$.

In summary, given the rotation matrices $\mathbf{C}_{21} = \mathbf{C}(\boldsymbol{\epsilon}_1, \eta_1)$, $\mathbf{C}_{32} = \mathbf{C}(\boldsymbol{\epsilon}_2, \eta_2)$ and $\mathbf{C}_{31} = \mathbf{C}(\boldsymbol{\epsilon}_3, \eta_3)$ and their respective quaternion parameterizations, given that $\mathbf{C}_{31} = \mathbf{C}_{32}\mathbf{C}_{21}$, the associated quaternions are related by

$$\begin{aligned}\boldsymbol{\epsilon}_3 &= \eta_2\boldsymbol{\epsilon}_1 + \eta_1\boldsymbol{\epsilon}_2 + \boldsymbol{\epsilon}_1^\times\boldsymbol{\epsilon}_2,\\ \eta_3 &= \eta_1\eta_2 - \boldsymbol{\epsilon}_2^T\boldsymbol{\epsilon}_1.\end{aligned} \tag{1.49}$$

1.4 Derivatives of Vectors

Now that we have examined physical vectors and their representations with respect to reference frames, we are in a position to look at the evolution of physical vectors with respect to time. Reference frames are crucial for this, since the evolution of a physical vector depends entirely on the point of view of an observer. For example, an observer attached to one reference frame will see a different evolution of the physical vector compared with an observer attached to a different reference frame. For this reason, when we talk about temporal derivatives of a physical vector, we will talk about the derivative as seen in a particular reference frame.

As such, let us consider a reference frame $\mathcal{F}_1$ defined by the unit vectors $\vec{\mathbf{x}}_1$, $\vec{\mathbf{y}}_1$ and $\vec{\mathbf{z}}_1$. Given a physical vector

$$\vec{\mathbf{r}} = \vec{\mathcal{F}}_1^T\mathbf{r}_1,$$

the time derivative of $\vec{\mathbf{r}}$ as seen in reference frame $\mathcal{F}_1$ is defined to be

$$\dot{\vec{\mathbf{r}}} \triangleq \lim_{\delta t \to 0} \frac{\delta \vec{\mathbf{r}}}{\delta t},$$

where

$$\delta \vec{\mathbf{r}} \triangleq \vec{\mathcal{F}}_1^T \left(\mathbf{r}_1(t + \delta t) - \mathbf{r}_1(t) \right).$$

Therefore, the time derivative of $\vec{\mathbf{r}}$ as seen in reference frame $\mathcal{F}_1$ is given by

$$\dot{\vec{\mathbf{r}}} = \vec{\mathcal{F}}_1^T \dot{\mathbf{r}}_1.$$

As a simple consequence of this, if we multiply the vector $\vec{\mathbf{r}}$ by a time-varying scalar a, we have

$$\begin{aligned} \overbrace{(a\vec{\mathbf{r}})}^{\cdot} &= \vec{\mathcal{F}}_1^T (a\dot{\mathbf{r}}_1) \\ &= \vec{\mathcal{F}}_1^T (\dot{a}\mathbf{r}_1 + a\dot{\mathbf{r}}_1) \\ &= \dot{a}\vec{\mathbf{r}} + a\dot{\vec{\mathbf{r}}}. \end{aligned}$$

Similarly, for sums of vectors,

$$\begin{aligned} \overbrace{\left(\vec{\mathbf{a}} + \vec{\mathbf{b}}\right)}^{\cdot} &= \vec{\mathcal{F}}_1^T \left[\dot{\mathbf{a}}_1 + \dot{\mathbf{b}}_1\right] \\ &= \dot{\vec{\mathbf{a}}} + \dot{\vec{\mathbf{b}}}. \end{aligned}$$

We also see immediately that the time derivatives of the unit vectors defining frame $\mathcal{F}_1$ are all $\vec{\mathbf{0}}$, and that therefore

$$\dot{\vec{\mathcal{F}}}_1 = \begin{bmatrix} \vec{\mathbf{0}} \\ \vec{\mathbf{0}} \\ \vec{\mathbf{0}} \end{bmatrix}.$$

Finally, we can obtain product rules for scalar- and cross-products of vectors:

$$\begin{aligned} \frac{d}{dt}\left(\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}\right) &= \frac{d}{dt}\left[\mathbf{a}^T\mathbf{b}\right] \\ &= \dot{\mathbf{a}}^T\mathbf{b} + \mathbf{a}^T\dot{\mathbf{b}} \\ &= \dot{\vec{\mathbf{a}}} \cdot \vec{\mathbf{b}} + \vec{\mathbf{a}} \cdot \dot{\vec{\mathbf{b}}}. \end{aligned}$$

and

$$\overbrace{\left(\vec{\mathbf{a}} \times \vec{\mathbf{b}}\right)}^{\cdot} = \vec{\mathcal{F}}_1^T \frac{d}{dt}\left[\mathbf{a}^\times \mathbf{b}\right]$$
$$= \vec{\mathcal{F}}_1^T \left[\dot{\mathbf{a}}^\times \mathbf{b} + \mathbf{a}^\times \dot{\mathbf{b}}\right]$$
$$= \dot{\vec{\mathbf{a}}} \times \vec{\mathbf{b}} + \vec{\mathbf{a}} \times \dot{\vec{\mathbf{b}}}.$$

1.4.1 Angular Velocity

Let us consider a second reference frame $\mathcal{F}_2$, which is rotating with respect to $\mathcal{F}_1$ with angular velocity $\vec{\boldsymbol{\omega}}_{21}$, as shown in Figure 1.18. The magnitude of $\vec{\boldsymbol{\omega}}_{21}$, $|\vec{\boldsymbol{\omega}}_{21}|$ is the rate of rotation, and the direction $\vec{\boldsymbol{\omega}}_{21}/|\vec{\boldsymbol{\omega}}_{21}|$ is the *instantaneous* axis of rotation. To understand what this means, consider the reference frame $\mathcal{F}_2$ at two times t and $t + \delta t$ as seen in reference frame $\mathcal{F}_1$. We know that the transformation from $\mathcal{F}_2(t)$ to $\mathcal{F}_2(t + \delta t)$ is a rotation about some axis $\vec{\mathbf{a}}(\delta t)$ through an angle $\phi(\delta t)$. We formally define the angular velocity to be

$$\vec{\boldsymbol{\omega}}_{21} = \lim_{\delta t \to 0} \vec{\mathbf{a}}(\delta t)\frac{\phi(\delta t)}{\delta t}.$$

Let us now consider the time-derivative of an arbitrary physical vector $\vec{\mathbf{v}}$ rotating with angular velocity $\vec{\boldsymbol{\omega}}$ as seen in reference frame $\mathcal{F}_1$. First, let us examine a finite rotation of the vector $\vec{\mathbf{v}}$ about a unit vector $\vec{\mathbf{a}}$ through an angle ϕ as seen in $\mathcal{F}_1$.

We have seen in (1.25) that the rotated vector is given in $\mathcal{F}_1$ coordinates by

$$\vec{\mathbf{v}}_{rot} = \vec{\mathcal{F}}_1^T \left[\cos\phi \mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T + \sin\phi \mathbf{a}^\times\right] \mathbf{v}. \tag{1.50}$$

Now, let δt be very small such that the angle $\phi(\delta t)$ is very small, and $\sin\phi \approx \phi$ and $\cos\phi \approx 1$. Under these approximations, we have

$$\vec{\mathbf{v}}(t + \delta t) = \vec{\mathcal{F}}_1^T \left[\mathbf{1} + \mathbf{a}^\times \phi\right] \mathbf{v}.$$

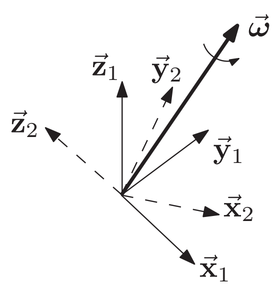

Figure 1.18 Angular velocity

From this we find that

$$\vec{\mathbf{v}}(t+\delta t)-\vec{\mathbf{v}}(t)=\vec{\mathcal{F}}_1^T\left[\mathbf{a}^{\times}\phi\right]\mathbf{v},$$

which leads to

$$\begin{aligned}\dot{\vec{\mathbf{v}}}(t)=\lim_{\delta t\to 0}\frac{\vec{\mathbf{v}}(t+\delta t)-\vec{\mathbf{v}}(t)}{\delta t}&=\vec{\mathcal{F}}_1^T\boldsymbol{\omega}^{\times}\mathbf{v},\\&=\vec{\boldsymbol{\omega}}\times\vec{\mathbf{v}}.\end{aligned}\tag{1.51}$$

Let us now return to the consideration of the two reference frames $\mathcal{F}_1$ and $\mathcal{F}_2$, with $\mathcal{F}_2$ rotating with angular velocity $\vec{\boldsymbol{\omega}}_{21}$ with respect to $\mathcal{F}_1$. Using the above result, we have that the time-derivatives as seen in $\mathcal{F}_1$ of the unit vectors defining $\mathcal{F}_2$ are

$$\dot{\vec{\mathbf{x}}}_2=\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathbf{x}}_2,\quad \dot{\vec{\mathbf{y}}}_2=\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathbf{y}}_2,\quad \dot{\vec{\mathbf{z}}}_2=\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathbf{z}}_2.$$

This can be written compactly as

$$\dot{\vec{\mathcal{F}}}_2^T=\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathcal{F}}_2^T.$$

As we have already discussed, observers in the two reference frames do not see the same motion, because of their own relative motions. We have already denoted the time-derivatives of vectors as seen in frame $\mathcal{F}_1$ by $(\dot{\ })$. Let us denote time derivatives of vectors as seen in frame $\mathcal{F}_2$ by $(\mathring{\ })$. We now consider the vector

$$\vec{\mathbf{r}}=\vec{\mathcal{F}}_1^T\mathbf{r}_1=\vec{\mathcal{F}}_2^T\mathbf{r}_2.$$

By definition, the time derivatives of the vector as seen in each of the frames are

$$\dot{\vec{\mathbf{r}}}=\vec{\mathcal{F}}_1^T\dot{\mathbf{r}}_1,\quad \mathring{\vec{\mathbf{r}}}=\vec{\mathcal{F}}_2^T\dot{\mathbf{r}}_2,$$

respectively. Let us now obtain a relationship between the time-derivatives of vectors as seen in the two frames. Making use of the rules for vector differentiation we obtained earlier, we have

$$\begin{aligned}\dot{\vec{\mathbf{r}}}&=\vec{\mathcal{F}}_2^T\dot{\mathbf{r}}_2+\dot{\vec{\mathcal{F}}}_2^T\mathbf{r}_2\\&=\mathring{\vec{\mathbf{r}}}+\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathcal{F}}_2^T\mathbf{r}_2\\&=\mathring{\vec{\mathbf{r}}}+\vec{\boldsymbol{\omega}}_{21}\times\vec{\mathbf{r}}.\end{aligned}\tag{1.52}$$

This very important relationship (1.52) is true for any vector $\vec{\mathbf{r}}$, and is referred to as the *Transport Theorem*. A very important application occurs when $\vec{\mathbf{r}}$ denotes the position (of say a spacecraft), $\mathcal{F}_1$ is a nonrotating inertial reference frame, and $\mathcal{F}_2$ is a frame that rotates with the body in question. Equation (1.52) also shows that provided two reference frames are not rotating with respect to each other (if $\vec{\boldsymbol{\omega}}_{21}=\vec{\mathbf{0}}$), then the time-derivative of a vector as seen in both frames is the same, that is $\dot{\vec{\mathbf{r}}}=\mathring{\vec{\mathbf{r}}}$.

Let us express the angular velocity in frame $\mathcal{F}_2$ as

$$\vec{\boldsymbol{\omega}}_{21} = \vec{\mathcal{F}}_2^T \boldsymbol{\omega}_{21}.$$

Making use of this, equation (1.52) becomes

$$\begin{aligned}\dot{\vec{\mathbf{r}}} = \vec{\mathcal{F}}_1^T \dot{\mathbf{r}}_1 &= \vec{\mathcal{F}}_2^T \left[\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_{21}^{\times} \mathbf{r}_2\right] \\ &= \vec{\mathcal{F}}_1^T \mathbf{C}_{12} \left[\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_{21}^{\times} \mathbf{r}_2\right],\end{aligned}$$

that is, the relationship between the coordinates of the time-derivatives with respect to the two frames is

$$\dot{\mathbf{r}}_1 = \mathbf{C}_{12} \left[\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_{21}^{\times} \mathbf{r}_2\right]. \tag{1.53}$$

We can make use of the above equation (1.53) to obtain the differential equation for the rotation matrix $\mathbf{C}_{21}$. To this end, let us consider an arbitrary constant vector $\vec{\mathbf{r}}$ as seen in frame $\mathcal{F}_1$, that is $\dot{\mathbf{r}}_1 = \mathbf{0}$. In this case, the equation (1.53) becomes

$$\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_{21}^{\times} \mathbf{r}_2 = \mathbf{0}$$

Making use of the transformation of coordinates relationship $\mathbf{r}_2 = \mathbf{C}_{21}\mathbf{r}_1$, this becomes

$$\left[\dot{\mathbf{C}}_{21} + \boldsymbol{\omega}_{21}^{\times} \mathbf{C}_{21}\right] \mathbf{r}_1 = \mathbf{0}.$$

Since the vector $\vec{\mathbf{r}}$ was arbitrary, this must hold for any $\mathbf{r}_1 \in R^3$. Therefore, we have the differential equation for the rotation matrix

$$\dot{\mathbf{C}}_{21} = -\boldsymbol{\omega}_{21}^{\times} \mathbf{C}_{21}. \tag{1.54}$$

This equation can be rewritten to give the angular velocity, when the rotation matrix is known as a function of time:

$$\boldsymbol{\omega}_{21}^{\times} = -\dot{\mathbf{C}}_{21} \mathbf{C}_{21}^T. \tag{1.55}$$

Finally, we can demonstrate that angular velocities are additive. Let us consider three reference frames $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$. Let frame $\mathcal{F}_2$ rotate with angular velocity $\vec{\boldsymbol{\omega}}_{21} = \vec{\mathcal{F}}_2^T \boldsymbol{\omega}_{21}$ with respect to $\mathcal{F}_1$, let frame $\mathcal{F}_3$ rotate with angular velocity $\vec{\boldsymbol{\omega}}_{32} = \vec{\mathcal{F}}_3^T \boldsymbol{\omega}_{32}$ with respect to $\mathcal{F}_2$, and let frame $\mathcal{F}_3$ rotate with angular velocity $\vec{\boldsymbol{\omega}}_{31} = \vec{\mathcal{F}}_3^T \boldsymbol{\omega}_{31}$ with respect to $\mathcal{F}_1$. Let $\mathbf{C}_{21}$, $\mathbf{C}_{32}$ and $\mathbf{C}_{31}$ be the associated rotation matrices respectively. From equation (1.54) we have

$$\dot{\mathbf{C}}_{21} = -\boldsymbol{\omega}_{21}^{\times} \mathbf{C}_{21}, \quad \dot{\mathbf{C}}_{32} = -\boldsymbol{\omega}_{32}^{\times} \mathbf{C}_{32}, \quad \dot{\mathbf{C}}_{31} = -\boldsymbol{\omega}_{31}^{\times} \mathbf{C}_{31}.$$

But, we also know from equation (1.19) that

$$\mathbf{C}_{31} = \mathbf{C}_{32} \mathbf{C}_{21}.$$

Differentiating this and using the above differential equations gives

$$\begin{aligned}-\boldsymbol{\omega}_{31}^{\times}\mathbf{C}_{31} = \dot{\mathbf{C}}_{31} &= \dot{\mathbf{C}}_{32}\mathbf{C}_{21} + \mathbf{C}_{32}\dot{\mathbf{C}}_{21} \\ &= -\boldsymbol{\omega}_{32}^{\times}\mathbf{C}_{32}\mathbf{C}_{21} - \mathbf{C}_{32}\boldsymbol{\omega}_{21}^{\times}\mathbf{C}_{21} \\ &= -\boldsymbol{\omega}_{32}^{\times}\mathbf{C}_{31} - \mathbf{C}_{32}\boldsymbol{\omega}_{21}^{\times}\mathbf{C}_{21}.\end{aligned}$$

Post-multiplying both sides by $\mathbf{C}_{31}^T$, and making use of the fact that $\mathbf{C}_{32} = \mathbf{C}_{31}\mathbf{C}_{21}^T$, this becomes

$$-\boldsymbol{\omega}_{31}^{\times} = -\boldsymbol{\omega}_{32}^{\times} - \mathbf{C}_{32}\boldsymbol{\omega}_{21}^{\times}\mathbf{C}_{32}^T.$$

Now, from equation (1.21) we have $(\mathbf{C}_{32}\boldsymbol{\omega}_{21})^{\times} = \mathbf{C}_{32}\boldsymbol{\omega}_{21}^{\times}\mathbf{C}_{32}^T$. Therefore, we obtain

$$-\boldsymbol{\omega}_{31}^{\times} = -\boldsymbol{\omega}_{32}^{\times} - (\mathbf{C}_{32}\boldsymbol{\omega}_{21})^{\times},$$

and

$$\boldsymbol{\omega}_{31} = \boldsymbol{\omega}_{32} + (\mathbf{C}_{32}\boldsymbol{\omega}_{21}).$$

Finally, recognizing that these are just the coordinates of the angular velocities expressed in $\mathcal{F}_3$, this leads to

$$\vec{\boldsymbol{\omega}}_{31} = \vec{\boldsymbol{\omega}}_{32} + \vec{\boldsymbol{\omega}}_{21}, \tag{1.56}$$

that is, **angular velocities are additive**. As a consequence, since a reference frame has zero angular velocity relative to itself, the angular velocity of frame $\mathcal{F}_1$ with respect to $\mathcal{F}_2$ is $\vec{\boldsymbol{\omega}}_{12} = -\vec{\boldsymbol{\omega}}_{21}$.

1.4.2 Angular Velocity in Terms of Euler Angle Rates

We have seen that any frame $\mathcal{F}_2$ can be obtained from another frame $\mathcal{F}_1$ through three successive principal rotations. The angles of these rotations are called Euler Angles. Making use of the fact that angular velocities are additive, we can now easily obtain an expression for the angular velocity in terms of the Euler Angle rates. We will demonstrate this using the 3-2-1 Euler rotation sequence presented previously. First, it will be useful to define some intermediate reference frames. Referring to Figure 1.15, let us denote frame $\mathcal{F}_i$ as the frame obtained from $\mathcal{F}_1$ by a principal rotation about the $\mathcal{F}_1$ z-axis. Next, let us denote frame $\mathcal{F}_t$ as the frame obtained from $\mathcal{F}_i$ by a principal rotation about the $\mathcal{F}_i$ y-axis. Finally, frame $\mathcal{F}_2$ is obtained from frame $\mathcal{F}_t$ by a principal rotation about the $\mathcal{F}_t$ x-axis.

It is now clear that frame $\mathcal{F}_i$ is rotating with respect to $\mathcal{F}_1$ with an angular velocity given by

$$\vec{\boldsymbol{\omega}}_{i1} = \vec{\mathcal{F}}_i^T \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} = \vec{\mathcal{F}}_2^T \mathbf{C}_x(\phi)\mathbf{C}_y(\theta) \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix},$$

frame $\mathcal{F}_t$ is rotating with respect to $\mathcal{F}_i$ with an angular velocity given by

$$\vec{\boldsymbol{\omega}}_{ti} = \vec{\mathcal{F}}_t^T \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} = \vec{\mathcal{F}}_2^T \mathbf{C}_x(\phi) \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix},$$

and frame $\mathcal{F}_2$ is rotating with respect to $\mathcal{F}_t$ with an angular velocity given by

$$\vec{\boldsymbol{\omega}}_{2t} = \vec{\mathcal{F}}_2^T \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix}.$$

Making use of the fact that the angular velocities are additive, we find that frame $\mathcal{F}_2$ rotates with respect to frame $\mathcal{F}_1$ with angular velocity

$$\vec{\boldsymbol{\omega}}_{21} = \vec{\boldsymbol{\omega}}_{2t} + \vec{\boldsymbol{\omega}}_{ti} + \vec{\boldsymbol{\omega}}_{i1}.$$

With $\vec{\boldsymbol{\omega}}_{21} = \vec{\mathcal{F}}_2^T \boldsymbol{\omega}_{21}$, we can obtain the relationship between the Euler Angle rates and the angular velocity in the coordinates of $\mathcal{F}_2$ as

$$\boldsymbol{\omega}_{21} = \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix} + \mathbf{C}_x(\phi) \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \mathbf{C}_x(\phi)\mathbf{C}_y(\theta) \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix}. \tag{1.57}$$

Written out in full, this is

$$\boldsymbol{\omega}_{21} = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix}. \tag{1.58}$$

This can be inverted to give the Euler Angle rates in terms of the angular velocity:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta \end{bmatrix} \boldsymbol{\omega}_{21}. \tag{1.59}$$

From this, it is clear that when the 3-2-1 rotation sequence is at the singularity (recall this is when $\theta = \pm 90^\circ$, the Euler Angle rates cannot be obtained from (1.59) due to the terms $\tan\theta$ and $\sec\theta$ being undefined.

1.4.3 Angular Velocity in Terms of Quaternion Rates

We have seen from (1.55) that if frame $\mathcal{F}_2$ is rotating with angular velocity $\vec{\boldsymbol{\omega}}_{21} = \vec{\mathcal{F}}_2^T \boldsymbol{\omega}_{21}$ relative to frame $\mathcal{F}_1$, then

$$\boldsymbol{\omega}_{21}^\times = -\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T. \tag{1.60}$$

It is important to note that $\boldsymbol{\omega}_{21}$ are the coordinates of $\vec{\boldsymbol{\omega}}_{21}$ in frame $\mathcal{F}_2$, and $\mathbf{C}_{21}$ is the rotation matrix transforming coordinates from frame $\mathcal{F}_1$ to frame $\mathcal{F}_2$.

Making use of the identity $\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times = \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T - \boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\mathbf{1}$ (see (1.14)), and the unit constraint (1.32), the rotation matrix $\mathbf{C}_{21}$ with quaternion parameterization $(\boldsymbol{\epsilon}, \eta)$, may be written as

$$\mathbf{C}_{21} = \mathbf{1} + 2\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times - 2\eta\boldsymbol{\epsilon}^\times. \tag{1.61}$$

Differentiating, we have

$$\dot{\mathbf{C}}_{21} = 2\left[\left(\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times + \boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times\right) - \dot{\eta}\boldsymbol{\epsilon}^\times - \eta\dot{\boldsymbol{\epsilon}}^\times\right]. \tag{1.62}$$

By post-multiplying (1.62) with (1.61) and expanding, we find that

$$\begin{aligned}\tfrac{1}{2}\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T = &\left(\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times + \boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times\right) - \dot{\eta}\boldsymbol{\epsilon}^\times - \eta\dot{\boldsymbol{\epsilon}}^\times \\ &+ 2\left(\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times + \boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times\right)\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times - 2\dot{\eta}\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times \\ &+ 2\eta\boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times - 2\eta\dot{\eta}\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times - 2\eta^2\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times.\end{aligned} \tag{1.63}$$

To reduce this expression, we need to make use of some column matrix identities. We have previously derived

$$\mathbf{a}^\times\mathbf{b}^\times = \mathbf{b}\mathbf{a}^T - \left(\mathbf{a}^T\mathbf{b}\right)\mathbf{1}.$$

Making use of this identity, we now expand the individual terms in (1.63). We have

$$\begin{aligned}\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times &= \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T - \left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right)\mathbf{1}, \\ \boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times &= -\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right)\boldsymbol{\epsilon}^\times, \\ \dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times &= \boldsymbol{\epsilon}\dot{\boldsymbol{\epsilon}}^T - \left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\mathbf{1}, \\ \boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times\boldsymbol{\epsilon}^\times &= -\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\boldsymbol{\epsilon}^\times, \\ \boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}^\times &= \dot{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T - \left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\mathbf{1},\end{aligned}$$

Substituting these into (1.63) and collecting like terms, we obtain

$$\begin{aligned}\tfrac{1}{2}\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T = &\dot{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T + \left[1 - 2\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) - 2\eta^2\right]\boldsymbol{\epsilon}\dot{\boldsymbol{\epsilon}}^T \\ &+ \left[-2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) + 4\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) + 2\eta\dot{\eta}\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) + 2\eta^2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\right]\mathbf{1} \\ &+ \left[-\dot{\eta} + 2\dot{\eta}\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) - 2\eta\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\right]\boldsymbol{\epsilon}^\times - \eta\dot{\boldsymbol{\epsilon}}^\times \\ &+ \left[-2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) - 2\eta\dot{\eta}\right]\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T.\end{aligned} \tag{1.64}$$

We can now further reduce each of the coefficients in (1.64), by making use of the unit magnitude constraint on the quaternion

$$\boldsymbol{\epsilon}^T\boldsymbol{\epsilon} + \eta^2 = 1, \tag{1.65}$$

and its derivative

$$\boldsymbol{\epsilon}^T \dot{\boldsymbol{\epsilon}} + \eta\dot{\eta} = 0. \tag{1.66}$$

Making use of (1.65), we find that

$$1 - 2\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) - 2\eta^2 = -1.$$

Making use of (1.65) and (1.66), we find that

$$\begin{aligned}
&-2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) + 4\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) + 2\eta\dot{\eta}\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) + 2\eta^2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) \\
&= -2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) + 2\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon} + \eta^2\right)\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) + 2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}} + \eta\dot{\eta}\right)\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) = 0.
\end{aligned}$$

$$\begin{aligned}
-\dot{\eta} + 2\dot{\eta}\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) - 2\eta\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) &= \dot{\eta}\left(2\left(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\right) - 1\right) - 2\eta\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) \\
&= \dot{\eta}\left(1 - 2\eta^2\right) - 2\eta\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) \\
&= \dot{\eta} - 2\left(\eta\dot{\eta} + \boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right)\eta \\
&= \dot{\eta}
\end{aligned}$$

$$-2\left(\boldsymbol{\epsilon}^T\dot{\boldsymbol{\epsilon}}\right) - 2\eta\dot{\eta} = 0.$$

Therefore, (1.64) becomes

$$\frac{1}{2}\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T = \left(\dot{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T - \boldsymbol{\epsilon}\dot{\boldsymbol{\epsilon}}^T\right) - \eta\dot{\boldsymbol{\epsilon}}^\times + \dot{\eta}\dot{\boldsymbol{\epsilon}}^\times.$$

We need one more identity to reduce the bracketed term. Making use of (1.42), we find that

$$\dot{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}^T - \boldsymbol{\epsilon}\dot{\boldsymbol{\epsilon}}^T = \left(\boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}\right)^\times,$$

and we finally have

$$\frac{1}{2}\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T = \left(\boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}\right)^\times - \eta\dot{\boldsymbol{\epsilon}}^\times + \dot{\eta}\boldsymbol{\epsilon}^\times.$$

Therefore, from (1.60), we have

$$\boldsymbol{\omega}_{21}^\times = -2\left[\left(\boldsymbol{\epsilon}^\times\dot{\boldsymbol{\epsilon}}\right)^\times - \eta\dot{\boldsymbol{\epsilon}}^\times + \dot{\eta}\boldsymbol{\epsilon}^\times\right]$$

Finally, since both sides of the above equation are cross-product operators, we can extract the equation for the angular velocity as

$$\boldsymbol{\omega}_{21} = 2\left(\eta\mathbf{1} - \boldsymbol{\epsilon}^\times\right)\dot{\boldsymbol{\epsilon}} - 2\boldsymbol{\epsilon}\dot{\eta}. \tag{1.67}$$

We can now invert (1.67) to find the quaternion rates in terms of the angular velocity. To do this, we write (1.67) in matrix form, and append the derivative constraint (1.66).

$$\begin{bmatrix} \boldsymbol{\omega}_{21} \\ 0 \end{bmatrix} = 2 \begin{bmatrix} \left(\eta \mathbf{1} - \boldsymbol{\epsilon}^{\times}\right) & -\boldsymbol{\epsilon} \\ \boldsymbol{\epsilon}^T & \eta \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\epsilon}} \\ \dot{\eta} \end{bmatrix}.$$

It is straightforward to show that the above matrix is orthonormal (just multiply it by its transpose and use the same identities as before). Therefore, we may invert the above relationship to obtain

$$\begin{bmatrix} \dot{\boldsymbol{\epsilon}} \\ \dot{\eta} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \left(\eta \mathbf{1} + \boldsymbol{\epsilon}^{\times}\right) & \boldsymbol{\epsilon} \\ -\boldsymbol{\epsilon}^T & \eta \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega}_{21} \\ 0 \end{bmatrix}.$$

From this, we can extract the equations for the quaternion rates in terms of the angular velocity

$$\begin{aligned} \dot{\boldsymbol{\epsilon}} &= \frac{1}{2}\left(\eta \mathbf{1} + \boldsymbol{\epsilon}^{\times}\right) \boldsymbol{\omega}_{21}, \\ \dot{\eta} &= -\frac{1}{2} \boldsymbol{\epsilon}^T \boldsymbol{\omega}_{21}. \end{aligned} \tag{1.68}$$

1.5 Velocity and Acceleration

We shall now obtain expressions for the velocity and acceleration as seen in different reference frames. As before, we consider the frames $\mathcal{F}_1$ and $\mathcal{F}_2$.

Let us denote the *velocity* as seen in frame $\mathcal{F}_1$ by

$$\vec{\mathbf{v}} = \dot{\vec{\mathbf{r}}} = \mathring{\vec{\mathbf{r}}} + \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{r}} \tag{1.69}$$

The *acceleration* as seen in frame $\mathcal{F}_1$ can be obtained by applying the rule in (1.52) to $\vec{\mathbf{v}}$:

$$\begin{aligned} \ddot{\vec{\mathbf{r}}} = \dot{\vec{\mathbf{v}}} &= \mathring{\vec{\mathbf{v}}} + \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{v}} \\ &= \left(\overset{\circ\circ}{\vec{\mathbf{r}}} + \vec{\boldsymbol{\omega}}_{21} \times \mathring{\vec{\mathbf{r}}} + \mathring{\vec{\boldsymbol{\omega}}}_{21} \times \vec{\mathbf{r}} \right) + \left(\vec{\boldsymbol{\omega}}_{21} \times \mathring{\vec{\mathbf{r}}} + \vec{\boldsymbol{\omega}}_{21} \times \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{r}} \right) \\ &= \overset{\circ\circ}{\vec{\mathbf{r}}} + 2\vec{\boldsymbol{\omega}}_{21} \times \mathring{\vec{\mathbf{r}}} + \mathring{\vec{\boldsymbol{\omega}}}_{21} \times \vec{\mathbf{r}} + \vec{\boldsymbol{\omega}}_{21} \times \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{r}}. \end{aligned} \tag{1.70}$$

The different terms in the above expression for the acceleration have special names:

$\overset{\circ\circ}{\vec{\mathbf{r}}}$	acceleration as seen in $\mathcal{F}_2$,
$2\vec{\boldsymbol{\omega}}_{21} \times \mathring{\vec{\mathbf{r}}}$	coriolis acceleration,
$\mathring{\vec{\boldsymbol{\omega}}}_{21} \times \vec{\mathbf{r}}$	angular acceleration,
$\vec{\boldsymbol{\omega}}_{21} \times \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{r}}$	centripetal acceleration.

The coordinates for acceleration in (1.70) given in the two reference frames can be obtained by using the expressions

$$\ddot{\vec{\mathbf{r}}} = \vec{\mathcal{F}}_1^T \ddot{\mathbf{r}}_1, \quad \overset{\circ\circ}{\vec{\mathbf{r}}} = \vec{\mathcal{F}}_2^T \ddot{\mathbf{r}}_2, \quad \vec{\boldsymbol{\omega}}_{21} = \vec{\mathcal{F}}_2^T \boldsymbol{\omega}_{21}.$$

The result is

$$\ddot{\mathbf{r}}_1 = \mathbf{C}_{12}\left[\ddot{\mathbf{r}}_2 + 2\boldsymbol{\omega}_{21}^{\times}\dot{\mathbf{r}}_2 + \dot{\boldsymbol{\omega}}_{21}^{\times}\mathbf{r}_2 + \boldsymbol{\omega}_{21}^{\times}\boldsymbol{\omega}_{21}^{\times}\mathbf{r}_2\right].$$

1.6 More Rigorous Definition of Angular Velocity

Consider two reference frames $\mathcal{F}_2$ and $\mathcal{F}_1$. Let the unit basis vectors of frame $\mathcal{F}_2$ be $\vec{\mathbf{x}}_2$, $\vec{\mathbf{y}}_2$ and $\vec{\mathbf{z}}_2$. Recall that the basis vectors must satisfy the normality and orthogonality constraints

$$\vec{\mathbf{x}}_2 \cdot \vec{\mathbf{x}}_2 = \vec{\mathbf{y}}_2 \cdot \vec{\mathbf{y}}_2 = \vec{\mathbf{z}}_2 \cdot \vec{\mathbf{z}}_2 = 1, \tag{1.71}$$

and

$$\vec{\mathbf{x}}_2 \cdot \vec{\mathbf{y}}_2 = \vec{\mathbf{x}}_2 \cdot \vec{\mathbf{z}}_2 = \vec{\mathbf{y}}_2 \cdot \vec{\mathbf{z}}_2 = 0. \tag{1.72}$$

respectively.

We are now going to examine the time-derivative of frame $\mathcal{F}_2$ as seen in frame $\mathcal{F}_1$. Let $(\dot{\vec{\mathbf{a}}})$ denote the time-derivative of the vector $\vec{\mathbf{a}}$ as seen in $\mathcal{F}_1$.

Let us now take the derivatives of the normality constraints (1.71). We see that

$$\vec{\mathbf{x}}_2 \cdot \dot{\vec{\mathbf{x}}}_2 = 0, \quad \vec{\mathbf{y}}_2 \cdot \dot{\vec{\mathbf{y}}}_2 = 0, \quad \vec{\mathbf{z}}_2 \cdot \dot{\vec{\mathbf{z}}}_2 = 0.$$

That is,

$$\vec{\mathbf{x}}_2 \perp \dot{\vec{\mathbf{x}}}_2, \quad \vec{\mathbf{y}}_2 \perp \dot{\vec{\mathbf{y}}}_2, \quad \vec{\mathbf{z}}_2 \perp \dot{\vec{\mathbf{z}}}_2.$$

Therefore, we can find vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ such that

$$\dot{\vec{\mathbf{x}}}_2 = \vec{\mathbf{a}} \times \vec{\mathbf{x}}_2, \quad \dot{\vec{\mathbf{y}}}_2 = \vec{\mathbf{b}} \times \vec{\mathbf{y}}_2, \quad \dot{\vec{\mathbf{z}}}_2 = \vec{\mathbf{c}} \times \vec{\mathbf{z}}_2. \tag{1.73}$$

The reason we can do this will be clear shortly. Let $\vec{\mathbf{a}}_{\perp x}$, $\vec{\mathbf{b}}_{\perp y}$ and $\vec{\mathbf{c}}_{\perp z}$ be the components of $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ that are perpendicular to $\vec{\mathbf{x}}_2$, $\vec{\mathbf{y}}_2$ and $\vec{\mathbf{z}}_2$ respectively. As we have seen in (1.6), we can rewrite the above derivatives as

$$\dot{\vec{\mathbf{x}}}_2 = \vec{\mathbf{a}}_{\perp x} \times \vec{\mathbf{x}}_2, \quad \dot{\vec{\mathbf{y}}}_2 = \vec{\mathbf{b}}_{\perp y} \times \vec{\mathbf{y}}_2, \quad \dot{\vec{\mathbf{z}}}_2 = \vec{\mathbf{c}}_{\perp z} \times \vec{\mathbf{z}}_2.$$

Now, let us express each of the vectors in $\mathcal{F}_2$ coordinates. We have

$$\vec{\mathbf{x}}_2 = \vec{\mathcal{F}}_2^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \vec{\mathbf{y}}_2 = \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \vec{\mathbf{z}}_2 = \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

and

$$\vec{\mathbf{a}}_{\perp x} = \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ a_y \\ a_z \end{bmatrix}, \ \vec{\mathbf{b}}_{\perp y} = \vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ 0 \\ b_z \end{bmatrix}, \ \vec{\mathbf{c}}_{\perp x} = \vec{\mathcal{F}}_2^T \begin{bmatrix} c_x \\ c_y \\ 0 \end{bmatrix}.$$

Let us now see why we can always find vectors $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ such that our derivatives are given by (1.73). Consider $\dot{\vec{\mathbf{x}}}_2$. Since $\vec{\mathbf{x}}_2 \perp \dot{\vec{\mathbf{x}}}_2$, when $\dot{\vec{\mathbf{x}}}_2$ is expressed in $\mathcal{F}_2$ coordinates, it must have the form

$$\dot{\vec{\mathbf{x}}}_2 = \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ p \\ q \end{bmatrix},$$

for some p and q. Let us now look at the derivative expression in $\mathcal{F}_2$. We have

$$\dot{\vec{\mathbf{x}}}_2 = \vec{\mathbf{a}}_{\perp x} \times \vec{\mathbf{x}}_2 = -\vec{\mathbf{x}}_2 \times \vec{\mathbf{a}}_{\perp x},$$

which leads to

$$\begin{bmatrix} 0 \\ p \\ q \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} 0 \\ a_z \\ +a_y \end{bmatrix}.$$

which is clearly always solvable for a_y and a_z, and hence $\vec{\mathbf{a}}$. We can similarly always solve for $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$.

Now, let us examine the orthogonality constraints (1.72). Differentiating these, we obtain

$$\begin{aligned} \dot{\vec{\mathbf{x}}}_2 \cdot \vec{\mathbf{y}}_2 + \vec{\mathbf{x}}_2 \cdot \dot{\vec{\mathbf{y}}}_2 = 0, \\ \dot{\vec{\mathbf{x}}}_2 \cdot \vec{\mathbf{z}}_2 + \vec{\mathbf{x}}_2 \cdot \dot{\vec{\mathbf{z}}}_2 = 0, \\ \dot{\vec{\mathbf{y}}}_2 \cdot \vec{\mathbf{z}}_2 + \vec{\mathbf{y}}_2 \cdot \dot{\vec{\mathbf{z}}}_2 = 0. \end{aligned}$$

Substituting the derivative expressions, these become

$$\begin{aligned} \left(\vec{\mathbf{a}}_{\perp x} \times \vec{\mathbf{x}}_2\right) \cdot \vec{\mathbf{y}}_2 + \vec{\mathbf{x}}_2 \cdot \left(\vec{\mathbf{b}}_{\perp y} \times \vec{\mathbf{y}}_2\right) = 0, \\ \left(\vec{\mathbf{a}}_{\perp x} \times \vec{\mathbf{x}}_2\right) \cdot \vec{\mathbf{z}}_2 + \vec{\mathbf{x}}_2 \cdot \left(\vec{\mathbf{c}}_{\perp z} \times \vec{\mathbf{z}}_2\right) = 0, \\ \left(\vec{\mathbf{b}}_{\perp y} \times \vec{\mathbf{y}}_2\right) \cdot \vec{\mathbf{z}}_2 + \vec{\mathbf{y}}_2 \cdot \left(\vec{\mathbf{c}}_{\perp z} \times \vec{\mathbf{z}}_2\right) = 0. \end{aligned}$$

Making use of the scalar triple product identity (see the end of Section 2.1), and the fact that for a right-handed coordinate system we have $\vec{\mathbf{x}}_1 \times \vec{\mathbf{y}}_1 = \vec{\mathbf{z}}_1$, $\vec{\mathbf{y}}_1 \times \vec{\mathbf{z}}_1 = \vec{\mathbf{x}}_1$ and $\vec{\mathbf{z}}_1 \times \vec{\mathbf{x}}_1 = \vec{\mathbf{y}}_1$, we obtain the constraints

$$\begin{aligned}\left(\vec{\mathbf{b}}_{\perp y} - \vec{\mathbf{a}}_{\perp x}\right) \cdot \vec{\mathbf{z}}_2 = 0,\\ \left(\vec{\mathbf{c}}_{\perp z} - \vec{\mathbf{a}}_{\perp x}\right) \cdot \vec{\mathbf{y}}_2 = 0,\\ \left(\vec{\mathbf{c}}_{\perp z} - \vec{\mathbf{b}}_{\perp y}\right) \cdot \vec{\mathbf{x}}_2 = 0.\end{aligned}$$

These constraints lead to

$$b_z = a_z, \quad c_x = b_x, \quad c_y = a_y,$$

so that we must have

$$\vec{\mathbf{a}}_{\perp x} = \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ a_y \\ a_z \end{bmatrix}, \quad \vec{\mathbf{b}}_{\perp y} = \vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ 0 \\ a_z \end{bmatrix}, \quad \vec{\mathbf{c}}_{\perp z} = \vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ a_y \\ 0 \end{bmatrix}.$$

Now, note that the vector $\vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ 0 \\ 0 \end{bmatrix}$ is parallel to $\vec{\mathbf{x}}_2$, the vector $\vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ a_y \\ 0 \end{bmatrix}$ is parallel to $\vec{\mathbf{y}}_2$ and the vector $\vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ 0 \\ a_z \end{bmatrix}$ is parallel to $\vec{\mathbf{z}}_2$. Therefore, see that we can write

$$\vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{x}}_2 = \vec{\mathbf{a}}_{\perp x} \times \vec{\mathbf{x}}_2, \quad \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{y}}_2 = \vec{\mathbf{b}}_{\perp y} \times \vec{\mathbf{y}}_2, \quad \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{z}}_2 = \vec{\mathbf{c}}_{\perp z} \times \vec{\mathbf{z}}_2,$$

where

$$\vec{\boldsymbol{\omega}}_{21} = \vec{\mathbf{a}}_{\perp x} + \vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ 0 \\ 0 \end{bmatrix} = \vec{\mathbf{b}}_{\perp y} + \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ a_y \\ 0 \end{bmatrix} = \vec{\mathbf{c}}_{\perp z} + \vec{\mathcal{F}}_2^T \begin{bmatrix} 0 \\ 0 \\ a_z \end{bmatrix} = \vec{\mathcal{F}}_2^T \begin{bmatrix} b_x \\ a_y \\ a_z \end{bmatrix}.$$

We call the vector $\vec{\boldsymbol{\omega}}_{21}$ the angular velocity of frame $\mathcal{F}_2$ relative to frame $\mathcal{F}_1$, and we have the relationships

$$\dot{\vec{\mathbf{x}}}_2 = \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{x}}_2, \quad \dot{\vec{\mathbf{y}}}_2 = \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{y}}_2, \quad \dot{\vec{\mathbf{z}}}_2 = \vec{\boldsymbol{\omega}}_{21} \times \vec{\mathbf{z}}_2.$$

These are the same as we have formally derived in Section 1.4.1.

Notes

In this chapter, we have developed the mathematics necessary to describe the kinematics of spacecraft motion. We have purposely chosen to develop the kinematics using the vectrix formalism. The use of vectrices provides a clear distinction between a physical vector and its coordinate representation in a particular reference frame. This becomes important in spacecraft dynamics problems where many different reference frames may be considered simultaneously. This then results in several different coordinate representations for the same physical vector. As will become evident in later chapters in this book, the use of vectrices allows the spacecraft equations of motion to be derived purely in terms of physical vectors, without consideration of its coordinates in any particular reference frame. Finally, once the equations of motion have been obtained in physical vector form, their coordinate representations in any particular reference frame may be obtained directly. The term vectrix was first coined by Peter C. Hughes, and the reader can find further treatment of vectrices in Hughes (2004). Similar formalisms may also be found in Likins (1973) and Wittenburg (1977). In Section 1.3, we have explained how the spacecraft attitude is described by a rotation matrix, which may be equivalently represented by a principal axis and angle of rotation, a quaternion or a set of Euler Angles. There are several other parameterizations of a rotation matrix. A detailed treatment is contained in Shuster (1993).

References

Hughes PC 2004 *Spacecraft Attitude Dynamics*. Dover Publications, Mineola, NY.
Likins PW 1973 *Elements of Engineering Mechanics*. McGraw-Hill, New York, NY.
Shuster MD 1993 A Survey of Attitude Representations. *Journal of the Astronautical Sciences* **41**(4), 439–517.
Wittenburg J 1977 *Dynamics of Systems of Rigid Bodies*. Teubner, Stuttgart.