6

Primer on Three-Dimensional Geometry

This chapter will introduce three-dimensional geometry and specifically the concept of a *rotation* and some of its representations. It pays particular attention to the establishment of *reference frames*. Sastry (1999) is a comprehensive reference on control for robotics that includes a background on three-dimensional geometry. Hughes (1986) also provides a good first-principles background.

6.1 Vectors and Reference Frames

Vehicles (e.g., robots, satellites, aircraft) are typically free to translate and rotate. Mathematically, they have six degrees of freedom: three in translation and three in rotation. This six-degree-of-freedom geometric configuration is known as the *pose* (position and orientation) of the vehicle. Some vehicles may have multiple bodies connected together; in this case each body has its own pose. We will consider only the single-body case here.

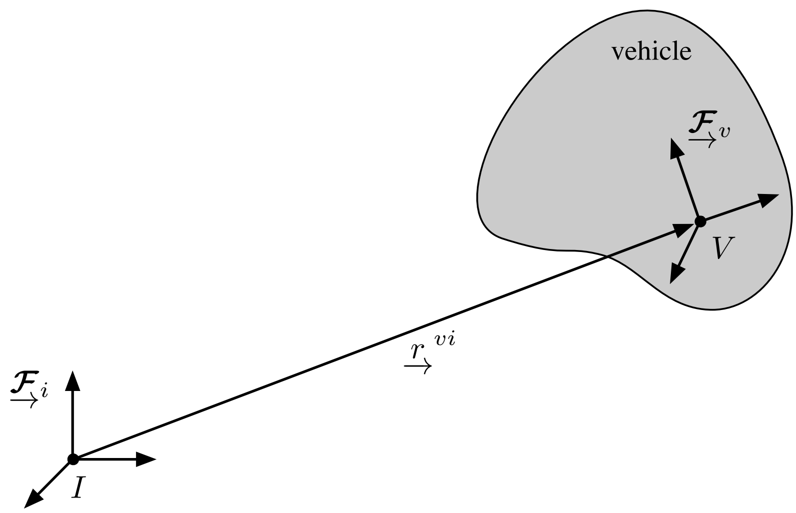
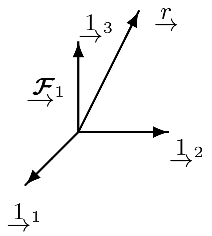


Figure 6.1 Vehicle and typical reference frames.

6.1.1 Reference Frames

The position of a point on a vehicle can be described with a vector, $\underrightarrow{r}^{vi}$, consisting of three components. Rotational motion is described by expressing the orientation of a reference frame on the vehicle, $\underrightarrow{\mathcal{F}}_v$, with respect to another frame, $\underrightarrow{\mathcal{F}}_i$. Figure 6.1 shows the typical setup for a single-body vehicle.

We will take a *vector* to be a quantity $\underrightarrow{r}$ having length and direction. This vector can be expressed in a reference frame as



$$\begin{aligned}\underrightarrow{r} &= r_1 \underrightarrow{1}_1 + r_2 \underrightarrow{1}_2 + r_3 \underrightarrow{1}_3 \\ &= [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} \underrightarrow{1}_1 \\ \underrightarrow{1}_2 \\ \underrightarrow{1}_3 \end{bmatrix} \\ &= \mathbf{r}_1^T \underrightarrow{\mathcal{F}}_1. \end{aligned} \tag{6.1}$$

The quantity

$$\underrightarrow{\mathcal{F}}_1 = \begin{bmatrix} \underrightarrow{1}_1 \\ \underrightarrow{1}_2 \\ \underrightarrow{1}_3 \end{bmatrix}$$

is a column containing the basis vectors forming the reference frame $\underrightarrow{\mathcal{F}}_1$; we will always use basis vectors that are unit length, orthogonal, and arranged in a dextral (right-handed) fashion. We shall refer to $\underrightarrow{\mathcal{F}}_1$ as a *vectrix* (Hughes, 1986). The quantity

$$\mathbf{r}_1 = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

is a column matrix containing the *components* or *coordinates* of $\underrightarrow{r}$ in reference frame $\underrightarrow{\mathcal{F}}_1$.

The vector can also be written as

$$\begin{aligned}\underrightarrow{r} &= \left[\underrightarrow{1}_1 \;\; \underrightarrow{1}_2 \;\; \underrightarrow{1}_3\right] \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \\ &= \underrightarrow{\mathcal{F}}_1^T \mathbf{r}_1. \end{aligned}$$

6.1.2 Dot Product

Consider two vectors, $\underrightarrow{r}$ and $\underrightarrow{s}$, expressed in the same reference frame $\underrightarrow{\mathcal{F}}_1$:

$$\underrightarrow{r} = [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} \underrightarrow{1}_1 \\ \underrightarrow{1}_2 \\ \underrightarrow{1}_3 \end{bmatrix}, \quad \underrightarrow{s} = \left[\underrightarrow{1}_1 \;\; \underrightarrow{1}_2 \;\; \underrightarrow{1}_3\right] \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}.$$

The *dot product* (a.k.a., inner product) is given by

$$
\begin{aligned}
\underset{\rightarrow}{r} \cdot \underset{\rightarrow}{s} &= [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} \underset{\rightarrow}{1}_1 \\ \underset{\rightarrow}{1}_2 \\ \underset{\rightarrow}{1}_3 \end{bmatrix} \cdot \begin{bmatrix} \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \\
&= [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} \underset{\rightarrow}{1}_1 \cdot \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_1 \cdot \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_1 \cdot \underset{\rightarrow}{1}_3 \\ \underset{\rightarrow}{1}_2 \cdot \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_2 \cdot \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_2 \cdot \underset{\rightarrow}{1}_3 \\ \underset{\rightarrow}{1}_3 \cdot \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_3 \cdot \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_3 \cdot \underset{\rightarrow}{1}_3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}.
\end{aligned}
$$

But

$$\underset{\rightarrow}{1}_1 \cdot \underset{\rightarrow}{1}_1 = \underset{\rightarrow}{1}_2 \cdot \underset{\rightarrow}{1}_2 = \underset{\rightarrow}{1}_3 \cdot \underset{\rightarrow}{1}_3 = 1,$$

and

$$\underset{\rightarrow}{1}_1 \cdot \underset{\rightarrow}{1}_2 = \underset{\rightarrow}{1}_2 \cdot \underset{\rightarrow}{1}_3 = \underset{\rightarrow}{1}_3 \cdot \underset{\rightarrow}{1}_1 = 0.$$

Therefore,

$$\underset{\rightarrow}{r} \cdot \underset{\rightarrow}{s} = \mathbf{r}_1{}^T \mathbf{1} \mathbf{s}_1 = \mathbf{r}_1{}^T \mathbf{s}_1 = r_1 s_1 + r_2 s_2 + r_3 s_3.$$

The notation **1** will be used to designate the *identity matrix*. Its dimension can be inferred from context.

6.1.3 Cross Product

The *cross product* of two vectors expressed in the same reference frame is given by

$$
\begin{aligned}
\underset{\rightarrow}{r} \times \underset{\rightarrow}{s} &= [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} \underset{\rightarrow}{1}_1 \times \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_1 \times \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_1 \times \underset{\rightarrow}{1}_3 \\ \underset{\rightarrow}{1}_2 \times \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_2 \times \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_2 \times \underset{\rightarrow}{1}_3 \\ \underset{\rightarrow}{1}_3 \times \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_3 \times \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_3 \times \underset{\rightarrow}{1}_3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \\
&= [r_1 \;\; r_2 \;\; r_3] \begin{bmatrix} 0 & \underset{\rightarrow}{1}_3 & -\underset{\rightarrow}{1}_2 \\ -\underset{\rightarrow}{1}_3 & 0 & \underset{\rightarrow}{1}_1 \\ \underset{\rightarrow}{1}_2 & -\underset{\rightarrow}{1}_1 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \\
&= \begin{bmatrix} \underset{\rightarrow}{1}_1 & \underset{\rightarrow}{1}_2 & \underset{\rightarrow}{1}_3 \end{bmatrix} \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \\
&= \underset{\rightarrow}{\mathcal{F}}_1^T \mathbf{r}_1{}^\times \mathbf{s}_1,
\end{aligned}
$$

where the fact that the basis vectors are orthogonal and arranged in a dextral fashion has been exploited. Hence, if $\underset{\rightarrow}{r}$ and $\underset{\rightarrow}{s}$ are expressed in the same reference frame, the 3×3 matrix

$$\mathbf{r}_1{}^\times = \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix}, \tag{6.2}$$

can be used to construct the components of the cross product. This matrix is *skew-symmetric*[1]; that is,

$$(\mathbf{r}_1{}^{\times})^T = -\mathbf{r}_1{}^{\times}.$$

It is easy to verify that

$$\mathbf{r}_1{}^{\times}\mathbf{r}_1 = \mathbf{0},$$

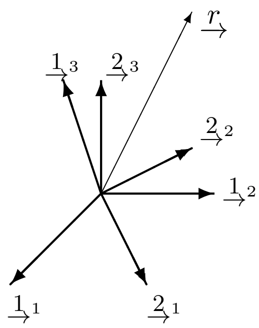
where $\mathbf{0}$ is a column matrix of zeros and

$$\mathbf{r}_1{}^{\times}\mathbf{s}_1 = -\mathbf{s}_1{}^{\times}\mathbf{r}_1.$$

6.2 Rotations

Critical to our ability to estimate how objects are moving in the world is the ability to parameterize the orientation, or rotation, of those objects. We begin by introducing rotation matrices and then provide some alternative representations.

6.2.1 Rotation Matrices



Let us consider two frames $\underrightarrow{\mathcal{F}}_1$ and $\underrightarrow{\mathcal{F}}_2$ with a common origin, and let us express $\underrightarrow{r}$ in each frame:

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

We seek to discover a relationship between the components in $\underrightarrow{\mathcal{F}}_1$, $\mathbf{r}_1$, and those in $\underrightarrow{\mathcal{F}}_2$, $\mathbf{r}_2$. We proceed as follows:

$$\begin{aligned}
\underrightarrow{\mathcal{F}}_2^T\mathbf{r}_2 &= \underrightarrow{\mathcal{F}}_1^T\mathbf{r}_1, \\
\underrightarrow{\mathcal{F}}_2 \cdot \underrightarrow{\mathcal{F}}_2^T\mathbf{r}_2 &= \underrightarrow{\mathcal{F}}_2 \cdot \underrightarrow{\mathcal{F}}_1^T\mathbf{r}_1, \\
\mathbf{r}_2 &= \mathbf{C}_{21}\mathbf{r}_1.
\end{aligned}$$

We have defined

$$\begin{aligned}
\mathbf{C}_{21} &= \underrightarrow{\mathcal{F}}_2 \cdot \underrightarrow{\mathcal{F}}_1^T \\
&= \begin{bmatrix} \underrightarrow{2}_1 \\ \underrightarrow{2}_2 \\ \underrightarrow{2}_3 \end{bmatrix} \cdot \begin{bmatrix} \underrightarrow{1}_1 & \underrightarrow{1}_2 & \underrightarrow{1}_3 \end{bmatrix} \\
&= \begin{bmatrix} \underrightarrow{2}_1 \cdot \underrightarrow{1}_1 & \underrightarrow{2}_1 \cdot \underrightarrow{1}_2 & \underrightarrow{2}_1 \cdot \underrightarrow{1}_3 \\ \underrightarrow{2}_2 \cdot \underrightarrow{1}_1 & \underrightarrow{2}_2 \cdot \underrightarrow{1}_2 & \underrightarrow{2}_2 \cdot \underrightarrow{1}_3 \\ \underrightarrow{2}_3 \cdot \underrightarrow{1}_1 & \underrightarrow{2}_3 \cdot \underrightarrow{1}_2 & \underrightarrow{2}_3 \cdot \underrightarrow{1}_3 \end{bmatrix}.
\end{aligned}$$

[1] There are many equivalent notations in the literature for this skew-symmetric definition: $\mathbf{r}_1^{\times} = \hat{\mathbf{r}}_1 = \mathbf{r}_1^{\wedge} = -[[\mathbf{r}_1]] = [\mathbf{r}_1]_{\times}$. For now, we use the first one, since it makes an obvious connection to the cross product; later we will also use $(\cdot)^{\wedge}$, as this is in common use in robotics.

The matrix $\mathbf{C}_{21}$ is called a *rotation matrix*. It is sometimes referred to as a 'direction cosine matrix' since the dot product of two unit vectors is just the cosine of the angle between them.

The unit vectors in $\underrightarrow{\mathcal{F}}_2$ can be related to those in $\underrightarrow{\mathcal{F}}_1$:

$$\underrightarrow{\mathcal{F}}_1^T = \underrightarrow{\mathcal{F}}_2^T \mathbf{C}_{21}. \tag{6.3}$$

Rotation matrices possess some special properties:

$$\mathbf{r}_1 = \mathbf{C}_{21}^{-1}\mathbf{r}_2 = \mathbf{C}_{12}\mathbf{r}_2.$$

But, $\mathbf{C}_{21}^T = \mathbf{C}_{12}$. Hence,

$$\mathbf{C}_{12} = \mathbf{C}_{21}^{-1} = \mathbf{C}_{21}^T. \tag{6.4}$$

We say that $\mathbf{C}_{21}$ is an *orthonormal* matrix because its inverse is equal to its transpose.

Consider three reference frames $\underrightarrow{\mathcal{F}}_1$, $\underrightarrow{\mathcal{F}}_2$, and $\underrightarrow{\mathcal{F}}_3$. The components of a vector $\underrightarrow{r}$ in these three frames are $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$. Now,

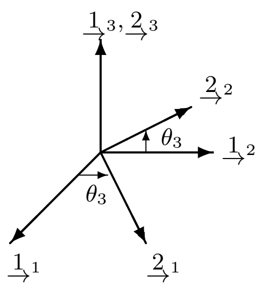
$$\mathbf{r}_3 = \mathbf{C}_{32}\mathbf{r}_2 = \mathbf{C}_{32}\mathbf{C}_{21}\mathbf{r}_1.$$

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

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

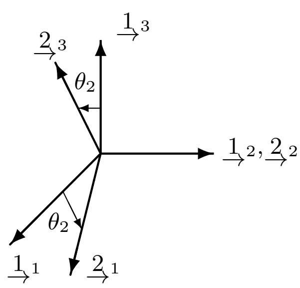
6.2.2 Principal Rotations

Before considering more general rotations, it is useful to consider rotations about one basis vector. The situation where $\underrightarrow{\mathcal{F}}_2$ has been rotated from $\underrightarrow{\mathcal{F}}_1$ through a rotation about the 3-axis is shown in the figure. The rotation matrix in this case is

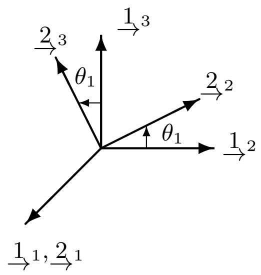


$$\mathbf{C}_3 = \begin{bmatrix} \cos\theta_3 & \sin\theta_3 & 0 \\ -\sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6.5}$$

For a rotation about the 2-axis, the rotation matrix is



$$\mathbf{C}_2 = \begin{bmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix}. \tag{6.6}$$



For a rotation about the 1-axis, the rotation matrix is

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix}. \tag{6.7}$$

6.2.3 Alternate Rotation Representations

We have seen one way of discussing the orientation of one reference frame with respect to another: the *rotation matrix*. The rotation matrix describes orientation both globally and uniquely. This requires nine parameters (they are not independent). There are a number of other alternatives.

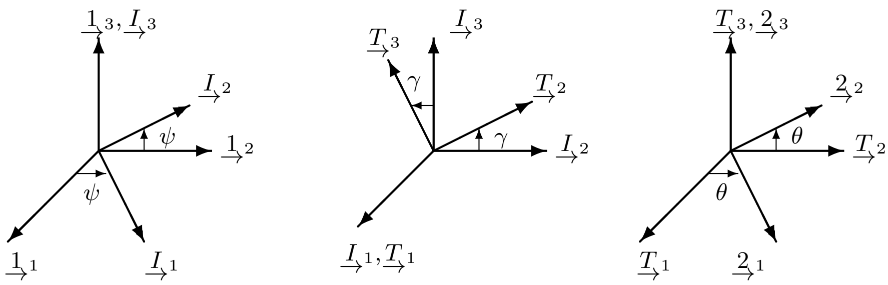
The key thing to realize about the different representations of rotations, is that there are always only three underlying degrees of freedom. The representations that have more than three parameters must have associated constraints to limit the number of degrees of freedom to three. The representations that have exactly three parameters have associated singularities. There is no perfect representation that is minimal (i.e., having only three parameters) and that is also free of singularities (Stuelpnagel, 1964).

Leonhard Euler (1707-1783) is considered to be the preeminent mathematician of the eighteenth century and one of the greatest mathematicians to have ever lived. He made important discoveries in fields as diverse as infinitesimal calculus and graph theory. He also introduced much of the modern mathematical terminology and notation, particularly for mathematical analysis, such as the notion of a mathematical function. He is also renowned for his work in mechanics, fluid dynamics, optics, astronomy, and music theory.

Euler Angles

The orientation of one reference frame with respect to another can also be specified by a sequence of three principal rotations. One possible sequence is as follows:

(i) A rotation ψ about the original 3-axis
(ii) A rotation γ about the intermediate 1-axis
(iii) A rotation θ about the transformed 3-axis



This is called a 3-1-3 sequence and is the one originally used by Euler.

In the classical mechanics literature, the angles are referred to by the following names:

$$\begin{aligned} \theta &: \text{spin angle} \\ \gamma &: \text{nutation angle} \\ \psi &: \text{precession angle} \end{aligned}$$

The rotation matrix from frame 1 to frame 2 is given by

$$\begin{aligned} \mathbf{C}_{21}(\theta, \gamma, \psi) &= \mathbf{C}_{2T}\mathbf{C}_{TI}\mathbf{C}_{I1} \\ &= \mathbf{C}_3(\theta)\mathbf{C}_1(\gamma)\mathbf{C}_3(\psi) \\ &= \begin{bmatrix} c_\theta c_\psi - s_\theta c_\gamma s_\psi & s_\psi c_\theta + c_\gamma s_\theta c_\psi & s_\gamma s_\theta \\ -c_\psi s_\theta - c_\theta c_\gamma s_\psi & -s_\psi s_\theta + c_\theta c_\gamma c_\psi & s_\gamma c_\theta \\ s_\psi s_\gamma & -s_\gamma c_\psi & c_\gamma \end{bmatrix}. \end{aligned} \quad (6.8)$$

We have made the abbreviations $s = \sin$, $c = \cos$.

Another possible sequence that can be used is as follows:

(i) A rotation θ_1 about the original 1-axis ('roll' rotation)
(ii) A rotation θ_2 about the intermediate 2-axis ('pitch' rotation)
(iii) A rotation θ_3 about the transformed 3-axis ('yaw' rotation)

This sequence, which is very common in aerospace applications, is called the 1-2-3 attitude sequence or the 'roll-pitch-yaw' convention. In this case, the rotation matrix from frame 1 to frame 2 is given by

$$\begin{aligned} \mathbf{C}_{21}(\theta_3, \theta_2, \theta_1) &= \mathbf{C}_3(\theta_3)\mathbf{C}_2(\theta_2)\mathbf{C}_1(\theta_1) \\ &= \begin{bmatrix} c_2c_3 & c_1s_3 + s_1s_2c_3 & s_1s_3 - c_1s_2c_3 \\ -c_2s_3 & c_1c_3 - s_1s_2s_3 & s_1c_3 + c_1s_2s_3 \\ s_2 & -s_1c_2 & c_1c_2 \end{bmatrix}, \end{aligned} \quad (6.9)$$

where $s_i = \sin\theta_i$, $c_i = \cos\theta_i$.

All Euler sequences have singularities. For instance, if $\gamma = 0$ for the 3-1-3 sequence, then the angles θ and ψ become associated with the same degree of freedom and cannot be uniquely determined.

For the 1-2-3 sequence, a singularity exists at $\theta_2 = \pi/2$. In this case,

$$\mathbf{C}_{21}(\theta_3, \frac{\pi}{2}, \theta_1) = \begin{bmatrix} 0 & \sin(\theta_1 + \theta_3) & -\cos(\theta_1 + \theta_3) \\ 0 & \cos(\theta_1 + \theta_3) & \sin(\theta_1 + \theta_3) \\ 1 & 0 & 0 \end{bmatrix}.$$

Therefore, θ_1 and θ_3 are associated with the same rotation. However, this is only a problem if we want to recover the rotation angles from the rotation matrix.

Infinitesimal Rotations

Consider the 1-2-3 transformation when the angles θ_1, θ_2, θ_3 are small. In this case, we make the approximations $c_i \approx 1$, $s_i \approx \theta_i$ and neglect products of small angles, $\theta_i\theta_j \approx 0$. Then we have

$$\begin{aligned}\mathbf{C}_{21} &\approx \begin{bmatrix} 1 & \theta_3 & -\theta_2 \\ -\theta_3 & 1 & \theta_1 \\ \theta_2 & -\theta_1 & 1 \end{bmatrix} \\ &\approx \mathbf{1} - \boldsymbol{\theta}^{\times}, \end{aligned} \tag{6.10}$$

where

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix},$$

which is referred to as a *rotation vector.*

It is easy to show that the form of the rotation matrix for infinitesimal rotations (i.e., 'small angle approximation') does not depend on the order in which the rotations are performed. For example, we can show that the same result is obtained for a 2-1-3 Euler sequence.

Euler Parameters

Euler's rotation theorem says that the most general motion of a rigid body with one point fixed is a rotation about an axis through that point.

Let us denote the *axis of rotation* by $\mathbf{a} = [a_1 \;\; a_2 \;\; a_3]^T$ and assume that it is a unit vector:

$$\mathbf{a}^T\mathbf{a} = a_1^2 + a_2^2 + a_3^2 = 1. \tag{6.11}$$

The *angle of rotation* is ϕ. We state, without proof, that the rotation matrix in this case is given by

$$\mathbf{C}_{21} = \cos\phi \mathbf{1} + (1 - \cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi \mathbf{a}^{\times}. \tag{6.12}$$

It does not matter in which frame $\mathbf{a}$ is expressed because

$$\mathbf{C}_{21}\mathbf{a} = \mathbf{a}. \tag{6.13}$$

The combination of variables,

$$\eta = \cos\frac{\phi}{2}, \quad \boldsymbol{\varepsilon} = \mathbf{a}\sin\frac{\phi}{2} = \begin{bmatrix} a_1\sin(\phi/2) \\ a_2\sin(\phi/2) \\ a_3\sin(\phi/2) \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}, \tag{6.14}$$

is particularly useful. The four parameters $\{\boldsymbol{\varepsilon}, \eta\}$ are called the *Euler parameters* associated with a rotation[2]. They are not independent because they satisfy the constraint

$$\eta^2 + \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 = 1.$$

[2] These are sometimes referred to as *unit-length quaternions* when stacked as $\mathbf{q} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}$. These are discussed in more detail below.

The rotation matrix can be expressed in terms of the Euler parameters as

$$\begin{aligned}\mathbf{C}_{21} &= (\eta^2 - \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})\mathbf{1} + 2\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T - 2\eta\boldsymbol{\varepsilon}^\times \\ &= \begin{bmatrix} 1-2(\varepsilon_2^2+\varepsilon_3^2) & 2(\varepsilon_1\varepsilon_2+\varepsilon_3\eta) & 2(\varepsilon_1\varepsilon_3-\varepsilon_2\eta) \\ 2(\varepsilon_2\varepsilon_1-\varepsilon_3\eta) & 1-2(\varepsilon_3^2+\varepsilon_1^2) & 2(\varepsilon_2\varepsilon_3+\varepsilon_1\eta) \\ 2(\varepsilon_3\varepsilon_1+\varepsilon_2\eta) & 2(\varepsilon_3\varepsilon_2-\varepsilon_1\eta) & 1-2(\varepsilon_1^2+\varepsilon_2^2)\end{bmatrix}. \end{aligned} \tag{6.15}$$

Euler parameters are useful in many spacecraft applications. There are no singularities associated with them, and the calculation of the rotation matrix does not involve trigonometric functions, which is a significant numerical advantage. The only drawback is the use of four parameters instead of three, as is the case with Euler angles; this makes it challenging to perform some estimation problems because the constraint must be enforced.

Quaternions were first described by Sir William Rowan Hamilton (1805-1865) in 1843 and applied to mechanics in three-dimensional space. Hamilton was an Irish physicist, astronomer, and mathematician, who made important contributions to classical mechanics, optics, and algebra. His studies of mechanical and optical systems led him to discover new mathematical concepts and techniques. His best known contribution to mathematical physics is the reformulation of Newtonian mechanics, now called Hamiltonian mechanics. This work has proven central to the modern study of classical field theories such as electromagnetism, and to the development of quantum mechanics. In pure mathematics, he is best known as the inventor of quaternions.

Quaternions

We will use the notation of Barfoot et al. (2011) for this section. A *quaternion* will be a 4×1 column that may be written as

$$\mathbf{q} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}, \tag{6.16}$$

where $\boldsymbol{\varepsilon}$ is a 3×1 and η is a scalar. The quaternion left-hand compound operator, $+$, and the right-hand compound operator, $\oplus$, will be defined as

$$\mathbf{q}^+ = \begin{bmatrix} \eta\mathbf{1} - \boldsymbol{\varepsilon}^\times & \boldsymbol{\varepsilon} \\ -\boldsymbol{\varepsilon}^T & \eta \end{bmatrix}, \quad \mathbf{q}^\oplus = \begin{bmatrix} \eta\mathbf{1} + \boldsymbol{\varepsilon}^\times & \boldsymbol{\varepsilon} \\ -\boldsymbol{\varepsilon}^T & \eta \end{bmatrix}. \tag{6.17}$$

The inverse operator, -1, will be defined by

$$\mathbf{q}^{-1} = \begin{bmatrix} -\boldsymbol{\varepsilon} \\ \eta \end{bmatrix}. \tag{6.18}$$

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be quaternions. Then some useful identities are

$$\mathbf{u}^+\mathbf{v} \equiv \mathbf{v}^\oplus\mathbf{u}, \tag{6.19}$$

and

$$\begin{aligned} &(\mathbf{u}^+)^T \equiv (\mathbf{u}^+)^{-1} \equiv (\mathbf{u}^{-1})^+, && (\mathbf{u}^\oplus)^T \equiv (\mathbf{u}^\oplus)^{-1} \equiv (\mathbf{u}^{-1})^\oplus, \\ &(\mathbf{u}^+\mathbf{v})^{-1} \equiv \mathbf{v}^{-1+}\mathbf{u}^{-1}, && (\mathbf{u}^\oplus\mathbf{v})^{-1} \equiv \mathbf{v}^{-1\oplus}\mathbf{u}^{-1}, \\ &(\mathbf{u}^+\mathbf{v})^+\mathbf{w} \equiv \mathbf{u}^+(\mathbf{v}^+\mathbf{w}) \equiv \mathbf{u}^+\mathbf{v}^+\mathbf{w}, && (\mathbf{u}^\oplus\mathbf{v})^\oplus\mathbf{w} \equiv \mathbf{u}^\oplus(\mathbf{v}^\oplus\mathbf{w}) \equiv \mathbf{u}^\oplus\mathbf{v}^\oplus\mathbf{w}, \\ &\alpha\mathbf{u}^+ + \beta\mathbf{v}^+ \equiv (\alpha\mathbf{u} + \beta\mathbf{v})^+, && \alpha\mathbf{u}^\oplus + \beta\mathbf{v}^\oplus \equiv (\alpha\mathbf{u} + \beta\mathbf{v})^\oplus, \end{aligned} \tag{6.20}$$

where α and β are scalars. We also have

$$\mathbf{u}^+\mathbf{v}^\oplus \equiv \mathbf{v}^\oplus\mathbf{u}^+. \tag{6.21}$$

The proofs are left to the reader.

Quaternions form a *non-commutative group*[3] under both the $+$ and $\oplus$ operations. Many of the identities above are prerequisites to showing this fact. The identity element of this group, $\boldsymbol{\iota} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$, is such that

$$\boldsymbol{\iota}^+ = \boldsymbol{\iota}^\oplus = \mathbf{1}, \tag{6.22}$$

where $\mathbf{1}$ is the 4×4 identity matrix.

Rotations may be represented in this notation by using a unit-length quaternion, $\mathbf{q}$, such that

$$\mathbf{q}^T\mathbf{q} = 1. \tag{6.23}$$

These form a *sub-group* that can be used to represent rotations.

To rotate a point (in homogeneous form)

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{6.24}$$

to another frame using the rotation, $\mathbf{q}$, we compute

$$\mathbf{u} = \mathbf{q}^+\mathbf{v}^+\mathbf{q}^{-1} = \mathbf{q}^+\mathbf{q}^{-1^\oplus}\mathbf{v} = \mathbf{R}\mathbf{v}, \tag{6.25}$$

where

$$\mathbf{R} = \mathbf{q}^+\mathbf{q}^{-1^\oplus} = \mathbf{q}^{-1^\oplus}\mathbf{q}^+ = \mathbf{q}^{\oplus^T}\mathbf{q}^+ = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{6.26}$$

and $\mathbf{C}$ is the 3×3 rotation matrix with which we are now familiar. We have included various forms for $\mathbf{R}$ to show the different structures this transformation can take.

Josiah Willard Gibbs (1839-1903) was an American scientist who made important theoretical contributions to physics, chemistry, and mathematics. As a mathematician, he invented modern vector calculus (independently of the British scientist Oliver Heaviside, who carried out similar work during the same period). The Gibbs vector is also sometimes known as the *Cayley-Rodrigues parameters*.

Gibbs Vector

Yet another way that we can parameterize rotations is through the *Gibbs vector*. In terms of axis/angle parameters discussed earlier, the Gibbs vector, $\mathbf{g}$, is given by

$$\mathbf{g} = \mathbf{a} \tan \frac{\phi}{2}, \tag{6.27}$$

which we note has a singularity at $\phi = \pi$, so this parameterization does not work well for all angles. The rotation matrix, $\mathbf{C}$, can then be written in terms of the Gibbs vector as

$$\mathbf{C} = \left(\mathbf{1} + \mathbf{g}^\times\right)^{-1}\left(\mathbf{1} - \mathbf{g}^\times\right) = \frac{1}{1 + \mathbf{g}^T\mathbf{g}}\left((1 - \mathbf{g}^T\mathbf{g})\mathbf{1} + 2\mathbf{g}\mathbf{g}^T - 2\mathbf{g}^\times\right). \tag{6.28}$$

[3] The next chapter will discuss group theory as it pertains to rotations in much more detail.

Substituting in the Gibbs vector definition, the right-hand expression becomes

$$\mathbf{C} = \frac{1}{1+\tan^2\frac{\phi}{2}}\left(\left(1-\tan^2\frac{\phi}{2}\right)\mathbf{1} + 2\tan^2\frac{\phi}{2}\mathbf{a}\mathbf{a}^T - 2\tan\frac{\phi}{2}\mathbf{a}^\times\right), \tag{6.29}$$

where we have used that $\mathbf{a}^T\mathbf{a} = 1$. Utilizing that $\left(1+\tan^2\frac{\phi}{2}\right)^{-1} = \cos^2\frac{\phi}{2}$, we have

$$\begin{aligned}\mathbf{C} &= \underbrace{\left(\cos^2\frac{\phi}{2} - \sin^2\frac{\phi}{2}\right)}_{\cos\phi}\mathbf{1} + \underbrace{2\sin^2\frac{\phi}{2}}_{1-\cos\phi}\mathbf{a}\mathbf{a}^T - \underbrace{2\sin\frac{\phi}{2}\cos\frac{\phi}{2}}_{\sin\phi}\mathbf{a}^\times \\ &= \cos\phi\mathbf{1} + (1-\cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi\mathbf{a}^\times, \end{aligned} \tag{6.30}$$

which is our usual expression for the rotation matrix in terms of the axis/angle parameters.

To relate the two expressions for $\mathbf{C}$ in terms of $\mathbf{g}$ given in (6.28), we first note that

$$\left(\mathbf{1}+\mathbf{g}^\times\right)^{-1} = \mathbf{1} - \mathbf{g}^\times + \mathbf{g}^\times\mathbf{g}^\times - \mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times + \cdots = \sum_{n=0}^{\infty}\left(-\mathbf{g}^\times\right)^n. \tag{6.31}$$

Then we observe that

$$\begin{aligned}&\mathbf{g}^T\mathbf{g}\left(\mathbf{1}+\mathbf{g}^\times\right)^{-1} \\ &\quad = (\mathbf{g}^T\mathbf{g})\mathbf{1} - \underbrace{(\mathbf{g}^T\mathbf{g})\mathbf{g}^\times}_{-\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times} + \underbrace{(\mathbf{g}^T\mathbf{g})\mathbf{g}^\times\mathbf{g}^\times}_{-\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times} - \underbrace{(\mathbf{g}^T\mathbf{g})\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times}_{-\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times} + \cdots \\ &\quad = \mathbf{1} + \mathbf{g}\mathbf{g}^T - \mathbf{g}^\times - \left(\mathbf{1}+\mathbf{g}^\times\right)^{-1}, \end{aligned} \tag{6.32}$$

where we have used the following manipulation several times:

$$(\mathbf{g}^T\mathbf{g})\mathbf{g}^\times = \left(-\mathbf{g}^\times\mathbf{g}^\times + \mathbf{g}\mathbf{g}^T\right)\mathbf{g}^\times = -\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times + \mathbf{g}\underbrace{\mathbf{g}^T\mathbf{g}^\times}_{\mathbf{0}} = -\mathbf{g}^\times\mathbf{g}^\times\mathbf{g}^\times. \tag{6.33}$$

Therefore we have that

$$\left(1+\mathbf{g}^T\mathbf{g}\right)\left(\mathbf{1}+\mathbf{g}^\times\right)^{-1} = \mathbf{1} + \mathbf{g}\mathbf{g}^T - \mathbf{g}^\times, \tag{6.34}$$

and thus

$$\begin{aligned}&\left(1+\mathbf{g}^T\mathbf{g}\right)\underbrace{\left(\mathbf{1}+\mathbf{g}^\times\right)^{-1}\left(\mathbf{1}-\mathbf{g}^\times\right)}_{\mathbf{C}} = \left(\mathbf{1}+\mathbf{g}\mathbf{g}^T - \mathbf{g}^\times\right)\left(\mathbf{1}-\mathbf{g}^\times\right) \\ &= \mathbf{1} + \mathbf{g}\mathbf{g}^T - 2\mathbf{g}^\times - \mathbf{g}\underbrace{\mathbf{g}^T\mathbf{g}^\times}_{\mathbf{0}} + \underbrace{\mathbf{g}^\times\mathbf{g}^\times}_{-\mathbf{g}^T\mathbf{g}\mathbf{1}+\mathbf{g}\mathbf{g}^T} = \left(1-\mathbf{g}^T\mathbf{g}\right)\mathbf{1} + 2\mathbf{g}\mathbf{g}^T - 2\mathbf{g}^\times. \end{aligned} \tag{6.35}$$

Dividing both sides by $(1+\mathbf{g}^T\mathbf{g})$ provides the desired result.

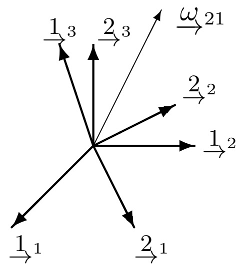
6.2.4 Rotational Kinematics

In the last section, we showed that the orientation of one frame $\underrightarrow{\mathcal{F}}_2$ with respect to another $\underrightarrow{\mathcal{F}}_1$ could be parameterized in different ways. In other words, the rotation matrix could be written as a function of Euler angles or Euler parameters. However, in most applications the orientation changes with time and thus we must introduce the vehicle *kinematics*, which form an important part of the vehicle's motion model.

We will first introduce the concept of angular velocity, then acceleration in a rotating frame. We will finish with expressions that relate the rate of change of the orientation parameterization to angular velocity.

Angular Velocity

Let frame $\underrightarrow{\mathcal{F}}_2$ rotate with respect to frame $\underrightarrow{\mathcal{F}}_1$. The angular velocity of frame 2 with respect to frame 1 is denoted by $\underrightarrow{\omega}_{21}$. The angular velocity of frame 1 with respect to 2 is $\underrightarrow{\omega}_{12} = -\underrightarrow{\omega}_{21}$.



The magnitude of $\underrightarrow{\omega}_{21}$, $|\underrightarrow{\omega}_{21}| = \sqrt{(\underrightarrow{\omega}_{21} \cdot \underrightarrow{\omega}_{21})}$, is the rate of rotation. The direction of $\underrightarrow{\omega}_{21}$ (i.e., the unit vector in the direction of $\underrightarrow{\omega}_{21}$, which is $|\underrightarrow{\omega}_{21}|^{-1}\,\underrightarrow{\omega}_{21}$) is the *instantaneous* axis of rotation.

Observers in the frames $\underrightarrow{\mathcal{F}}_2$ and $\underrightarrow{\mathcal{F}}_1$ do not see the same motion because of their own relative motions. Let us denote the *vector time derivative* as seen in $\underrightarrow{\mathcal{F}}_1$ by $(\cdot)^{\bullet}$ and that seen in $\underrightarrow{\mathcal{F}}_2$ by $(\cdot)^{\circ}$. Therefore,

$$\underrightarrow{\mathcal{F}}^{\bullet}_1 = \underrightarrow{0}\,, \quad \underrightarrow{\mathcal{F}}^{\circ}_2 = \underrightarrow{0}.$$

It can be shown that

$$\underrightarrow{2}^{\bullet}_1 = \underrightarrow{\omega}_{21} \times \underrightarrow{2}_1, \quad \underrightarrow{2}^{\bullet}_2 = \underrightarrow{\omega}_{21} \times \underrightarrow{2}_2, \quad \underrightarrow{2}^{\bullet}_3 = \underrightarrow{\omega}_{21} \times \underrightarrow{2}_3,$$

or equivalently

$$\begin{bmatrix} \underrightarrow{2}^{\bullet}_1 & \underrightarrow{2}^{\bullet}_2 & \underrightarrow{2}^{\bullet}_3 \end{bmatrix} = \underrightarrow{\omega}_{21} \times \begin{bmatrix} \underrightarrow{2}_1 & \underrightarrow{2}_2 & \underrightarrow{2}_3 \end{bmatrix},$$

or

$$\underrightarrow{\mathcal{F}}^{\bullet T}_2 = \underrightarrow{\omega}_{21} \times \underrightarrow{\mathcal{F}}^{T}_2. \tag{6.36}$$

We want to determine the time derivative of an arbitrary vector expressed in both frames:

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

Therefore, the time derivative as seen in $\underrightarrow{\mathcal{F}}_1$ is

$$\underrightarrow{r}^{\bullet} = \underrightarrow{\mathcal{F}}_1^{\bullet T} \mathbf{r}_1 + \underrightarrow{\mathcal{F}}_1^T \dot{\mathbf{r}}_1 = \underrightarrow{\mathcal{F}}_1^T \dot{\mathbf{r}}_1. \tag{6.37}$$

In a similar way,

$$\underrightarrow{r}^{\circ} = \underrightarrow{\mathcal{F}}_2^{\circ T} \mathbf{r}_2 + \underrightarrow{\mathcal{F}}_2^T \mathring{\mathbf{r}}_2 = \underrightarrow{\mathcal{F}}_2^T \mathring{\mathbf{r}}_2 = \underrightarrow{\mathcal{F}}_2^T \dot{\mathbf{r}}_2. \tag{6.38}$$

(Note that for nonvectors, $(\dot{\ }) = (\mathring{\ })$, i.e., $\mathring{\mathbf{r}}_2 = \dot{\mathbf{r}}_2$.)

Alternatively, the time derivative as seen in $\underrightarrow{\mathcal{F}}_1$, but expressed in $\underrightarrow{\mathcal{F}}_2$, is

$$\begin{aligned} \underrightarrow{r}^{\bullet} &= \underrightarrow{\mathcal{F}}_2^T \dot{\mathbf{r}}_2 + \underrightarrow{\mathcal{F}}_2^{\bullet T} \mathbf{r}_2 \\ &= \underrightarrow{\mathcal{F}}_2^T \dot{\mathbf{r}}_2 + \underrightarrow{\omega}_{21} \times \underrightarrow{\mathcal{F}}_2^T \mathbf{r}_2 \\ &= \underrightarrow{r}^{\circ} + \underrightarrow{\omega}_{21} \times \underrightarrow{r}. \end{aligned} \tag{6.39}$$

The above is true for any vector $\underrightarrow{r}$. The most important application occurs when $\underrightarrow{r}$ denotes position, $\underrightarrow{\mathcal{F}}_1$ is a nonrotating inertial reference frame, and $\underrightarrow{\mathcal{F}}_2$ is a frame that rotates with a body, vehicle, etc. In this case, (6.39) expresses the velocity in the inertial frame in terms of the motion in the second frame.

Now, express the angular velocity in $\underrightarrow{\mathcal{F}}_2$:

$$\underrightarrow{\omega}_{21} = \underrightarrow{\mathcal{F}}_2^T \boldsymbol{\omega}_2^{21}. \tag{6.40}$$

Therefore,

$$\begin{aligned} \underrightarrow{r}^{\bullet} = \underrightarrow{\mathcal{F}}_1^T \dot{\mathbf{r}}_1 &= \underrightarrow{\mathcal{F}}_2^T \dot{\mathbf{r}}_2 + \underrightarrow{\omega}_{21} \times \underrightarrow{r} \\ &= \underrightarrow{\mathcal{F}}_2^T \dot{\mathbf{r}}_2 + \underrightarrow{\mathcal{F}}_2^T \boldsymbol{\omega}_2^{21^\times} \mathbf{r}_2 \\ &= \underrightarrow{\mathcal{F}}_2^T (\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_2^{21^\times} \mathbf{r}_2). \end{aligned} \tag{6.41}$$

If we want to express the 'inertial time derivative' (that seen in $\underrightarrow{\mathcal{F}}_1$) in $\underrightarrow{\mathcal{F}}_1$, then we can use the rotation matrix $\mathbf{C}_{12}$:

$$\dot{\mathbf{r}}_1 = \mathbf{C}_{12}(\dot{\mathbf{r}}_2 + \boldsymbol{\omega}_2^{21^\times} \mathbf{r}_2). \tag{6.42}$$

Acceleration

Let us denote the *velocity* by

$$\underrightarrow{v} = \underrightarrow{r}^{\bullet} = \underrightarrow{r}^{\circ} + \underrightarrow{\omega}_{21} \times \underrightarrow{r}.$$

The *acceleration* can be calculated by applying (6.39) to $\underrightarrow{v}$:

$$\begin{aligned}\underrightarrow{r}^{\bullet\bullet} = \underrightarrow{v}^{\bullet} &= \underrightarrow{v}^{\circ} + \underrightarrow{\omega}_{21} \times \underrightarrow{v} \\ &= (\underrightarrow{r}^{\circ\circ} + \underrightarrow{\omega}_{21} \times \underrightarrow{r}^{\circ} + \underrightarrow{\omega}^{\circ}_{21} \times \underrightarrow{r}) \\ &\qquad\qquad + (\underrightarrow{\omega}_{21} \times \underrightarrow{r}^{\circ} + \underrightarrow{\omega}_{21} \times (\underrightarrow{\omega}_{21} \times \underrightarrow{r})) \\ &= \underrightarrow{r}^{\circ\circ} \; + \; 2\underrightarrow{\omega}_{21} \times \underrightarrow{r}^{\circ} \; + \; \underrightarrow{\omega}^{\circ}_{21} \times \underrightarrow{r} \; + \; \underrightarrow{\omega}_{21} \times (\underrightarrow{\omega}_{21} \times \underrightarrow{r}).\end{aligned} \tag{6.43}$$

The matrix equivalent in terms of components can be had by making the following substitutions:

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

The result for the components is

$$\ddot{\mathbf{r}}_1 = \mathbf{C}_{12} \left[\ddot{\mathbf{r}}_2 + 2\boldsymbol{\omega}_2^{21^\times} \dot{\mathbf{r}}_2 + \dot{\boldsymbol{\omega}}_2^{21^\times} \mathbf{r}_2 + \boldsymbol{\omega}_2^{21^\times} \boldsymbol{\omega}_2^{21^\times} \mathbf{r}_2 \right] . \tag{6.44}$$

The various terms in the expression for the acceleration have been given special names:

$$\begin{aligned} \underrightarrow{r}^{\circ\circ} &: \text{acceleration with respect to } \underrightarrow{\boldsymbol{\mathcal{F}}}_2 \\ 2\underrightarrow{\omega}_{21} \times \underrightarrow{r}^{\circ} &: \text{Coriolis acceleration} \\ \underrightarrow{\omega}^{\circ}_{21} \times \underrightarrow{r} &: \text{angular acceleration} \\ \underrightarrow{\omega}_{21} \times \left(\underrightarrow{\omega}_{21} \times \underrightarrow{r} \right) &: \text{centripetal acceleration} \end{aligned}$$

Angular Velocity Given Rotation Matrix

Begin with (6.3), which relates relates two reference frames via the rotation matrix:

$$\underrightarrow{\boldsymbol{\mathcal{F}}}_1^T = \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \mathbf{C}_{21}.$$

Now take the time derivative of both sides as seen in $\underrightarrow{\boldsymbol{\mathcal{F}}}_1$:

$$\underrightarrow{0} = \underrightarrow{\boldsymbol{\mathcal{F}}}_2^{\bullet T} \mathbf{C}_{21} + \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \dot{\mathbf{C}}_{21}.$$

Substitute (6.36) for $\underrightarrow{\boldsymbol{\mathcal{F}}}_2^{\bullet T}$:

$$\underrightarrow{0} = \underrightarrow{\omega}_{21} \times \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \mathbf{C}_{21} + \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \dot{\mathbf{C}}_{21}.$$

Now use (6.40) to get

$$\begin{aligned} \underrightarrow{0} &= \boldsymbol{\omega}_2^{21^T} \underrightarrow{\boldsymbol{\mathcal{F}}}_2 \times \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \mathbf{C}_{21} + \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \dot{\mathbf{C}}_{21} \\ &= \underrightarrow{\boldsymbol{\mathcal{F}}}_2^T \left(\boldsymbol{\omega}_2^{21^\times} \mathbf{C}_{21} + \dot{\mathbf{C}}_{21} \right) . \end{aligned}$$

Siméon Denis Poisson (1781-1840) was a French mathematician, geometer, and physicist.

Therefore, we conclude that

$$\dot{\mathbf{C}}_{21} = -\boldsymbol{\omega}_2^{21^\times} \mathbf{C}_{21}, \tag{6.45}$$

which is known as *Poisson's equation.* Given the angular velocity as measured in the frame $\underrightarrow{\mathcal{F}}_2$, the rotation matrix relating $\underrightarrow{\mathcal{F}}_1$ to $\underrightarrow{\mathcal{F}}_2$ can be determined by integrating the above expression[4].

We can also rearrange to obtain an explicit function of $\boldsymbol{\omega}_2^{21}$:

$$\begin{aligned}\boldsymbol{\omega}_2^{21^\times} &= -\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^{-1} \\ &= -\dot{\mathbf{C}}_{21}\mathbf{C}_{21}^T,\end{aligned} \tag{6.46}$$

which gives the angular velocity when the rotation matrix is known as a function of time.

Euler Angles

Consider the 1-2-3 Euler angle sequence and its associated rotation matrix. In this case, (6.46) becomes

$$\boldsymbol{\omega}_2^{21^\times} = -\mathbf{C}_3\mathbf{C}_2\dot{\mathbf{C}}_1\mathbf{C}_1^T\mathbf{C}_2^T\mathbf{C}_3^T - \mathbf{C}_3\dot{\mathbf{C}}_2\mathbf{C}_2^T\mathbf{C}_3^T - \dot{\mathbf{C}}_3\mathbf{C}_3^T. \tag{6.47}$$

Then, using

$$-\dot{\mathbf{C}}_i\mathbf{C}_i^T = \mathbf{1}_i^\times\dot{\theta}_i, \tag{6.48}$$

for each principal axis rotation (where $\mathbf{1}_i$ is column i of $\mathbf{1}$) and the identity

$$(\mathbf{C}_i\mathbf{r})^\times \equiv \mathbf{C}_i\mathbf{r}^\times\mathbf{C}_i^T, \tag{6.49}$$

we can show that

$$\boldsymbol{\omega}_2^{21^\times} = \left(\mathbf{C}_3\mathbf{C}_2\mathbf{1}_1\dot{\theta}_1\right)^\times + \left(\mathbf{C}_3\mathbf{1}_2\dot{\theta}_2\right)^\times + \left(\mathbf{1}_3\dot{\theta}_3\right)^\times, \tag{6.50}$$

which can be simplified to

$$\begin{aligned}\boldsymbol{\omega}_2^{21} &= \underbrace{\begin{bmatrix}\mathbf{C}_3(\theta_3)\mathbf{C}_2(\theta_2)\mathbf{1}_1 & \mathbf{C}_3(\theta_3)\mathbf{1}_2 & \mathbf{1}_3\end{bmatrix}}_{\mathbf{S}(\theta_2,\theta_3)}\underbrace{\begin{bmatrix}\dot{\theta}_1\\ \dot{\theta}_2\\ \dot{\theta}_3\end{bmatrix}}_{\dot{\boldsymbol{\theta}}} \\ &= \mathbf{S}(\theta_2,\theta_3)\dot{\boldsymbol{\theta}},\end{aligned} \tag{6.51}$$

which gives the angular velocity in terms of the Euler angles and the *Euler rates*, $\dot{\boldsymbol{\theta}}$. In scalar detail we have

$$\mathbf{S}(\theta_2,\theta_3) = \begin{bmatrix}\cos\theta_2\cos\theta_3 & \sin\theta_3 & 0\\ -\cos\theta_2\sin\theta_3 & \cos\theta_3 & 0\\ \sin\theta_2 & 0 & 1\end{bmatrix}. \tag{6.52}$$

[4] This is termed ‘strapdown navigation’ because the sensors that measure $\boldsymbol{\omega}_2^{21}$ are strapped down in the rotating frame, $\underrightarrow{\mathcal{F}}_2$.

By inverting the matrix $\mathbf{S}$, we arrive at a system of differential equations that can be integrated to yield the Euler angles, assuming $\boldsymbol{\omega}_2^{21}$ is known:

$$\begin{aligned}\dot{\boldsymbol{\theta}} &= \mathbf{S}^{-1}(\theta_2,\theta_3)\boldsymbol{\omega}_2^{21} \\ &= \begin{bmatrix} \sec\theta_2\cos\theta_3 & -\sec\theta_2\sin\theta_3 & 0 \\ \sin\theta_3 & \cos\theta_3 & 0 \\ -\tan\theta_2\cos\theta_3 & \tan\theta_2\sin\theta_3 & 1 \end{bmatrix} \boldsymbol{\omega}_2^{21}. \end{aligned} \tag{6.53}$$

Note that $\mathbf{S}^{-1}$ does not exist at $\theta_2 = \pi/2$, which is precisely the singularity associated with the 1-2-3 sequence.

It should be noted that the above developments hold true in general for any Euler sequence. If we pick an α-β-γ set,

$$\mathbf{C}_{21}(\theta_1,\theta_2,\theta_3) = \mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{C}_\alpha(\theta_1), \tag{6.54}$$

then

$$\mathbf{S}(\theta_2,\theta_3) = \begin{bmatrix}\mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{1}_\alpha & \mathbf{C}_\gamma(\theta_3)\mathbf{1}_\beta & \mathbf{1}_\gamma\end{bmatrix}, \tag{6.55}$$

and $\mathbf{S}^{-1}$ does not exist at the singularities of $\mathbf{S}$.

6.2.5 Perturbing Rotations

Now that we have some basic notation built up for handling quantities in three-dimensional space, we will turn our focus to an issue that is often handled incorrectly or simply ignored altogether. We have shown in the previous section that the state of a single-body vehicle involves a translation, which has three degrees of freedom, as well as a rotation, which also has three degrees of freedom. The problem is that the degrees of freedom associated with rotations are a bit unique and must be handled carefully. The reason is that rotations do not live in a *vector space*[5]; rather, they form the *non-commutative group* called $SO(3)$.

As we have seen above, there are many ways of representing rotations mathematically, including rotation matrices, axis-angle formulations, Euler angles, and Euler parameters/unit-length quaternions. The most important fact to remember is that all these representations have the same underlying rotation, which only has three degrees of freedom. A 3×3 rotation matrix has nine elements, but only three are independent. Euler parameters have four scalar parameters, but only three are independent. Of all the common rotation representations, Euler angles are the only ones that have exactly three parameters; the problem is that Euler sequences have singularities, so for some problems, one must choose an appropriate sequence that avoids the singularities.

The fact that rotations do not live in a vector space is actually quite fundamental when it comes to linearizing motion and observation models involving rotations. What are we to do about linearizing rotations?

[5] Here we mean a vector space in the sense of linear algebra.

Fortunately, there is a way forwards. The key is to consider what is happening on a small, in fact infinitesimal, level. We will begin by deriving a few key identities and then turn to linearizing a rotation matrix built from a sequence of Euler angles.

Some Key Identities

Euler's rotation theorem allows us to write a rotation matrix, $\mathbf{C}$, in terms of a rotation about an axis, $\mathbf{a}$, through an angle, ϕ:

$$\mathbf{C} = \cos\phi\mathbf{1} + (1-\cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi\mathbf{a}^\times. \tag{6.56}$$

We now take the partial derivative of $\mathbf{C}$ with respect to the angle, ϕ:

$$\begin{aligned}
\frac{\partial \mathbf{C}}{\partial \phi} &= -\sin\phi\mathbf{1} + \sin\phi\mathbf{a}\mathbf{a}^T - \cos\phi\mathbf{a}^\times && \text{(6.57a)}\\
&= \sin\phi\underbrace{\left(-\mathbf{1} + \mathbf{a}\mathbf{a}^T\right)}_{\mathbf{a}^\times\mathbf{a}^\times} - \cos\phi\mathbf{a}^\times && \text{(6.57b)}\\
&= -\cos\phi\mathbf{a}^\times - (1-\cos\phi)\underbrace{\mathbf{a}^\times\mathbf{a}}_{\mathbf{0}}\mathbf{a}^T + \sin\phi\mathbf{a}^\times\mathbf{a}^\times && \text{(6.57c)}\\
&= -\mathbf{a}^\times\underbrace{\left(\cos\phi\mathbf{1} + (1-\cos\phi)\mathbf{a}\mathbf{a}^T - \sin\phi\mathbf{a}^\times\right)}_{\mathbf{C}}. && \text{(6.57d)}
\end{aligned}$$

Thus, our first important identity is

$$\frac{\partial \mathbf{C}}{\partial \phi} \equiv -\mathbf{a}^\times\mathbf{C}. \tag{6.58}$$

An immediate application of this is that for any principal-axis rotation, about axis α, we have

$$\frac{\partial \mathbf{C}_\alpha(\theta)}{\partial \theta} \equiv -\mathbf{1}_\alpha^\times\mathbf{C}_\alpha(\theta), \tag{6.59}$$

where $\mathbf{1}_\alpha$ is column α of the identity matrix.

Let us now consider an α-β-γ Euler sequence:

$$\mathbf{C}(\boldsymbol{\theta}) = \mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{C}_\alpha(\theta_1), \tag{6.60}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. Furthermore, we select an arbitrary constant vector, $\mathbf{v}$. Applying (6.59), we have

$$\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\theta_3} = -\mathbf{1}_\gamma^\times\mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{C}_\alpha(\theta_1)\mathbf{v} \;=\; \left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)^\times\mathbf{1}_\gamma, \tag{6.61a}$$

$$\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\theta_2} = -\mathbf{C}_\gamma(\theta_3)\mathbf{1}_\beta^\times\mathbf{C}_\beta(\theta_2)\mathbf{C}_\alpha(\theta_1)\mathbf{v} \;=\; \left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)^\times\mathbf{C}_\gamma(\theta_3)\mathbf{1}_\beta, \tag{6.61b}$$

$$\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\theta_1} = -\mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{1}_\alpha^\times\mathbf{C}_\alpha(\theta_1)\mathbf{v} \;=\; \left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)^\times\mathbf{C}_\gamma(\theta_3)\mathbf{C}_\beta(\theta_2)\mathbf{1}_\alpha, \tag{6.61c}$$

where we have made use of the two general identities

$$\mathbf{r}^{\times}\mathbf{s} \equiv -\mathbf{s}^{\times}\mathbf{r}, \tag{6.62a}$$
$$(\mathbf{R}\mathbf{s})^{\times} \equiv \mathbf{R}\mathbf{s}^{\times}\mathbf{R}^{T} \tag{6.62b}$$

for any vectors $\mathbf{r}$, $\mathbf{s}$ and any rotation matrix $\mathbf{R}$. Combining the results in (6.61), we have

$$\begin{aligned}\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\boldsymbol{\theta}} &= \left[\tfrac{\partial(\mathbf{C}(\boldsymbol{\theta})\mathbf{v})}{\partial\theta_1} \quad \tfrac{\partial(\mathbf{C}(\boldsymbol{\theta})\mathbf{v})}{\partial\theta_2} \quad \tfrac{\partial(\mathbf{C}(\boldsymbol{\theta})\mathbf{v})}{\partial\theta_3}\right] \\ &= \left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)^{\times} \underbrace{\left[\mathbf{C}_{\gamma}(\theta_3)\mathbf{C}_{\beta}(\theta_2)\mathbf{1}_{\alpha} \quad \mathbf{C}_{\gamma}(\theta_3)\mathbf{1}_{\beta} \quad \mathbf{1}_{\gamma}\right]}_{\mathbf{S}(\theta_2,\theta_3)},\end{aligned} \tag{6.63}$$

and thus another very important identity that we can state is

$$\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\boldsymbol{\theta}} \equiv \Big(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\Big)^{\times}\mathbf{S}(\theta_2,\theta_3), \tag{6.64}$$

which we note is true regardless of the choice of Euler set. This will prove critical in the next section, when we discuss linearization of a rotation matrix.

Perturbing a Rotation Matrix

Let us return to first principles and consider carefully how to linearize a rotation. If we have a function, $\mathbf{f}(\mathbf{x})$, of some variable, $\mathbf{x}$, then perturbing $\mathbf{x}$ slightly from its nominal value, $\bar{\mathbf{x}}$, by an amount $\delta\mathbf{x}$ will result in a change in the function. We can express this in terms of a Taylor-series expansion of $\mathbf{f}$ about $\bar{\mathbf{x}}$:

$$\mathbf{f}(\bar{\mathbf{x}} + \delta\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}}) + \left.\frac{\partial\mathbf{f}(\mathbf{x})}{\partial\mathbf{x}}\right|_{\bar{\mathbf{x}}} \delta\mathbf{x} + (\text{higher order terms}) \tag{6.65}$$

and so if $\delta\mathbf{x}$ is small, a 'first-order' approximation is

$$\mathbf{f}(\bar{\mathbf{x}} + \delta\mathbf{x}) \approx \mathbf{f}(\bar{\mathbf{x}}) + \left.\frac{\partial\mathbf{f}(\mathbf{x})}{\partial\mathbf{x}}\right|_{\bar{\mathbf{x}}} \delta\mathbf{x}. \tag{6.66}$$

This presupposes that $\delta\mathbf{x}$ is not constrained in any way. The trouble with carrying out the same process with rotations is that most of the representations involve constraints and thus are not easily perturbed (without enforcing the constraint). The notable exceptions are the Euler angle sets. These contain exactly three parameters, and thus each can be varied independently. For this reason, we choose to use Euler angles in our perturbation of functions involving rotations.

Consider perturbing $\mathbf{C}(\boldsymbol{\theta})\mathbf{v}$ with respect to Euler angles $\boldsymbol{\theta}$, where $\mathbf{v}$ is an arbitrary constant vector. Letting $\bar{\boldsymbol{\theta}} = (\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3)$ and $\delta\boldsymbol{\theta} = (\delta\theta_1, \delta\theta_2, \delta\theta_3)$, then applying a first-order Taylor-series approximation, we have

$$
\begin{aligned}
\mathbf{C}(\bar{\boldsymbol{\theta}} + \delta\boldsymbol{\theta})\mathbf{v} &\approx \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v} + \left.\frac{\partial\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)}{\partial\boldsymbol{\theta}}\right|_{\bar{\boldsymbol{\theta}}} \delta\boldsymbol{\theta} \\
&= \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v} + \left.\Big(\left(\mathbf{C}(\boldsymbol{\theta})\mathbf{v}\right)^{\times} \mathbf{S}(\theta_2, \theta_3)\Big)\right|_{\bar{\boldsymbol{\theta}}} \delta\boldsymbol{\theta} \\
&= \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v} + \left(\mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v}\right)^{\times} \mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\,\delta\boldsymbol{\theta} \\
&= \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v} - \left(\mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\,\delta\boldsymbol{\theta}\right)^{\times} \left(\mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v}\right) \\
&= \Big(\mathbf{1} - \left(\mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\,\delta\boldsymbol{\theta}\right)^{\times}\Big)\, \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v},
\end{aligned}
\tag{6.67}
$$

where we have used (6.64) to get to the second line. Observing that $\mathbf{v}$ is arbitrary, we can drop it from both sides and write

$$
\mathbf{C}(\bar{\boldsymbol{\theta}} + \delta\boldsymbol{\theta}) \approx \underbrace{\Big(\mathbf{1} - \left(\mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\,\delta\boldsymbol{\theta}\right)^{\times}\Big)}_{\text{infinitesimal rot. mat.}}\, \mathbf{C}(\bar{\boldsymbol{\theta}}),
\tag{6.68}
$$

which we see is the product (not the sum) of an infinitesimal rotation matrix and the unperturbed rotation matrix, $\mathbf{C}(\bar{\boldsymbol{\theta}})$. Notationally, it is simpler to write

$$
\mathbf{C}(\bar{\boldsymbol{\theta}} + \delta\boldsymbol{\theta}) \approx \left(\mathbf{1} - \delta\boldsymbol{\phi}^{\times}\right)\, \mathbf{C}(\bar{\boldsymbol{\theta}}),
\tag{6.69}
$$

with $\delta\boldsymbol{\phi} = \mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\,\delta\boldsymbol{\theta}$. Equation (6.68) is extremely important. It tells us exactly how to perturb a rotation matrix (in terms of perturbations to its Euler angles) when it appears inside any function.

Example 6.1 The following example shows how we can apply our linearized rotation expression in an arbitrary expression. Suppose we have a scalar function, J, given by

$$
J(\boldsymbol{\theta}) = \mathbf{u}^T \mathbf{C}(\boldsymbol{\theta})\mathbf{v},
\tag{6.70}
$$

where $\mathbf{u}$ and $\mathbf{v}$ are arbitrary vectors. Applying our approach to linearizing rotations, we have

$$
J(\bar{\boldsymbol{\theta}} + \delta\boldsymbol{\theta}) \approx \mathbf{u}^T \left(\mathbf{1} - \delta\boldsymbol{\phi}^{\times}\right)\, \mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v} = \underbrace{\mathbf{u}^T\mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v}}_{J(\bar{\boldsymbol{\theta}})} + \underbrace{\mathbf{u}^T \left(\mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v}\right)^{\times}\, \delta\boldsymbol{\phi}}_{\delta J(\delta\boldsymbol{\theta})},
\tag{6.71}
$$

so that the linearized function is

$$
\delta J(\delta\boldsymbol{\theta}) = \underbrace{\Big(\mathbf{u}^T \left(\mathbf{C}(\bar{\boldsymbol{\theta}})\mathbf{v}\right)^{\times} \mathbf{S}(\bar{\theta}_2, \bar{\theta}_3)\Big)}_{\text{constant}}\, \delta\boldsymbol{\theta},
\tag{6.72}
$$

where we see that the factor in front of $\delta\boldsymbol{\theta}$ is indeed constant; in fact, it is $\left.\frac{\partial J}{\partial \boldsymbol{\theta}}\right|_{\bar{\boldsymbol{\theta}}}$, the Jacobian of J with respect to $\boldsymbol{\theta}$.