3D Attitude for Aircraft Navigation, Guidance and Control

Miguel Alonso
TEAYC-TL2

April 22, 2024

Contents

Chapter 1

Linear Algebra

1.1 Vectors

1.1.1 Vector Spaces

A *field* is a mathematical structure consisting of a set $\mathcal{F} = \{a, b, c, ...\}$ and the two following operations:

a) *Addition*, which takes two elements $a, b \in \mathcal{F}$ and yields another element $c = a + b \in \mathcal{F}$. The addition operation must satisfy the following properties:

   1) Associativity:
   $$a + (b + c) = (a + b) + c$$

   2) Commutativity:
   $$a + b = b + a$$

   3) Existence of an identity element:
   $$\exists \{0 \in \mathcal{F} \mid \forall a \in \mathcal{F}, a + 0 = a\}$$

   4) Existence of an inverse element:
   $$\forall a \in \mathcal{F},\ \exists \{-a \in \mathcal{F} \mid a + (-a) = 0\}$$

b) *Multiplication*, which takes two elements $a, b \in \mathcal{F}$ and yields another element $c = ab \in \mathcal{F}$. The multiplication operation must satisfy the following properties:

   1) Associativity:
   $$a(bc) = (ab)c$$

   2) Commutativity:
   $$ab = ba$$

   3) Existence of an identity element:
   $$\exists \{1 \in \mathcal{F} \mid \forall a \in \mathcal{F}, (a)(1) = a\}$$

4) Existence of an inverse element:

$$\forall a \in \mathcal{F},\ \exists \left\{ a^{-1} \in \mathcal{F} \mid aa^{-1} = 1 \right\}$$

5) Distributivity over addition:

$$(a+b)c = ac + bc$$

Some familiar examples of fields are those of rational numbers ($\mathbb{Q}$), real numbers ($\mathbb{R}$) and complex numbers ($\mathbb{C}$).

A *vector space* is a mathematical structure consisting of a set $\mathcal{V} = \{\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}, \vec{\boldsymbol{z}}, \dots\}$, whose members are called *vectors*, a field $\mathcal{F} = \{a, b, c, \dots\}$, whose members are called *scalars*, and the two following operations:

a) *Vector addition*, which takes two vectors $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in \mathcal{V}$ and produces another vector $\vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} + \vec{\boldsymbol{w}} \in \mathcal{V}$. The vector addition operation must satisfy the following properties:

   1) Associativity:
   $$\vec{\boldsymbol{v}} + (\vec{\boldsymbol{w}} + \vec{\boldsymbol{z}}) = (\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) + \vec{\boldsymbol{z}} \tag{1.1}$$

   2) Commutativity:
   $$\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}} = \vec{\boldsymbol{w}} + \vec{\boldsymbol{v}} \tag{1.2}$$

   3) Existence of an identity element:
   $$\exists \left\{ \vec{\boldsymbol{0}} \in \mathcal{V} \mid \forall \vec{\boldsymbol{v}} \in \mathcal{V}, \vec{\boldsymbol{v}} + \vec{\boldsymbol{0}} = \vec{\boldsymbol{v}} \right\} \tag{1.3}$$

   4) Existence of an inverse element:
   $$\forall \vec{\boldsymbol{v}} \in \mathcal{V},\ \exists \left\{ -\vec{\boldsymbol{v}} \in \mathcal{V} \mid \vec{\boldsymbol{v}} + (-\vec{\boldsymbol{v}}) = \vec{\boldsymbol{0}} \right\} \tag{1.4}$$

b) *Scalar multiplication*, which takes a scalar $a \in \mathcal{F}$ and a vector $\vec{\boldsymbol{v}} \in \mathcal{V}$, and produces another vector $\vec{\boldsymbol{w}} = a\vec{\boldsymbol{v}} \in \mathcal{V}$. The scalar multiplication operation must satisfy the following properties:

   1) Compatibility with field multiplication:
   $$a(b\vec{\boldsymbol{v}}) = (ab)\vec{\boldsymbol{v}} \tag{1.5}$$

   2) Existence of an identity element:
   $$1\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}, \text{ where 1 denotes the multiplicative identity in } \mathcal{F} \tag{1.6}$$

   3) Distributivity with respect to vector addition
   $$a(\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) = a\vec{\boldsymbol{v}} + a\vec{\boldsymbol{w}} \tag{1.7}$$

4) Distributivity with respect to field addition

$$(a+b)\vec{\boldsymbol{v}} = a\vec{\boldsymbol{v}} + b\vec{\boldsymbol{v}} \tag{1.8}$$

A set of vectors $\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \dots, \vec{\boldsymbol{v}}_n\} \in \mathcal{V}$ is said to be *linearly independent* if the following equation is only satisfied by $a_i = 0$ for $i = 1, 2, \dots, n$:

$$\sum_{i=1}^{n} a_i \vec{\boldsymbol{v}}_i = \vec{\boldsymbol{0}} \tag{1.9}$$

An ordered set of linearly independent vectors $\mathcal{E} = \{\vec{\boldsymbol{e}}_1, \vec{\boldsymbol{e}}_2, \dots, \vec{\boldsymbol{e}}_n\} \in \mathcal{V}$ is a *basis* of $\mathcal{V}$ if it *spans* $\mathcal{V}$, that is, if every vector $\vec{\boldsymbol{v}} \in \mathcal{V}$ can be written as a linear combination of the elements in $\mathcal{E}$:

$$\vec{\boldsymbol{v}} = \sum_{i=1}^{n} v_{(i)} \vec{\boldsymbol{e}}_i \tag{1.10}$$

Where the ordered set $\left\{v_{(1)}, v_{(2)}, \dots v_{(n)}\right\}$ is unique and its elements are called the *components* or *coordinates* of $\vec{\boldsymbol{v}}$ in $\mathcal{E}$. When we specify $\vec{\boldsymbol{v}}$ in terms of its components in basis $\mathcal{E}$, we say $\vec{\boldsymbol{v}}$ is *projected* or *resolved* in $\mathcal{E}$.

A vector space $\mathcal{V}$ can generally have multiple bases, all of which have the same number of elements. This number is the *dimension* of $\mathcal{V}$. When the need arises to work with multiple bases, vector components will be superscripted to identify the basis to which they correspond. For instance, the components of $\vec{\boldsymbol{v}}$ in basis $\mathcal{E}_\alpha$ will be written $\{v^\alpha_{(1)}, v^\alpha_{(2)}, \dots, v^\alpha_{(n)}\}$.

For most practical applications, the scalar field $\mathcal{F}$ is the field of real numbers $\mathbb{R}$, in which case the vector space is called a *real vector space*. In a real vector space, an *inner* or *scalar* product can be defined, which takes two vectors $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in \mathcal{V}$ and produces a scalar $a = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}$. The scalar product allows to introduce additional notions such as the norm of a vector or the angle between two vectors. A vector space with a scalar product operation is called an *inner product space*.

The scalar product must satisfy the following properties:

1) Commutativity:

$$\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}} = \vec{\boldsymbol{w}} \cdot \vec{\boldsymbol{v}} \tag{1.11}$$

2) Distributivity with respect to vector addition:

$$(\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{z}} + \vec{\boldsymbol{w}} \cdot \vec{\boldsymbol{z}} \tag{1.12}$$

3) Linearity with respect to scalar multiplication:

$$(a\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{w}} = a(\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}) \tag{1.13}$$

4) Positive-definiteness:

$$\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}} \geq 0, \text{ with } \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}} = 0 \iff \vec{\boldsymbol{v}} = \vec{\boldsymbol{0}} \tag{1.14}$$

The norm of a vector is defined as:

$$|\vec{\boldsymbol{v}}| = \sqrt{\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{v}}} \tag{1.15}$$

If $|\vec{\boldsymbol{v}}| = 1$, $\vec{\boldsymbol{v}}$ is called a *unit vector*.

The angle between two vectors $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ is defined as:

$$\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}} = \arccos \frac{\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}}{|\vec{\boldsymbol{v}}|\,|\vec{\boldsymbol{w}}|} \tag{1.16}$$

When $\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}} = \vec{\mathbf{0}}$, $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ are said to be *orthogonal*. Hence, $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ are orthogonal if and only if $\left|\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}}\right| = \pi/2$.

A basis is said to be orthogonal if its vectors are mutually orthogonal, that is:

$$\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j = 0, \forall i \neq j \tag{1.17}$$

An orthogonal basis is also *orthonormal* if all its vectors are unit vectors, that is:

$$\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_i = 1, \forall i \in \{1, \dots, n\} \tag{1.18}$$

Taking the scalar product of (1.10) with basis element $\vec{\boldsymbol{e}}_j$ and using properties (1.12) and (1.13):

$$\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{e}}_j = \left(\sum_{i=1}^{n} v_{(i)} \vec{\boldsymbol{e}}_i \right) \cdot \vec{\boldsymbol{e}}_j = \sum_{i=1}^{n} v_{(i)} \left(\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j \right)$$

If $\mathcal{E}$ is an orthonormal basis, (1.17) and (1.18) can be applied to yield:

$$\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{e}}_j = \sum_{i=1}^{n} v_{(i)} \left(\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j \right) = v_{(j)} \tag{1.19}$$

Thus, the j-th component of a vector in an orthonormal basis is simply its scalar product with the j-th basis vector.

Applying (1.10), together with (1.12), (1.13) and (1.19), the essential operations for an inner product space can be expressed in terms of components in an arbitrary orthonormal basis:

a) Vector addition:

$$z_{(i)} = \vec{\boldsymbol{z}} \cdot \vec{\boldsymbol{e}}_i = (\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{e}}_i = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{e}}_i + \vec{\boldsymbol{w}} \cdot \vec{\boldsymbol{e}}_i = v_{(i)} + w_{(i)} \tag{1.20}$$

b) Scalar multiplication:

$$z_{(i)} = \vec{\boldsymbol{z}} \cdot \vec{\boldsymbol{e}}_i = (a\vec{\boldsymbol{v}}) \cdot \vec{\boldsymbol{e}}_i = a(\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{e}}_i) = av_{(i)} \tag{1.21}$$

c) Inner product:

$$\begin{aligned} a &= \left(\sum_{i=1}^{n} v_{(i)} \vec{\boldsymbol{e}}_i \right) \cdot \left(\sum_{j=1}^{n} w_{(j)} \vec{\boldsymbol{e}}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_{(i)} w_{(j)} \left(\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j \right) \\ &= \sum_{k=1}^{n} v_{(k)} w_{(k)} \end{aligned} \tag{1.22}$$

1.1.2 Euclidean Vectors

An *Euclidean vector* is typically defined as a geometric object that has both length and direction. Two-dimensional and three-dimensional Euclidean vectors are often used as abstract representations of physical quantities that are fully defined by their magnitude and direction. Throughout this document, whenever we talk about the Euclidean vector space without any additional qualifiers, we will be referring specifically to the three-dimensional Euclidean vector space.

The essential vector operations introduced in section 1.1.1 can be defined for Euclidean vectors and real numbers as purely geometrical constructions, without the need for additional mathematical formalisms. Hence, Euclidean vectors comprise, in and of themselves, an inner product space.

However, for Euclidean vectors to be useful in practice, we need to connect them to a systematic mathematical description. To this end, we may introduce a basis $\mathcal{E} = \{\vec{\boldsymbol{e}}_1, \vec{\boldsymbol{e}}_2, \vec{\boldsymbol{e}}_3\}$, which enables us to represent any Euclidean vector by a unique set of three components. For such a basis, the orthonormality conditions (1.17) and (1.18) can be written more compactly as:

$$\vec{\boldsymbol{e}}_i \cdot \vec{\boldsymbol{e}}_j = \delta_{ij}, \forall i, j \in \{1, 2, 3\} \tag{1.23}$$

Where δ_{ij} is the *Kronecker delta*, defined as:

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases} \tag{1.24}$$

Expressions (1.20), (1.21) and (1.22), which were obtained for any inner product space using an orthonormal basis, are directly applicable to the Euclidean vector space.

In the Euclidean vector space, an additional *cross product* or *vector product* operation can be constructed. The cross product takes two vectors $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ and produces another vector $\vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}$, which is orthogonal to both $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$. The direction of $\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}$ is given by the right hand rule, and its norm is:

$$|\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}| = |\vec{\boldsymbol{v}}|\,|\vec{\boldsymbol{w}}| \sin\theta_{xy} \tag{1.25}$$

The cross product satisfies the following properties:

$$\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}} = -\vec{\boldsymbol{w}} \times \vec{\boldsymbol{v}} \tag{1.26a}$$

$$(a\vec{\boldsymbol{v}}) \times \vec{\boldsymbol{w}} = a(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) \tag{1.26b}$$

$$(\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) \times \vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} \times \vec{\boldsymbol{z}} + \vec{\boldsymbol{w}} \times \vec{\boldsymbol{z}} \tag{1.26c}$$

$$\vec{\boldsymbol{v}} \times (\vec{\boldsymbol{w}} \times \vec{\boldsymbol{z}}) = (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{z}})\vec{\boldsymbol{w}} - (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}})\vec{\boldsymbol{z}} \tag{1.26d}$$

$$\vec{\boldsymbol{v}} \times (\vec{\boldsymbol{w}} \times \vec{\boldsymbol{z}}) + \vec{\boldsymbol{w}} \times (\vec{\boldsymbol{z}} \times \vec{\boldsymbol{v}}) + \vec{\boldsymbol{z}} \times (\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) = \vec{\mathbf{0}} \tag{1.26e}$$

$$\vec{\boldsymbol{v}} \cdot (\vec{\boldsymbol{w}} \times \vec{\boldsymbol{z}}) = \vec{\boldsymbol{z}} \cdot (\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) = \vec{\boldsymbol{w}} \cdot (\vec{\boldsymbol{z}} \times \vec{\boldsymbol{v}}) \tag{1.26f}$$

The cross product allows us to define the *orientation* or *handedness* of an Euclidean vector basis.

An orthonormal basis is called *right-handed* if its vectors further satisfy:

$$(\vec{\boldsymbol{e}}_1 \times \vec{\boldsymbol{e}}_2) \cdot \vec{\boldsymbol{e}}_3 = 1 \tag{1.27}$$

Conversely, a *left-handed* orthonormal basis satisfies:

$$(\vec{\boldsymbol{e}}_1 \times \vec{\boldsymbol{e}}_2) \cdot \vec{\boldsymbol{e}}_3 = -1 \tag{1.28}$$

Note that, due to the invariance of the triple product under cyclic permutations (property (1.26f)):

$$(\vec{\boldsymbol{e}}_1 \times \vec{\boldsymbol{e}}_2) \cdot \vec{\boldsymbol{e}}_3 = (\vec{\boldsymbol{e}}_2 \times \vec{\boldsymbol{e}}_3) \cdot \vec{\boldsymbol{e}}_1 = (\vec{\boldsymbol{e}}_3 \times \vec{\boldsymbol{e}}_1) \cdot \vec{\boldsymbol{e}}_2 \tag{1.29}$$

The orthonormality and right-handedness conditions (1.23) and (1.27), together with the invariance property (1.29), can be combined as:

$$\left(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j\right) \cdot \vec{\boldsymbol{e}}_k = \epsilon_{ijk} \tag{1.30}$$

Where the n-dimensional *Levi-Civita symbol* is defined as:

$$\epsilon_{i_1 \dots i_n} = \prod_{1 \le j \le k \le n} \operatorname{sgn}(i_k - i_j) \tag{1.31}$$

In particular, the three-dimensional the Levi-Civita ϵ_{ijk} is given by:

$$\epsilon_{ijk} = \begin{cases} +1, & \text{if } (i,j,k) \in \{(1,2,3),(2,3,1),(3,1,2)\}; \\ -1, & \text{if } (i,j,k) \in \{(1,3,2),(2,1,3),(3,2,1)\}; \\ 0, & \text{otherwise} \end{cases} \tag{1.32}$$

For an orthonormal right-handed basis, using (1.10), (1.26c) and (1.26b) the cross product can be expressed as:

$$\vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}} = \left(\sum_{i=1}^{3} v_{(i)} \vec{\boldsymbol{e}}_i\right) \times \left(\sum_{j=1}^{3} w_{(j)} \vec{\boldsymbol{e}}_j\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} v_{(i)} w_{(j)} \vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j \tag{1.33}$$

Its k-th component is found by applying (1.19) and (1.30):

$$\begin{aligned} z_{(k)} = \vec{\boldsymbol{z}} \cdot \vec{\boldsymbol{e}}_k = (\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{e}}_k &= \sum_{i=1}^{3}\sum_{j=1}^{3} v_{(i)} w_{(j)} \left(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j\right) \cdot \vec{\boldsymbol{e}}_k \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} v_{(i)} w_{(j)} \epsilon_{ijk}, \forall k \in \{1,2,3\} \end{aligned} \tag{1.34}$$

Using (1.32), (1.34) can be expanded as:

$$\begin{aligned} z_{(1)} &= v_{(2)} w_{(3)} - v_{(3)} w_{(2)} \\ z_{(2)} &= v_{(3)} w_{(1)} - v_{(1)} w_{(3)} \\ z_{(3)} &= v_{(1)} w_{(2)} - v_{(2)} w_{(1)} \end{aligned} \tag{1.35}$$

The triple product $(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{z}}$ can be computed as follows:

$$\begin{aligned}(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{z}} &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} v_{(i)} w_{(j)} z_{(k)} \left(\vec{\boldsymbol{e}}_i \times \vec{\boldsymbol{e}}_j\right) \cdot \vec{\boldsymbol{e}}_k \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} v_{(i)} w_{(j)} z_{(k)} \epsilon_{ijk}\end{aligned} \tag{1.36}$$

1.2 Matrices

1.2.1 Definitions and Properties

A *matrix* $\boldsymbol{A}$ of size $n \times m$ is a set of nm scalar values $A_{(i,j)}$, with $(i,j) \in \{1,\dots,n\} \times \{1,\dots,m\}$. It is typically represented as a two-dimensional array with n rows and m columns:

$$\boldsymbol{A} = \begin{pmatrix} A_{(1,1)} & A_{(1,2)} & \cdots & A_{(1,m)} \\ A_{(2,1)} & A_{(2,2)} & \cdots & A_{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{(n,1)} & A_{(n,2)} & \cdots & A_{(n,m)} \end{pmatrix} \tag{1.37}$$

$\boldsymbol{A}$ is said to be *square* if $m = n$. If $n = 1$, it is called a *row matrix*, and if $m = 1$, a *column matrix*. Row and column matrices will be written in lowercase.

A *diagonal* matrix $\boldsymbol{A}$ is a square matrix with $A_{(i,j)} = 0, \forall i \neq j$.

An $n \times n$ *identity matrix* is a diagonal matrix with $A_{(i,i)} = 1, \forall i \in \{1,\dots,n\}$. An identity matrix will be written as $\boldsymbol{I}_n$ or, wherever it does not result in ambiguity, simply as $\boldsymbol{I}$.

An $n \times m$ matrix whose entries are all equal to zero is called a *null matrix*. A null matrix will be written as $\mathbf{0}_{n\times m}$ or, wherever it does not result in ambiguity, simply as $\mathbf{0}$.

Given a scalar b and a matrix $\boldsymbol{C}$ of size $n \times m$, their *product* $\boldsymbol{A} = b\boldsymbol{C}$ is a matrix of size $n \times m$ with $A_{(i,j)} = bC_{(i,j)}, \forall (i,j) \in \{1,\dots,n\} \times \{1,\dots,m\}$.

Given two matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ of size $n \times m$, their *sum* is another matrix $\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{C}$ with $A_{(i,j)} = B_{(i,j)} + C_{(i,j)}, \forall (i,j) \in \{1,\dots,n\} \times \{1,\dots,m\}$.

Given two matrices $\boldsymbol{B}$ of sizes $n \times l$ and $l \times m$, their *product* is a matrix $\boldsymbol{A} = \boldsymbol{BC}$ of size $n \times m$ with:

$$A_{(i,j)} = \sum_{k=1}^{k=l} B_{(i,k)} C_{(k,j)}, \forall (i,j) \in \{1,\dots,n\} \times \{1,\dots,m\}.$$

The matrix product satisfies:

$$\boldsymbol{AI} = \boldsymbol{A} \tag{1.38a}$$

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) \tag{1.38b}$$

$$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC} \tag{1.38c}$$

The matrix product is non-commutative in general:

$$\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}$$

Given a matrix $\boldsymbol{A}$ of size $n \times m$, its *transpose* $\boldsymbol{B} = \boldsymbol{A}^T$ is a matrix of size $m \times n$ with $B_{(j,i)} = A_{(i,j)}, \forall (j,i) \in \{1, \dots, m\} \times \{1, \dots, n\}$.

The transpose operation satisfies:

$$\left(\boldsymbol{A}^T\right)^T = \boldsymbol{A} \tag{1.39a}$$

$$\left(\boldsymbol{A} + \boldsymbol{B}\right)^T = \boldsymbol{A}^T + \boldsymbol{B}^T \tag{1.39b}$$

$$\left(\boldsymbol{A}\boldsymbol{B}\right)^T = \boldsymbol{B}^T \boldsymbol{A}^T \tag{1.39c}$$

If $\boldsymbol{A}^T = \boldsymbol{A}$, then $\boldsymbol{A}$ is a *symmetric matrix*. If $\boldsymbol{A}^T = -\boldsymbol{A}$, then $\boldsymbol{A}$ is a *skew-symmetric matrix*. Note that symmetric and skew-symmetric matrices are necessarily square. A skew-symmetric matrix $\boldsymbol{A}$ of size $n \times n$ must have $A_{(i,i)} = 0, \forall i \in \{1, \dots, n\}$.

The trace of a $n \times n$ matrix is defined as:

$$\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} A_{(i,i)} \tag{1.40}$$

The trace satisfies:

$$\operatorname{tr}\left(\boldsymbol{A} + \boldsymbol{B}\right) = \operatorname{tr} \boldsymbol{A} + \operatorname{tr} \boldsymbol{B} \tag{1.41a}$$

$$\operatorname{tr}\left(a\boldsymbol{B}\right) = a \operatorname{tr} \boldsymbol{B} \tag{1.41b}$$

$$\operatorname{tr}\left(\boldsymbol{A}\boldsymbol{B}\right) = \operatorname{tr}\left(\boldsymbol{B}\boldsymbol{A}\right) \tag{1.41c}$$

The determinant of a $n \times n$ matrix $\boldsymbol{A}$ is a scalar defined by either of the following recursive expressions:

$$\det \boldsymbol{A} = \sum_{k=1}^{n} (-1)^{i+k} A_{(i,k)} m_{ik} = \sum_{k=1}^{n} (-1)^{k+j} A_{(k,j)} m_{kj} \tag{1.42}$$

Where m_{ij}, called the (i,j) minor of $\boldsymbol{A}$, is the determinant of the $(n-1) \times (n-1)$ matrix resulting from deleting row i and column j of $\boldsymbol{A}$.

An equivalent definition for the determinant is:

$$\det \boldsymbol{A} = \sum_{i_1=1}^{n} \dots \sum_{i_n=1}^{n} \epsilon_{i_1 \dots i_n} A_{(i_1,1)} \dots A_{(i_n,n)} \tag{1.43}$$

Where the Levi-Civita symbol $\epsilon_{i_1 \dots i_n}$ was defined in (1.31).

For $n = 3$, (1.43) reduces to:

$$\det \boldsymbol{A} = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \epsilon_{ijk} A_{(i,1)} A_{(j,2)} A_{(k,3)} \tag{1.44}$$

If we arrange the components of three vectors $\vec{\boldsymbol{v}}$, $\vec{\boldsymbol{w}}$ and $\vec{\boldsymbol{z}}$ in an orthonormal right-handed basis along the columns of a matrix $\boldsymbol{A}$, so that $A_{(i,1)} = v_{(i)}$, $A_{(j,2)} = w_{(j)}$, $A_{(k,3)} = z_{(k)}$, then from (1.36) and (1.44):

$$(\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}) \cdot \vec{\boldsymbol{z}} = \det \boldsymbol{A} \tag{1.45}$$

The determinant satisfies the following properties:

$$\det \boldsymbol{I} = 1 \tag{1.46a}$$
$$\det \boldsymbol{A}^T = \det \boldsymbol{A} \tag{1.46b}$$
$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \det \boldsymbol{B} \tag{1.46c}$$
$$\det(a\boldsymbol{B}) = a^n \det \boldsymbol{B} \tag{1.46d}$$

Given a $n \times n$ matrix $\boldsymbol{A}$, the scalar λ and the column matrix $\boldsymbol{v}$ are respectively an *eigenvalue* and a *right eigenvector* of $\boldsymbol{A}$ if they satisfy:

$$\boldsymbol{Av} = \lambda \boldsymbol{v} \tag{1.47}$$

Since this condition is equally satisfied by λ and $k\boldsymbol{v}$ for any nonzero scalar k, eigenvectors are usually specified as unit vectors.

Equation (1.47) can also be written as:

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\, \boldsymbol{v} = \boldsymbol{0} \tag{1.48}$$

For (1.48) to have a nontrivial solution, $\boldsymbol{A} - \lambda\boldsymbol{I}$ must be singular, that is:

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0 \tag{1.49}$$

Equation (1.49) is a polynomial equation of degree n in λ. It is known as the *characteristic equation* of $\boldsymbol{A}$. Its solutions are the eigenvalues of $\boldsymbol{A}$, which may be real or complex.

The trace and determinant of $\boldsymbol{A}$ are given in terms of its eigenvalues as:

$$\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \tag{1.50}$$

$$\det \boldsymbol{A} = \prod_{i=1}^{n} \lambda_i \tag{1.51}$$

The exponential of a square matrix $\boldsymbol{A}$ is defined as:

$$\exp(\boldsymbol{A}) = \sum_{i=0}^{\infty} \frac{1}{i!} \boldsymbol{A}^i \tag{1.52}$$

Where, by definition, $\boldsymbol{A}^0 = \boldsymbol{I}$.

The matrix exponential satisfies:

$$\exp\left(\boldsymbol{A}^T\right) = \left(\exp\left(\boldsymbol{A}\right)\right)^T \tag{1.53}$$

An $n \times n$ matrix $\boldsymbol{A}$ is said to be *nonsingular* or *invertible* if there exists an $n \times n$ matrix $\boldsymbol{A}^{-1}$ such that:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_n \tag{1.54}$$

The matrix inverse satisfies:

$$\left(\boldsymbol{A}^{-1}\right)^T = \left(\boldsymbol{A}^T\right)^{-1} \tag{1.55a}$$

$$\left(\boldsymbol{A}\boldsymbol{B}\right)^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \tag{1.55b}$$

A matrix is invertible if and only if its determinant is nonzero.

A matrix is invertible if and only if its columns are linearly independent.

A matrix for which $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$ is called *orthogonal* (or sometimes *orthonormal*).

The set of $n \times n$ orthogonal matrices, together with the matrix product, jointly satisfy the following properties:

a) Closure: The product of two orthogonal matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is also an orthogonal matrix:

$$(\boldsymbol{A}\boldsymbol{B})(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$$

b) Identity: The identity matrix $\boldsymbol{I}_n$, which is the identity element for the matrix product, is also an orthogonal matrix.

c) Associativity: This follows directly from the associativity of the general matrix product.

d) Invertibility: For an orthogonal matrix, the existence of an inverse is guaranteed, since by definition its inverse is equal to its transpose, and the transpose always exists.

These properties characterize the set of $n \times n$ orthogonal matrices as a *group*, known as the *orthogonal group* and denoted by $O(n)$.

If $\boldsymbol{A}$ is an orthogonal matrix, from (1.46c) and (1.46b):

$$(\det \boldsymbol{A})^2 = \det \boldsymbol{A} \det \boldsymbol{A}^T = \det(\boldsymbol{A}\boldsymbol{A}^T) = \det \boldsymbol{I} = 1$$

Hence, an orthogonal matrix $\boldsymbol{A}$ has $\det \boldsymbol{A} = \pm 1$.

The subset of $n \times n$ orthogonal matrices whose determinant is $+1$ are called *proper*. They form themselves a group, called the *special orthogonal group* and denoted by $SO(n)$. Orthogonal matrices whose determinant is -1 are called *improper*.

The derivative of a matrix $\boldsymbol{A}$ with respect to some scalar variable v is another matrix $\boldsymbol{A}'$ defined as:

$$A'_{(i,j)} = \frac{dA_{(i,j)}}{dx}$$

The matrix derivative satisfies:

$$\left(\boldsymbol{A}^T\right)' = \left(\boldsymbol{A}'\right)^T \tag{1.56a}$$

$$(\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{A}'\boldsymbol{B} + \boldsymbol{A}\boldsymbol{B}' \tag{1.56b}$$

1.2.2 Matrix Representation of Vectors

Let $\mathcal{V}$ be an n-dimensional vector space with a basis $\mathcal{E}$. The coordinates of an arbitrary vector $\vec{\boldsymbol{v}} \in \mathcal{V}$ in $\mathcal{E}$ may be arranged in a column matrix $\boldsymbol{v}$ as follows:

$$\boldsymbol{v} = \begin{pmatrix} v_{(1)} \\ dots \\ v_{(n)} \end{pmatrix} = \begin{pmatrix} v_{(1)} & \dots & v_{(n)} \end{pmatrix}^T$$

If $\mathcal{V}$ is an inner product space, and $\mathcal{E}$ is an orthonormal basis, then the component-wise expressions (1.20), (1.21) and (1.22) apply, and they can be written in terms of matrix operations as:

a) Vector addition ($\vec{\boldsymbol{z}} = \vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}$):

$$\boldsymbol{z} = \boldsymbol{v} + \boldsymbol{w} \tag{1.57}$$

b) Scalar multiplication ($\vec{\boldsymbol{z}} = a\vec{\boldsymbol{v}}$):

$$\boldsymbol{z} = a\boldsymbol{v} \tag{1.58}$$

c) Inner product ($a = \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}}$):

$$a = \boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^T\boldsymbol{w} \tag{1.59}$$

If $\mathcal{V}$ is the three-dimensional Euclidean space, and the orthonormal basis $\mathcal{E}$ is also right-handed, the component-wise expressions (1.35) for the cross product apply, and they can be written in matrix form as:

$$\boldsymbol{z} = \boldsymbol{v} \times \boldsymbol{w} = [\boldsymbol{v}^\times]\, \boldsymbol{w} \tag{1.60}$$

Where the *cross product operator* $[^\times]$ takes column matrix $\boldsymbol{v}$ as an input and outputs the skew-symmetric *cross product matrix*:

$$[\boldsymbol{v}^\times] = \begin{pmatrix} 0 & -v_{(3)} & v_{(2)} \\ v_{(3)} & 0 & -v_{(1)} \\ -v_{(2)} & v_{(1)} & 0 \end{pmatrix} \tag{1.61}$$

The cross product matrix can be written element-wise in terms of the Levi-Civita symbol:

$$[v^\times]_{(i,j)} = -\sum_{k=1}^{3} v_{(k)}\epsilon_{ijk} \tag{1.62}$$

The cross product matrix satisfies the following properties:

$$[a\boldsymbol{v}^{\times}] = a\left[\boldsymbol{v}^{\times}\right] \tag{1.63a}$$

$$\left[\boldsymbol{v}^{\times}\right]\boldsymbol{v} = \boldsymbol{0} \tag{1.63b}$$

$$\left[\boldsymbol{v}^{\times}\right]\boldsymbol{w} = -\left[\boldsymbol{w}^{\times}\right]\boldsymbol{v} \tag{1.63c}$$

$$\left[\boldsymbol{v}^{\times}\right]^{T} = -\left[\boldsymbol{v}^{\times}\right] \tag{1.63d}$$

$$\left[\boldsymbol{v}^{\times}\right]\left[\boldsymbol{w}^{\times}\right] = -\left(\boldsymbol{v}^{T}\boldsymbol{w}\right)\boldsymbol{I}_3 + \boldsymbol{w}\boldsymbol{v}^{T} \tag{1.63e}$$

$$\left[\boldsymbol{v}^{\times}\right]\left[\boldsymbol{w}^{\times}\right] - \left[\boldsymbol{w}^{\times}\right]\left[\boldsymbol{v}^{\times}\right] = \left[\left(\boldsymbol{v}\times\boldsymbol{w}\right)^{\times}\right] \tag{1.63f}$$

$$\left[\boldsymbol{v}^{\times}\right]^{3} = -\left|\boldsymbol{v}\right|^{2}\left[\boldsymbol{v}^{\times}\right] \tag{1.63g}$$

Applying (1.63g) recursively, it is easy to verify that:

$$\left[\boldsymbol{v}^{\times}\right]^{2k} = -(-1)^{k}\left|\boldsymbol{v}\right|^{2(k-1)}\left[\boldsymbol{v}^{\times}\right]^{2}, \forall k = 1, 2, \ldots \tag{1.64a}$$

$$\left[\boldsymbol{v}^{\times}\right]^{2k+1} = (-1)^{k}\left|\boldsymbol{v}\right|^{2k}\left[\boldsymbol{v}^{\times}\right], \forall k = 0, 1, \ldots \tag{1.64b}$$

The cross product matrix also satisfies the following identity, where $\boldsymbol{A}$ is any nonsingular 3×3 matrix:

$$\boldsymbol{A}\left[\boldsymbol{v}^{\times}\right]\boldsymbol{A}^{T} = (\det \boldsymbol{A})\left[\left(\left(\boldsymbol{A}^{T}\right)^{-1}\boldsymbol{v}\right)^{\times}\right] \tag{1.65}$$

Finally, the linear independence condition (1.9) translates directly into the matrix representation: the set of vectors $\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \ldots, \vec{\boldsymbol{v}}_n\}$ is linearly independent if and only if the following equation is satisfied only by $a_i = 0$ for $i = 1, 2, \ldots, n$:

$$\sum_{i=1}^{n} a_i \boldsymbol{v}_i = \boldsymbol{0} \tag{1.66}$$

Whenever we need to emphasize that $\boldsymbol{v}$ holds the coordinates of $\vec{\boldsymbol{v}}$ in a specific basis $\mathcal{E}_{\alpha}$, we shall superscript it as $\boldsymbol{v}^{\alpha}$, following the same convention as for individual coordinates.

1.3 Linear Transformations

1.3.1 Definitions and Properties

Given two vector spaces $\mathcal{V}$ and $\mathcal{W}$ over a scalar field $\mathcal{F}$, a function $f : \mathcal{V} \to \mathcal{W}$ is said to be a *linear map* or a *linear transformation* if, for any two vectors $\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2 \in \mathcal{V}$ and any scalar $a \in \mathcal{F}$, the following two conditions are satisfied:

a) Additivity:

$$f\left(\vec{\boldsymbol{v}}_1 + \vec{\boldsymbol{v}}_2\right) = f\left(\vec{\boldsymbol{v}}_1\right) + f\left(\vec{\boldsymbol{v}}_2\right) \tag{1.67}$$

b) Homogeinity:

$$f\left(a\vec{\boldsymbol{v}}_1\right) = af\left(\vec{\boldsymbol{v}}_1\right) \tag{1.68}$$

Let $\mathcal{V}$ and $\mathcal{W}$ be two vector spaces with respective dimensions $n_{\mathcal{V}}$ and $n_{\mathcal{W}}$ and bases $\mathcal{E}_{\mathcal{V}}$ and $\mathcal{E}_{\mathcal{W}}$. Let $f : \mathcal{V} \to \mathcal{W}$ be a linear transformation that maps basis vectors $\vec{\boldsymbol{e}}_{\mathcal{V}i}$ into another set of vectors $\vec{\boldsymbol{\gamma}}_i \in \mathcal{W}$, that is, $\vec{\boldsymbol{\gamma}}_i = f\left(\vec{\boldsymbol{e}}_{\mathcal{V}i}\right), \forall i \in \{1, \dots, n_{\mathcal{V}}\}$.

Each $\vec{\boldsymbol{\gamma}}_i \in \mathcal{W}$ can be expressed in terms of its coordinates in basis $\mathcal{E}_{\mathcal{W}}$ as:

$$\vec{\boldsymbol{\gamma}}_i = \sum_{j=1}^{n_{\mathcal{W}}} \gamma_{i(j)} \vec{\boldsymbol{e}}_{\mathcal{W}j} \tag{1.69}$$

Now, consider an arbitrary vector $\vec{\boldsymbol{v}} \in \mathcal{V}$ and its transformed counterpart $\vec{\boldsymbol{w}} = f\left(\vec{\boldsymbol{v}}\right) \in \mathcal{W}$, which we can express respectively in terms of their $\mathcal{E}_{\mathcal{V}}$ and $\mathcal{E}_{\mathcal{W}}$ coordinates as:

$$\vec{\boldsymbol{v}} = \sum_{i=1}^{n_{\mathcal{V}}} v_{(i)} \vec{\boldsymbol{e}}_{\mathcal{V}i} \tag{1.70}$$

$$\vec{\boldsymbol{w}} = \sum_{j=1}^{n_{\mathcal{W}}} w_{(j)} \vec{\boldsymbol{e}}_{\mathcal{W}j} \tag{1.71}$$

The components of $\vec{\boldsymbol{w}}$ can be found from (1.70) and (1.71) by applying the linearity properties (1.67) and (1.68):

$$\begin{aligned}
\vec{\boldsymbol{w}} = f\left(\vec{\boldsymbol{v}}\right) = f\left(\sum_{i=1}^{n_{\mathcal{V}}} v_{(i)} \vec{\boldsymbol{e}}_{\mathcal{V}i}\right) &= \sum_{i=1}^{n_{\mathcal{V}}} v_{(i)} f\left(\vec{\boldsymbol{e}}_{\mathcal{V}i}\right) = \sum_{i=1}^{n_{\mathcal{V}}} v_{(i)} \vec{\boldsymbol{\gamma}}_i \\
= \sum_{i=1}^{n_{\mathcal{V}}} v_{(i)} \left(\sum_{j=1}^{n_{\mathcal{W}}} \gamma_{i(j)} \vec{\boldsymbol{e}}_{\mathcal{W}j}\right) &= \sum_{i=1}^{n_{\mathcal{V}}} \sum_{j=1}^{n_{\mathcal{W}}} \gamma_{i(j)} v_{(i)} \vec{\boldsymbol{e}}_{\mathcal{W}j} \\
= \sum_{j=1}^{n_{\mathcal{W}}} \sum_{i=1}^{n_{\mathcal{V}}} \gamma_{i(j)} v_{(i)} \vec{\boldsymbol{e}}_{\mathcal{W}j} &= \sum_{j=1}^{n_{\mathcal{W}}} \left(\sum_{i=1}^{n_{\mathcal{V}}} \gamma_{i(j)} v_{(i)}\right) \vec{\boldsymbol{e}}_{\mathcal{W}j}
\end{aligned} \tag{1.72}$$

Equating (1.71) and (1.72) gives:

$$w_{(j)} = \sum_{i=1}^{n_{\mathcal{V}}} \gamma_{j(i)} v_{(j)}, \forall j \in \{1, \dots, n_{\mathcal{W}}\} \tag{1.73}$$

This can be written in matrix form as:

$$\boldsymbol{w} = \boldsymbol{\Gamma} \boldsymbol{v} \tag{1.74}$$

Where:

- $\boldsymbol{v}$ is the $n_{\mathcal{V}} \times 1$ column matrix containing the coordinates of $\vec{\boldsymbol{v}}$ in $\mathcal{E}_{\mathcal{V}}$
- $\boldsymbol{w}$ is the $n_{\mathcal{W}} \times 1$ column matrix containing the coordinates of $\vec{\boldsymbol{w}}$ in $\mathcal{E}_{\mathcal{W}}$
- $\boldsymbol{\Gamma}$ is a $n_{\mathcal{W}} \times n_{\mathcal{V}}$ transformation matrix, defined by:

$$\mathit{\Gamma}_{(i,j)} = \gamma_{j(i)} \tag{1.75}$$

Given the coordinates of an arbitrary vector $\vec{\boldsymbol{v}}$ in $\mathcal{E}_{\mathcal{V}}$, transformation matrix $\boldsymbol{\Gamma}$ allows us to compute the coordinates of $\vec{\boldsymbol{w}} = f(\vec{\boldsymbol{v}})$ in basis $\mathcal{E}_{\mathcal{W}}$. Therefore, the linear transformation f is fully defined by $\boldsymbol{\Gamma}$. Conversely, any $n_{\mathcal{W}} \times n_{\mathcal{V}}$ matrix can be interpreted as the representation of a linear transformation between two vector spaces of dimensions $n_{\mathcal{V}}$ and $n_{\mathcal{W}}$.

From (1.75) we can see that the j-th column of $\boldsymbol{\Gamma}$ holds the coordinates of the transformed basis vector $\vec{\boldsymbol{\gamma}}_j = f\left(\vec{\boldsymbol{e}}_{\mathcal{V}j}\right)$ in basis $\mathcal{E}_{\mathcal{W}}$:

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\gamma}_1 & \dots & \boldsymbol{\gamma}_{n_{\mathcal{V}}} \end{pmatrix} \tag{1.76}$$

Note that $\boldsymbol{\Gamma}$ clearly depends on the choice of bases $\mathcal{E}_{\mathcal{V}}$ and $\mathcal{E}_{\mathcal{W}}$, and this is not made explicit by the notation. However, since so far we have defined only one basis for each vector space, no confusion should arise from this fact.

Now let $\mathcal{Z}$ be a third vector space with dimension $n_{\mathcal{Z}}$ and basis $\mathcal{E}_{\mathcal{Z}}$, and let $g : \mathcal{W} \to \mathcal{Z}$ be another linear transformation that maps basis vectors $\vec{\boldsymbol{e}}_{\mathcal{W}j}$ into a set of vectors $\vec{\boldsymbol{\sigma}}_j \in \mathcal{Z}$, that is, $\vec{\boldsymbol{\sigma}}_j = g\left(\vec{\boldsymbol{e}}_{\mathcal{W}j}\right), \forall j \in \{1, \dots, n_{\mathcal{W}}\}$. This new transformation will be described by a $n_{\mathcal{Z}} \times n_{\mathcal{W}}$ matrix $\boldsymbol{\Sigma}$.

Applying f and g consecutively to an arbitrary vector $\vec{\boldsymbol{v}} \in \mathcal{V}$ yields another vector $\vec{\boldsymbol{z}} = g(\vec{\boldsymbol{w}}) = g(f(\vec{\boldsymbol{v}})) = g \circ f(\vec{\boldsymbol{v}}) \in \mathcal{Z}$ whose coordinates in basis $\mathcal{E}_{\mathcal{Z}}$ are given by:

$$\boldsymbol{z} = \boldsymbol{\Sigma}\boldsymbol{w} = \boldsymbol{\Sigma}\boldsymbol{\Gamma}\boldsymbol{v} \tag{1.77}$$

It follows that the composition of linear transformations $g \circ f : \mathcal{V} \to \mathcal{Z}$ is another linear transformation with transformation matrix $\boldsymbol{\Sigma}\boldsymbol{\Gamma}$. Thus, the composition of linear transformations corresponds to matrix multiplication. This implies that it is generally non-commutative.

A linear transformation $f : \mathcal{V} \to \mathcal{V}$, that is, one which maps a vector space $\mathcal{V}$ onto itself, is called an *endomorphism.* The transformation matrix for an endomorphism is necessarily square. Conversely, any $n \times n$ matrix can be interpreted as an endomorphism within a vector space of dimension n.

Note that, for an endomorphism, both $\boldsymbol{v}$ and $\boldsymbol{w}$ in (1.74), as well as the columns $\boldsymbol{\gamma}_i$ in (1.76), represent coordinates in the same basis (whichever we have chosen for $\mathcal{V}$).

A column matrix $\boldsymbol{v}$ is a right eigenvector of a square matrix $\boldsymbol{\Gamma}$ with eigenvalue λ if it satisfies:

$$\boldsymbol{\Gamma}\boldsymbol{v} = \lambda\boldsymbol{v} \tag{1.78}$$

Thus, if we interpret $\boldsymbol{\Gamma}$ as an endomorphism $f : \mathcal{V} \to \mathcal{V}$, then $\boldsymbol{v}$ represents the coordinates of a vector $\vec{\boldsymbol{v}} \in \mathcal{V}$ along which the effect of f is simply a scaling by a factor λ.

1.3.2 Invertible Transformations and Basis Changes

Let $\mathcal{V}$ be a vector space of dimension n and let $\mathcal{E}_\alpha$ denote our basis for $\mathcal{V}$. Consider an endomorphism $f : \mathcal{V} \to \mathcal{V}$, defined by a matrix $\boldsymbol{\Gamma}$, that transforms the basis vectors $\vec{\boldsymbol{e}}_{\alpha i}$ into a set of *linearly independent* vectors $\vec{\boldsymbol{\gamma}}_i \in \mathcal{V}$.

Since $\vec{\boldsymbol{\gamma}}_i$ are linearly independent, so are their respective coordinates $\boldsymbol{\gamma}_i$, which make up the columns of $\boldsymbol{\Gamma}$. Thus, $\boldsymbol{\Gamma}$ is nonsingular. This means we can premultiply (1.74) by $(\boldsymbol{\Gamma})^{-1}$ to recover the $\mathcal{E}_\alpha$ coordinates of the original vector $\vec{\boldsymbol{v}}$ from those of its transformed counterpart $\vec{\boldsymbol{w}}$:

$$\boldsymbol{v} = (\boldsymbol{\Gamma})^{-1}\boldsymbol{w} \tag{1.79}$$

Therefore, if the above holds, the inverse endomorphism $f^{-1} : \mathcal{V} \to \mathcal{V}$ exists, and its transformation matrix is $(\boldsymbol{\Gamma})^{-1}$.

Being a set of n linearly independent vectors, the transformed vectors $\vec{\boldsymbol{\gamma}}_i$ can be used to construct another basis for $\mathcal{V}$. Let this basis be $\mathcal{E}_\beta$, with basis vectors $\vec{\boldsymbol{e}}_{\beta i} = \vec{\boldsymbol{\gamma}}_i = f(\vec{\boldsymbol{e}}_{\alpha i})$.

Since we are now considering two bases for the same vector space, we should make our notation more precise to avoid confusion. From now on, we shall superscript both scalar components and column matrices to indicate whether they correspond to basis $\mathcal{E}_\alpha$ or basis $\mathcal{E}_\beta$. Furthermore, to emphasize the fact that f maps the original basis $\mathcal{E}_\alpha$ into $\mathcal{E}_\beta$, we shall denote its transformation matrix $\boldsymbol{\Gamma}^\alpha_\beta$.

The original and transformed vectors $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}} = f(\vec{\boldsymbol{v}})$ are expressed in terms of their $\mathcal{E}_\alpha$ components as:

$$\vec{\boldsymbol{v}} = \sum_{i=1}^{n} v^\alpha_{(i)} \vec{\boldsymbol{e}}_{\alpha i} \tag{1.80}$$

$$\vec{\boldsymbol{w}} = \sum_{i=1}^{n} w^\alpha_{(i)} \vec{\boldsymbol{e}}_{\alpha i} \tag{1.81}$$

The same goes for the $\mathcal{E}_\beta$ basis vectors:

$$\vec{\boldsymbol{e}}_{\beta j} = \sum_{i=1}^{n} e^\alpha_{\beta j(i)} \vec{\boldsymbol{e}}_{\alpha i} \tag{1.82}$$

With these notational conventions, (1.75), (1.76) and (1.74) become:

$$\Gamma^\alpha_{\beta(i,j)} = e^\alpha_{\beta j(i)} \tag{1.83}$$

$$\boldsymbol{\Gamma}^\alpha_\beta = \begin{pmatrix} \boldsymbol{e}^\alpha_{\beta 1} & \dots & \boldsymbol{e}^\alpha_{\beta n} \end{pmatrix} \tag{1.84}$$

$$\boldsymbol{w}^\alpha = \boldsymbol{\Gamma}^\alpha_\beta \boldsymbol{v}^\alpha \tag{1.85}$$

Because the inverse endomorphism f^{-1} maps $\mathcal{E}_\beta$ back into $\mathcal{E}_\alpha$, we shall denote its matrix $\left(\boldsymbol{\Gamma}^\alpha_\beta\right)^{-1}$ by $\boldsymbol{\Gamma}^\beta_\alpha$. Then, by analogy with (1.83), (1.84) and (1.85)

$$\Gamma^\beta_{\alpha(i,j)} = e^\beta_{\alpha j(i)} \tag{1.86}$$

$$\boldsymbol{\Gamma}^\beta_\alpha = \begin{pmatrix} \boldsymbol{e}^\beta_{\alpha 1} & \dots & \boldsymbol{e}^\beta_{\alpha n} \end{pmatrix} \tag{1.87}$$

$$\boldsymbol{v}^\alpha = \left(\boldsymbol{\Gamma}^\alpha_\beta\right)^{-1} \boldsymbol{w}^\alpha = \boldsymbol{\Gamma}^\beta_\alpha \boldsymbol{w}^\alpha \tag{1.88}$$

Now that we can express any vector in both bases $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$, we may ask ourselves about the connection between the $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ coordinates of an arbitrary vector $\vec{\boldsymbol{x}}$. Applying (1.80) to $\vec{\boldsymbol{x}}$ for both $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$:

$$\vec{\boldsymbol{x}} = \sum_{j=1}^{n} x^{\alpha}_{(j)} \vec{\boldsymbol{e}}_{\alpha j} \tag{1.89}$$

$$\vec{\boldsymbol{x}} = \sum_{j=1}^{n} x^{\beta}_{(j)} \vec{\boldsymbol{e}}_{\beta j} \tag{1.90}$$

Substituting (1.82) into (1.90):

$$\begin{aligned} \vec{\boldsymbol{x}} &= \sum_{j=1}^{n} x^{\beta}_{(j)} \left(\sum_{i=1}^{n} e^{\alpha}_{\beta j(i)} \vec{\boldsymbol{e}}_{\alpha j} \right) = \sum_{j=1}^{n} \sum_{i=1}^{n} x^{\beta}_{(j)} e^{\alpha}_{\beta j(i)} \vec{\boldsymbol{e}}_{\alpha i} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x^{\beta}_{(j)} e^{\alpha}_{\beta j(i)} \vec{\boldsymbol{e}}_{\alpha i} = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} e^{\alpha}_{\beta j(i)} x^{\beta}_{(j)} \right) \vec{\boldsymbol{e}}_{\alpha i} \end{aligned} \tag{1.91}$$

By comparison with (1.89) and (1.83) we see that:

$$x^{\alpha}_{(i)} = \sum_{j=1}^{3} e^{\alpha}_{\beta j(i)} x^{\beta}_{(j)} = \sum_{j=1}^{3} \mathit{\Gamma}^{\alpha}_{\beta(i,j)} x^{\beta}_{(j)}$$

Which can be writen in matrix form as:

$$\boldsymbol{x}^{\alpha} = \boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta} \boldsymbol{x}^{\beta} \tag{1.92}$$

And, reciprocally:

$$\boldsymbol{x}^{\beta} = \left(\boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta}\right)^{-1} \boldsymbol{x}^{\alpha} = \boldsymbol{\mathit{\Gamma}}^{\beta}_{\alpha} \boldsymbol{x}^{\alpha} \tag{1.93}$$

An important insight follows from the above results: for an invertible endomorphism f that maps basis $\mathcal{E}_\alpha$ into another basis $\mathcal{E}_\beta$, the transformation matrix $\boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta}$, as defined by (1.83), has a dual interpretation:

- When used to relate the $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ components of the same vector $\vec{\boldsymbol{x}}$, as in (1.92), $\boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta}$ is said to represent a *passive* or *alias* (Latin for "also known as") transformation.
- When used to relate the components of a vector $\vec{\boldsymbol{v}}$ and its transformed counterpart $\vec{\boldsymbol{w}} = f(\vec{\boldsymbol{v}})$ in the same basis, as in (1.85), $\boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta}$ is said to represent an *active* or *alibi* (Latin for "elsewhere") transformation.

Since (1.93) holds for any vector, we can write:

$$\boldsymbol{w}^{\beta} = \left(\boldsymbol{\mathit{\Gamma}}^{\alpha}_{\beta}\right)^{-1} \boldsymbol{w}^{\alpha} \tag{1.94}$$

Inserting (1.85) into (1.94):

$$\boldsymbol{w}^{\beta} = \left(\boldsymbol{\Gamma}^{\alpha}_{\beta}\right)^{-1} \boldsymbol{w}^{\alpha} = \left(\boldsymbol{\Gamma}^{\alpha}_{\beta}\right)^{-1} \boldsymbol{\Gamma}^{\alpha}_{\beta} \boldsymbol{v}^{\alpha} = \boldsymbol{v}^{\alpha} \tag{1.95}$$

Equation (1.95) shows that the coordinates of the transformed vector in the transformed basis equal those of the original vector in the original basis, as one would intuitively expect.

Substituting (1.92) into (1.95):

$$\boldsymbol{w}^{\beta} = \boldsymbol{\Gamma}^{\alpha}_{\beta} \boldsymbol{v}^{\beta} \tag{1.96}$$

Thus, $\boldsymbol{\Gamma}^{\alpha}_{\beta}$ relates not only the $\mathcal{E}_{\alpha}$ components of $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$, as seen in (1.85), but also their $\mathcal{E}_{\beta}$ components.

Finally, consider another invertible endomorphism $g : \mathcal{V} \to \mathcal{V}$, which maps $\mathcal{E}_{\beta}$ into a third basis $\mathcal{E}_{\delta}$, and whose transformation matrix can be therefore denoted by $\boldsymbol{\Gamma}^{\beta}_{\delta}$.

The composition of invertible endomorphisms $g \circ f : \mathcal{V} \to \mathcal{V}$ is also an invertible endomorphism that maps basis $\mathcal{E}_{\alpha}$ directly into basis $\mathcal{E}_{\delta}$. As a particular case of the general composition of linear transformations, the transformation matrix $\boldsymbol{\Gamma}^{\alpha}_{\delta}$ for $g \circ f$ is found through matrix multiplication:

$$\boldsymbol{\Gamma}^{\alpha}_{\delta} = \boldsymbol{\Gamma}^{\alpha}_{\beta} \boldsymbol{\Gamma}^{\beta}_{\delta}$$

Chapter 2

Attitude Representation

2.1 Fundamentals

2.1.1 Proper Rotations in Euclidean Space

Let $\mathcal{V}$ be an inner product space of dimension n and let $\mathcal{E}_\alpha$ be an *orthonormal* basis for $\mathcal{V}$. Let $f : \mathcal{V} \to \mathcal{V}$ be an invertible endomorphism, defined by a non-singular transformation matrix $\boldsymbol{R}^\alpha_\beta$, that maps $\mathcal{E}_\alpha$ into *another* orthonormal basis $\mathcal{E}_\beta$.

Since $\mathcal{E}_\beta$ is orthonormal, we can apply (1.19) in (1.83) and (1.86) to yield:

$$R^\alpha_{\beta(i,j)} = e^\alpha_{\beta j(i)} = \vec{\boldsymbol{e}}_{\alpha i} \cdot \vec{\boldsymbol{e}}_{\beta j} \tag{2.1a}$$

$$R^\beta_{\alpha(i,j)} = e^\beta_{\alpha j(i)} = \vec{\boldsymbol{e}}_{\beta i} \cdot \vec{\boldsymbol{e}}_{\alpha j} = \vec{\boldsymbol{e}}_{\alpha j} \cdot \vec{\boldsymbol{e}}_{\beta i} \tag{2.1b}$$

Which means:

$$\boldsymbol{R}^\beta_\alpha = \left(\boldsymbol{R}^\alpha_\beta\right)^T \tag{2.2}$$

Since $\boldsymbol{R}^\beta_\alpha = \left(\boldsymbol{R}^\alpha_\beta\right)^{-1}$, we have that:

$$\left(\boldsymbol{R}^\alpha_\beta\right)^{-1} = \left(\boldsymbol{R}^\alpha_\beta\right)^T \tag{2.3}$$

Thus, $\boldsymbol{R}^\alpha_\beta$ is an orthogonal matrix, that is, $\boldsymbol{R}^\alpha_\beta \in O(n)$

Now, let $\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2 \in \mathcal{V}$ be two arbitrary vectors, and let $\vec{\boldsymbol{w}}_1 = f(\vec{\boldsymbol{v}}_1)$ and $\vec{\boldsymbol{w}}_2 = f(\vec{\boldsymbol{v}}_2)$, so that:

$$\boldsymbol{w}^\alpha_1 = \boldsymbol{R}^\alpha_\beta \boldsymbol{v}^\alpha_1$$
$$\boldsymbol{w}^\alpha_2 = \boldsymbol{R}^\alpha_\beta \boldsymbol{v}^\alpha_2$$

Applying (1.59) and (1.39c), their inner product can be written as:

$$\vec{\boldsymbol{w}}_1 \cdot \vec{\boldsymbol{w}}_2 = (\boldsymbol{w}^\alpha_1)^T \boldsymbol{w}^\alpha_2 = \left(\boldsymbol{R}^\alpha_\beta \boldsymbol{v}^\alpha_1\right)^T \boldsymbol{R}^\alpha_\beta \boldsymbol{v}^\alpha_2 = (\boldsymbol{v}^\alpha_1)^T \left(\boldsymbol{R}^\alpha_\beta\right)^T \boldsymbol{R}^\alpha_\beta \boldsymbol{v}^\alpha_2 = (\boldsymbol{v}^\alpha_1)^T \boldsymbol{v}^\alpha_2 = \vec{\boldsymbol{v}}_1 \cdot \vec{\boldsymbol{v}}_2$$

This shows that the inner product remains invariant under f. Since distances and angles are defined from the inner product, they are both preserved. It follows that an invertible endomorphism that maps an orthonormal basis into another, and whose matrix is therefore orthogonal, preserves distances and angles. Such an endomorphism is called a *rotation*.

Using (1.16), we can also write (2.1a) and (2.1b) as:

$$\begin{aligned} R^{\alpha}_{\beta(i,j)} &= \cos\theta_{\vec{\boldsymbol{e}}_{\alpha i},\vec{\boldsymbol{e}}_{\beta j}} \\ R^{\beta}_{\alpha(i,j)} &= \cos\theta_{\vec{\boldsymbol{e}}_{\beta i},\vec{\boldsymbol{e}}_{\alpha j}} \end{aligned}$$

Thus, the elements of a rotation matrix are the cosines of the angles between the original and transformed basis vectors. For this reason, a rotation matrix is sometimes known as a *direction cosine matrix*.

The composition of rotations is a particular case of the composition of linear transformations, and therefore it is realized through matrix multiplication:

$$\boldsymbol{R}^{\alpha}_{\beta} = \boldsymbol{R}^{\alpha}_{\delta}\boldsymbol{R}^{\delta}_{\beta} \tag{2.4}$$

It follows from (2.4) and the closure property of $O(n)$ that the composition of rotations is also a rotation.

We now focus in the particular case in which $\mathcal{V}$ is the three-dimensional Euclidean space and $\boldsymbol{R}^{\alpha}_{\beta}$ is a proper orthogonal matrix, that is, $\boldsymbol{R}^{\alpha}_{\beta} \in SO(3)$.

Under these assumptions, orientation, as defined by (1.27) and (1.28), is preserved by the rotation represented by $\boldsymbol{R}^{\alpha}_{\beta}$, which means that the transformed basis $\mathcal{E}_{\beta}$ has the same orientation as $\mathcal{E}_{\alpha}$.

To see why, let us return to (2.4). Noting that $\det \boldsymbol{R}^{\alpha}_{\beta} = 1$ and using properties (1.46b) and (1.46c), we can write:

$$\begin{aligned} 1 = \det \boldsymbol{R}^{\alpha}_{\beta} &= \det\left(\boldsymbol{R}^{\alpha}_{\delta}\boldsymbol{R}^{\delta}_{\beta}\right) = \det\left(\left(\boldsymbol{R}^{\delta}_{\alpha}\right)^{T}\boldsymbol{R}^{\delta}_{\beta}\right) \\ &= \det\left(\boldsymbol{R}^{\delta}_{\alpha}\right)^{T} \det \boldsymbol{R}^{\delta}_{\beta} = \det \boldsymbol{R}^{\delta}_{\alpha} \det \boldsymbol{R}^{\delta}_{\beta} \end{aligned} \tag{2.5}$$

From (1.84) we know that $\boldsymbol{R}^{\delta}_{\alpha}$ and $\boldsymbol{R}^{\delta}_{\beta}$ have the following structure:

$$\begin{aligned} \boldsymbol{R}^{\delta}_{\alpha} &= \begin{pmatrix} \boldsymbol{e}^{\delta}_{\alpha 1} & \boldsymbol{e}^{\delta}_{\alpha 2} & \boldsymbol{e}^{\delta}_{\alpha 3} \end{pmatrix} \\ \boldsymbol{R}^{\delta}_{\beta} &= \begin{pmatrix} \boldsymbol{e}^{\delta}_{\beta 1} & \boldsymbol{e}^{\delta}_{\beta 2} & \boldsymbol{e}^{\delta}_{\beta 3} \end{pmatrix} \end{aligned}$$

If we choose $\mathcal{E}_{\delta}$ to be orthonormal and right-handed, we can apply (1.45) to write:

$$\det \boldsymbol{R}^{\delta}_{\alpha} = \left(\vec{\boldsymbol{e}}_{\alpha 1} \times \vec{\boldsymbol{e}}_{\alpha 2}\right) \cdot \vec{\boldsymbol{e}}_{\alpha 3} \tag{2.6a}$$

$$\det \boldsymbol{R}^{\delta}_{\beta} = \left(\vec{\boldsymbol{e}}_{\beta 1} \times \vec{\boldsymbol{e}}_{\beta 2}\right) \cdot \vec{\boldsymbol{e}}_{\beta 3} \tag{2.6b}$$

Inserting (2.6a) and (2.6b) in (2.5) yields:

$$\left(\left(\vec{\boldsymbol{e}}_{\alpha 1} \times \vec{\boldsymbol{e}}_{\alpha 2}\right) \cdot \vec{\boldsymbol{e}}_{\alpha 3}\right)\left(\left(\vec{\boldsymbol{e}}_{\beta 1} \times \vec{\boldsymbol{e}}_{\beta 2}\right) \cdot \vec{\boldsymbol{e}}_{\beta 3}\right) = 1$$

And, for the above to hold, $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ must have the same orientation, as we intended to show.

A rotation in Euclidean space whose matrix is proper orthogonal, and therefore preserves orientation, is called a *proper rotation*. From the closure property of $SO(3)$, it follows that the composition of proper rotations yields a proper rotation.

When a proper rotation in Euclidean space transforming basis $\mathcal{E}_\alpha$ into $\mathcal{E}_\beta$ is interpreted in its passive sense, it is said to represent the *attitude* of $\mathcal{E}_\beta$ with respect to $\mathcal{E}_\alpha$. The corresponding basis changes (1.92) and (1.93) become:

$$\boldsymbol{x}^\alpha = \boldsymbol{R}^\alpha_\beta \boldsymbol{x}^\beta \tag{2.7}$$

$$\boldsymbol{x}^\beta = \left(\boldsymbol{R}^\alpha_\beta\right)^T \boldsymbol{x}^\alpha \tag{2.8}$$

All the rotations we will be dealing with in the remaining sections are proper rotations in Euclidean space.

2.1.2 Euler's Rotation Theorem

Euler's Rotation Theorem asserts that, for every proper rotation in three-dimensional Euclidean space, there is a direction which remains invariant under it. Such direction is called the *axis of rotation*. As any direction in space, the axis of rotation has two degrees of freedom. It may be defined by means of a unit vector $\vec{\boldsymbol{n}}$, whose three components are subject to a unit norm constraint.

Additionally specifying a rotation angle $\vartheta \in [-\pi, \pi)$ fully defines the rotation. The rotation sense is given by the right-hand rule: with $\vec{\boldsymbol{n}}$ pointing towards the viewer, $\vartheta > 0$ yields a counter-clockwise rotation.

It follows from the above that a three-dimensional rotation has three degrees of freedom. Therefore, the 9 elements of a rotation matrix $\boldsymbol{R}^\alpha_\beta$ cannot possibly be independent, but rather they must satisfy some set of 6 scalar constraints. Indeed, these constraints arise from the fact that $\boldsymbol{R}^\alpha_\beta$ is orthogonal: since its columns $\boldsymbol{e}^\alpha_{\beta j}$ represent the components of an orthonormal basis, they are related by the orthonormality conditions (1.23). Hence, the rotation matrix is a redundant or *non-minimal* attitude descriptor.

An equivalent algebraic statement of Euler's theorem is: for any rotation matrix $\boldsymbol{R}^\alpha_\beta$, there exists a unit vector $\vec{\boldsymbol{n}}_{\alpha\beta}$ such that $\boldsymbol{n}^\alpha_{\alpha\beta}$ is an eigenvector of $\boldsymbol{R}^\alpha_\beta$ with eigenvalue $\lambda = 1$:

$$\boldsymbol{R}^\alpha_\beta \boldsymbol{n}^\alpha_{\alpha\beta} = \boldsymbol{n}^\alpha_{\alpha\beta} \tag{2.9}$$

From (2.7):

$$\boldsymbol{n}^\alpha_{\alpha\beta} = \boldsymbol{R}^\alpha_\beta \boldsymbol{n}^\beta_{\alpha\beta} \tag{2.10}$$

Since $\boldsymbol{R}^\alpha_\beta$ is nonsingular, equating (2.9) and (2.10) we see that:

$$\boldsymbol{n}^\alpha_{\alpha\beta} = \boldsymbol{n}^\beta_{\alpha\beta} = \boldsymbol{n}_{\alpha\beta} \tag{2.11}$$

Where the omitted superscript indicates that $\vec{\boldsymbol{n}}_{\alpha\beta}$ can be expressed indistinctly in either $\mathcal{E}_\alpha$ or $\mathcal{E}_\beta$. Equation (2.11) embodies an alternative statement of Euler's theorem, connected to the alias interpretation of rotations: any vector along the rotation axis has the same components in both the original and transformed bases.

To prove the theorem one must show that $\lambda = 1$ is an eigenvalue of $\boldsymbol{R}_\beta^\alpha$, that is, $\det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right) = 0$. Using (1.46b), (1.46c) and (1.46d):

$$
\begin{aligned}
\det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right) &= \det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right)^T = \det\left(\left(\boldsymbol{R}_\beta^\alpha\right)^T - \boldsymbol{I}\right) \\
&= \det\left(\left(\boldsymbol{R}_\beta^\alpha\right)^T\left(\boldsymbol{I} - \boldsymbol{R}_\beta^\alpha\right)\right) = \det\left(\boldsymbol{R}_\beta^\alpha\right)^T \det\left(\boldsymbol{I} - \boldsymbol{R}_\beta^\alpha\right) \\
&= \det\left(\boldsymbol{R}_\beta^\alpha\right) \det\left(-\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right)\right) = (-1)^3 \det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right) \\
&= -\det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right)
\end{aligned}
$$

Hence, $\det\left(\boldsymbol{R}_\beta^\alpha - \boldsymbol{I}\right) = 0$. Note that the number of dimensions ($n = 3$ in this case) is critical for the proof; the theorem only holds for spaces with an uneven number of dimensions. In particular, it is intuitively obvious that no invariant directions exist in a rotation within the two-dimensional plane.

2.1.3 Rodrigues' Rotation Formula

Let f be a proper rotation around the axis defined by unit vector $\vec{\boldsymbol{n}}$, and let ϑ denote the rotation angle. An arbitrary vector $\vec{\boldsymbol{v}}$ can be expressed as the sum of a component parallel to $\vec{\boldsymbol{n}}$ and another perpendicular to $\vec{\boldsymbol{n}}$:

$$\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}_\parallel + \vec{\boldsymbol{v}}_\perp$$

Where:

$$\vec{\boldsymbol{v}}_\parallel = \left(\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}}\right)\vec{\boldsymbol{n}} \tag{2.12}$$

$$\vec{\boldsymbol{v}}_\perp = \vec{\boldsymbol{v}} - \vec{\boldsymbol{v}}_\parallel = \vec{\boldsymbol{v}} - \left(\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}}\right)\vec{\boldsymbol{n}} \tag{2.13}$$

We now construct the following unit vectors:

$$\vec{\boldsymbol{n}}_\perp = \frac{\vec{\boldsymbol{v}}_\perp}{\left|\vec{\boldsymbol{v}}_\perp\right|} \tag{2.14}$$

$$\vec{\boldsymbol{n}}_\times = \vec{\boldsymbol{n}} \times \vec{\boldsymbol{n}}_\perp = \frac{\vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}}_\perp}{\left|\vec{\boldsymbol{v}}_\perp\right|} = \frac{\vec{\boldsymbol{n}} \times \left(\vec{\boldsymbol{v}}_\perp + \vec{\boldsymbol{v}}_\parallel\right)}{\left|\vec{\boldsymbol{v}}_\perp\right|} = \frac{\vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}}}{\left|\vec{\boldsymbol{v}}_\perp\right|} \tag{2.15}$$

The set $\{\vec{\boldsymbol{n}}, \vec{\boldsymbol{n}}_\perp, \vec{\boldsymbol{n}}_\times\}$ forms an orthonormal, right-handed basis.

The rotated vector $\vec{\boldsymbol{w}} = f\left(\vec{\boldsymbol{v}}\right)$ can similarly be expressed as:

$$\vec{\boldsymbol{w}} = \vec{\boldsymbol{w}}_\parallel + \vec{\boldsymbol{w}}_\perp \tag{2.16}$$

From Euler's theorem, we know that $\vec{\boldsymbol{v}}_{\parallel}$ is unchanged by the rotation, so:

$$\vec{\boldsymbol{w}}_{\parallel} = \vec{\boldsymbol{v}}_{\parallel} = (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})\, \vec{\boldsymbol{n}} \tag{2.17}$$

Since the rotation preserves lengths, $|\vec{\boldsymbol{w}}| = |\vec{\boldsymbol{v}}|$, and therefore $|\vec{\boldsymbol{w}}_{\perp}| = |\vec{\boldsymbol{v}}_{\perp}|$.

Working on the plane of rotation, we can express $\vec{\boldsymbol{w}}_{\perp}$ as:

$$\vec{\boldsymbol{w}}_{\perp} = |\vec{\boldsymbol{w}}_{\perp}| \,(\vec{\boldsymbol{n}}_{\perp} \cos\vartheta + \vec{\boldsymbol{n}}_{\times} \sin\vartheta) = |\vec{\boldsymbol{v}}_{\perp}| \,(\vec{\boldsymbol{n}}_{\perp} \cos\vartheta + \vec{\boldsymbol{n}}_{\times} \sin\vartheta)$$

Using (2.14), (2.15) and (2.13):

$$\vec{\boldsymbol{w}}_{\perp} = \vec{\boldsymbol{v}}_{\perp} \cos\vartheta + \vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}} \sin\vartheta = (\vec{\boldsymbol{v}} - (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})\, \vec{\boldsymbol{n}}) \cos\vartheta + \vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}} \sin\vartheta \tag{2.18}$$

Gathering (2.17) and (2.18):

$$\vec{\boldsymbol{w}} = (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})\, \vec{\boldsymbol{n}} + (\vec{\boldsymbol{v}} - (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})\, \vec{\boldsymbol{n}}) \cos\vartheta + \vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}} \sin\vartheta \tag{2.19}$$

Equation (2.19) is known as *Rodrigues' rotation formula*. It is a component-free description of the result of rotating an Euclidean vector $\vec{\boldsymbol{v}}$ an angle ϑ around the axis defined by $\vec{\boldsymbol{n}}$.

Rearranging (2.19) yields the alternative form:

$$\vec{\boldsymbol{w}} = \vec{\boldsymbol{v}} \cos\vartheta + (1 - \cos\vartheta)\, (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})\, \vec{\boldsymbol{n}} + \vec{\boldsymbol{n}} \times \vec{\boldsymbol{v}} \sin\vartheta \tag{2.20}$$

Taking the scalar product with $\vec{\boldsymbol{v}}$ in (2.19) gives:

$$\begin{aligned}\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{w}} &= (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})^2 + \left(|\vec{\boldsymbol{v}}|^2 - (\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}})^2\right) \cos\vartheta \\ |\vec{\boldsymbol{v}}|^2 \cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}} &= |\vec{\boldsymbol{v}}|^2 \left(\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}\right)^2 + |\vec{\boldsymbol{v}}|^2 \left(1 - \left(\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}\right)^2\right) \cos\vartheta\end{aligned}$$

Solving for $\cos\vartheta$ we arrive at the following relation:

$$\vartheta = \arccos\left(\frac{\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}} - \left(\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}\right)^2}{1 - \left(\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}\right)^2}\right) = f\left(\left(\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}\right)^2\right) \tag{2.21}$$

It is easy to verify that f is monotonically increasing. Thus, ϑ is minimum for $\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}} = 0$, which yields $\vartheta_{min} = \theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}}$. In this case, $\vec{\boldsymbol{n}} \perp \vec{\boldsymbol{v}}$, so $\vec{\boldsymbol{v}}_{\parallel} = \vec{\boldsymbol{w}}_{\parallel} = 0$. Therefore, $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ are contained in the plane of rotation, and the direction of $\vec{\boldsymbol{n}}$ is given by $\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}$.

As $|\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{n}}}|$ increases, so does ϑ, until $\vartheta_{max} = \pi$. Setting $\vartheta = \pi$ in (2.19) yields:

$$\vec{\boldsymbol{n}} = \frac{\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}}{2\vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{n}}}$$

Thus, in this case $\vec{\boldsymbol{n}}$ is coplanar with $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$, and parallel to their bisecting line.

The above results show that there are infinite rotations mapping $\vec{\boldsymbol{v}}$ into $\vec{\boldsymbol{w}}$. Of these, the one which requires the minimum angle ($\vartheta_{min} = \theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}}$) corresponds to $\vec{\boldsymbol{n}} \perp \vec{\boldsymbol{v}}$.

Now let rotation f act on an orthonormal basis $\mathcal{E}_\alpha$ to yield another basis $\mathcal{E}_\beta$, so that we may write $\vec{\boldsymbol{n}} = \vec{\boldsymbol{n}}_{\alpha\beta}$ and $\vartheta = \vartheta_{\alpha\beta}$. From the $\mathcal{E}_\alpha$ components of $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$, the axis for the minimum-angle rotation can be computed as:

$$\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{v}^\alpha \times \boldsymbol{w}^\alpha}{|\boldsymbol{v}^\alpha \times \boldsymbol{w}^\alpha|} \tag{2.22}$$

With the direction of $\vec{\boldsymbol{n}}$ given by $\vec{\boldsymbol{v}} \times \vec{\boldsymbol{w}}$, the angle $\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}}$ measured from $\vec{\boldsymbol{v}}$ to $\vec{\boldsymbol{w}}$ lies within $[0, \pi]$. Therefore, we can write:

$$\boldsymbol{v}^\alpha \times \boldsymbol{w}^\alpha = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\,\left|\sin\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}}\right| = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\sin\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}} = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\sin\vartheta_{\alpha\beta} \tag{2.23}$$

Additionally, we have:

$$\boldsymbol{v}^\alpha \cdot \boldsymbol{w}^\alpha = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\cos\theta_{\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}} = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\cos\vartheta_{\alpha\beta} = |\boldsymbol{v}^\alpha|\,|\boldsymbol{w}^\alpha|\cos\vartheta_{\alpha\beta} \tag{2.24}$$

From (2.23) and (2.24), $\vartheta_{\alpha\beta}$ can be found as:

$$\vartheta_{\alpha\beta} = \operatorname{atan2}\left(\boldsymbol{v}^\alpha \times \boldsymbol{w}^\alpha, \boldsymbol{v}^\alpha \cdot \boldsymbol{w}^\alpha\right) \tag{2.25}$$

Finally, by comparing (1.85) and (1.92), we see that the passive interpretation equivalents of (2.22) and (2.25) can be obtained simply by replacing $\boldsymbol{w}^\alpha$ with $\boldsymbol{x}^\alpha$ and $\boldsymbol{v}^\alpha$ with $\boldsymbol{x}^\beta$:

$$\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{x}^\beta \times \boldsymbol{x}^\alpha}{|\boldsymbol{x}^\beta \times \boldsymbol{x}^\alpha|} \tag{2.26}$$

$$\vartheta_{\alpha\beta} = \operatorname{atan2}\left(\boldsymbol{x}^\beta \times \boldsymbol{x}^\alpha, \boldsymbol{x}^\beta \cdot \boldsymbol{x}^\alpha\right) \tag{2.27}$$

2.2 Parameterizations

2.2.1 Axis-Angle

Having established that every three-dimensional rotation can be expressed as an axis-angle combination, we now seek the connection between these two parameters and the rotation matrix.

Let $\mathcal{E}_\alpha$ be an arbitrary orthonormal right-handed basis, and let $\mathcal{E}_\beta$ be the basis that results from applying to $\mathcal{E}_\alpha$ the rotation with axis $\vec{\boldsymbol{n}}_{\alpha\beta}$ and angle $\vartheta_{\alpha\beta}$.

Setting $\vec{\boldsymbol{n}} = \vec{\boldsymbol{n}}_{\alpha\beta}$, $\vartheta = \vartheta_{\alpha\beta}$, $\vec{\boldsymbol{v}} = \vec{\boldsymbol{e}}_{\alpha j}$ and $\vec{\boldsymbol{w}} = \vec{\boldsymbol{e}}_{\beta j}$ in (2.20), we have:

$$\vec{\boldsymbol{e}}_{\beta j} = \vec{\boldsymbol{e}}_{\alpha j}\cos\vartheta_{\alpha\beta} + (1-\cos\vartheta_{\alpha\beta})\left(\vec{\boldsymbol{e}}_{\alpha j}\cdot\vec{\boldsymbol{n}}_{\alpha\beta}\right)\vec{\boldsymbol{n}}_{\alpha\beta} + \vec{\boldsymbol{n}}_{\alpha\beta}\times\vec{\boldsymbol{e}}_{\alpha j}\sin\vartheta_{\alpha\beta}$$

Taking the scalar product with $\vec{\boldsymbol{e}}_{\alpha i}$:

$$\begin{gathered}\vec{\boldsymbol{e}}_{\beta j}\cdot\vec{\boldsymbol{e}}_{\alpha i} = \vec{\boldsymbol{e}}_{\alpha j}\cdot\vec{\boldsymbol{e}}_{\alpha i}\cos\vartheta_{\alpha\beta} + (1-\cos\vartheta_{\alpha\beta})\left(\vec{\boldsymbol{e}}_{\alpha j}\cdot\vec{\boldsymbol{n}}_{\alpha\beta}\right)\left(\vec{\boldsymbol{n}}_{\alpha\beta}\cdot\vec{\boldsymbol{e}}_{\alpha i}\right) \\ +(\vec{\boldsymbol{n}}_{\alpha\beta}\times\vec{\boldsymbol{e}}_{\alpha j})\cdot\vec{\boldsymbol{e}}_{\alpha i}\sin\vartheta_{\alpha\beta}\end{gathered}$$

Applying (1.19), (2.1a) and (1.23):

$$\begin{aligned} e^{\alpha}_{\beta j(i)} = R^{\alpha}_{\beta(i,j)} = \delta_{ij} \cos\vartheta_{\alpha\beta} + \left(1-\cos\vartheta_{\alpha\beta}\right) n^{\alpha}_{\alpha\beta(i)} n^{\alpha}_{\alpha\beta(j)} \\ + \left(\vec{\boldsymbol{n}}_{\alpha\beta} \times \vec{\boldsymbol{e}}_{\alpha j}\right) \cdot \vec{\boldsymbol{e}}_{\alpha i} \sin\vartheta_{\alpha\beta} \end{aligned} \tag{2.28}$$

The last term on the right-hand side can be developed using (1.26f), (1.26a), (1.10), (1.30) and (1.62):

$$\begin{aligned} \left(\vec{\boldsymbol{n}}_{\alpha\beta} \times \vec{\boldsymbol{e}}_{\alpha j}\right) \cdot \vec{\boldsymbol{e}}_{\alpha i} &= \left(\vec{\boldsymbol{e}}_{\alpha j} \times \vec{\boldsymbol{e}}_{\alpha i}\right) \cdot \vec{\boldsymbol{n}}_{\alpha\beta} = -\left(\vec{\boldsymbol{e}}_{\alpha i} \times \vec{\boldsymbol{e}}_{\alpha j}\right) \cdot \vec{\boldsymbol{n}}_{\alpha\beta} \\ &= -\left(\vec{\boldsymbol{e}}_{\alpha i} \times \vec{\boldsymbol{e}}_{\alpha j}\right) \cdot \sum_{k=1}^{3} n^{\alpha}_{\alpha\beta(i)} \vec{\boldsymbol{e}}_{\alpha k} = -\sum_{k=1}^{3} n^{\alpha}_{\alpha\beta(i)} \left(\vec{\boldsymbol{e}}_{\alpha i} \times \vec{\boldsymbol{e}}_{\alpha j}\right) \cdot \vec{\boldsymbol{e}}_{\alpha k} \\ &= -\sum_{k=1}^{3} n^{\alpha}_{\alpha\beta(k)} \epsilon_{ijk} = \left[{n^{\alpha}_{\alpha\beta}}^{\times} \right]_{(i,j)} \end{aligned} \tag{2.29}$$

Substituting (2.29) into (2.28) and rearranging leads to:

$$\begin{aligned} R^{\alpha}_{\beta(i,j)} &= \delta_{ij} \cos\vartheta_{\alpha\beta} + \left[{n^{\alpha}_{\alpha\beta}}^{\times} \right]_{(i,j)} \sin\vartheta_{\alpha\beta} + \left(1-\cos\vartheta_{\alpha\beta}\right) n^{\alpha}_{\alpha\beta(i)} n^{\alpha}_{\alpha\beta(j)} \\ &= \delta_{ij} + \left[{n^{\alpha}_{\alpha\beta}}^{\times} \right]_{(i,j)} \sin\vartheta_{\alpha\beta} + \left(1-\cos\vartheta_{\alpha\beta}\right) \left(n^{\alpha}_{\alpha\beta(i)} n^{\alpha}_{\alpha\beta(j)} - \delta_{ij} \right) \end{aligned} \tag{2.30}$$

Which can be written in matrix form as:

$$\boldsymbol{R}^{\alpha}_{\beta} = \boldsymbol{I} + \left[{\boldsymbol{n}^{\alpha}_{\alpha\beta}}^{\times} \right] \sin\vartheta_{\alpha\beta} + \left(\boldsymbol{n}^{\alpha}_{\alpha\beta} \left(\boldsymbol{n}^{\alpha}_{\alpha\beta} \right)^{T} - \boldsymbol{I} \right) \left(1-\cos\vartheta_{\alpha\beta}\right) \tag{2.31}$$

Drawing on (2.11), we can drop the superscript to write:

$$\boldsymbol{R}^{\alpha}_{\beta} = \boldsymbol{I} + \left[{\boldsymbol{n}_{\alpha\beta}}^{\times} \right] \sin\vartheta_{\alpha\beta} + \left(\boldsymbol{n}_{\alpha\beta} \left(\boldsymbol{n}_{\alpha\beta} \right)^{T} - \boldsymbol{I} \right) \left(1-\cos\vartheta_{\alpha\beta}\right) \tag{2.32}$$

Now, from (1.63e):

$$\boldsymbol{n}_{\alpha\beta} \left(\boldsymbol{n}_{\alpha\beta} \right)^{T} = \left| \vec{\boldsymbol{n}}_{\alpha\beta} \right|^{2} \boldsymbol{I} + \left[{\boldsymbol{n}_{\alpha\beta}}^{\times} \right]^{2} = \boldsymbol{I} + \left[{\boldsymbol{n}_{\alpha\beta}}^{\times} \right]^{2} \tag{2.33}$$

Substituting (2.33) into (2.31) we finally get:

$$\boldsymbol{R}^{\alpha}_{\beta} = \boldsymbol{I} + \left[{\boldsymbol{n}_{\alpha\beta}}^{\times} \right] \sin\vartheta_{\alpha\beta} + \left[{\boldsymbol{n}_{\alpha\beta}}^{\times} \right]^{2} \left(1-\cos\vartheta_{\alpha\beta}\right) = \boldsymbol{R}_{n\vartheta} \left(\boldsymbol{n}_{\alpha\beta}, \vartheta_{\alpha\beta} \right) \tag{2.34}$$

Where we have defined the following operator, which maps the axis-angle representation onto the special orthogonal group $SO(3)$:

$$\boldsymbol{R}_{n\vartheta} \left(\boldsymbol{n}, \vartheta \right) = \boldsymbol{I} + \sin\vartheta \left[\boldsymbol{n}^{\times} \right] + (1-\cos\vartheta) \left[\boldsymbol{n}^{\times} \right]^{2} \tag{2.35}$$

The operator $\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right)$ has the following properties:

$$\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},0\right)=\boldsymbol{I} \tag{2.36a}$$
$$\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right)=\boldsymbol{R}_{n\vartheta}\left(-\boldsymbol{n},-\vartheta\right) \tag{2.36b}$$
$$\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\pi\right)=\boldsymbol{R}_{n\vartheta}\left(-\boldsymbol{n},\pi\right) \tag{2.36c}$$
$$\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},-\vartheta\right)=\boldsymbol{R}_{n\vartheta}\left(-\boldsymbol{n},\vartheta\right)=\left(\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right)\right)^{T} \tag{2.36d}$$

Finally, note that the i-th element in the main diagonal of $\boldsymbol{R}_{\beta}^{\alpha}$ is:

$$R_{\beta(i,i)}^{\alpha}=1+\left(\left(n_{\alpha\beta(i)}^{\alpha}\right)^{2}-1\right)\left(1-\cos\vartheta_{\alpha\beta}\right)=\left(n_{\alpha\beta(i)}^{\alpha}\right)^{2}\left(1-\cos\vartheta_{\alpha\beta}\right)+\cos\vartheta_{\alpha\beta}$$

So the trace of the rotation matrix is simply:

$$\operatorname{tr}\boldsymbol{R}_{\beta}^{\alpha}=\sum_{i=1}^{3}R_{\beta(i,i)}^{\alpha}=\left|\vec{\boldsymbol{n}}_{\alpha\beta}\right|^{2}\left(1-\cos\vartheta_{\alpha\beta}\right)+3\cos\vartheta_{\alpha\beta}=1+2\cos\vartheta_{\alpha\beta} \tag{2.37}$$

2.2.2 Quaternion

2.2.2.1 Introduction to Quaternions

Quaternions are a four-dimensional hypercomplex number system. Much like complex numbers extend real numbers by defining an imaginary unit **i**, quaternions can be regarded as an extension of complex numbers introducing two additional imaginary units, **j** and **k**. The mathematical structure and properties of quaternions are, however, significantly different from those of two-dimensional complex numbers.

A generic quaternion **q** is defined by an expression of the form:

$$\mathbf{q}=q_{(0)}+q_{(1)}\mathbf{i}+q_{(2)}\mathbf{j}+q_{(3)}\mathbf{k} \tag{2.38}$$

The coefficients $q_{(i)}$ are real numbers; $q_{(0)}$ is known as the *real* or *scalar* part of **q**, and $q_{(1)}$, $q_{(2)}$ and $q_{(3)}$ comprise its *imaginary* or *vector* part.

A quaternion with zero imaginary part is called a *real quaternion*. Real numbers can be viewed as a subset of quaternions, so a real quaternion **q** directly identifies with its real part $q_{(0)}$, and it is typically denoted by it. A quaternion with zero real part is called a *pure* or *vector* quaternion.

Quaternions form a 4-dimensional vector space over the field of real numbers $\mathbb{R}$. The set $\mathcal{E}=\{1,\mathbf{i},\mathbf{j},\mathbf{k}\}$ is a basis for this vector space, and the coefficients $q_{(i)}$ are the coordinates of **q** in $\mathcal{E}$. The two fundamental vector space operations are defined for quaternions and real numbers as follows:

a) Quaternion addition:

$$\mathbf{q}+\mathbf{p}=\left(q_{(0)}+p_{(0)}\right)+\left(q_{(1)}+p_{(1)}\right)\mathbf{i}+\left(q_{(2)}+p_{(2)}\right)\mathbf{j}+\left(q_{(3)}+p_{(3)}\right)\mathbf{k}$$

b) Multiplication by a scalar:

$$a\mathbf{q} = aq_{(0)} + aq_{(1)}\mathbf{i} + aq_{(2)}\mathbf{j} + aq_{(3)}\mathbf{k}$$

It is easily verified that these operations satisfy properties (1.1) to (1.8), as required of a vector space.

However, the unique mathematical structure and properties of quaternions arise from the *quaternion product* operation, which is built upon the following multiplicative identities between basis elements:

$$\begin{aligned}
\mathbf{i} \odot \mathbf{i} = \mathbf{j} \odot \mathbf{j} = \mathbf{k} \odot \mathbf{k} &= -1 \\
\mathbf{i} \odot 1 = 1 \odot \mathbf{i} &= \mathbf{i} \\
\mathbf{j} \odot 1 = 1 \odot \mathbf{j} &= \mathbf{j} \\
\mathbf{k} \odot 1 = 1 \odot \mathbf{k} &= \mathbf{k} \\
\mathbf{i} \odot \mathbf{j} = -\mathbf{j} \odot \mathbf{i} &= \mathbf{k} \\
\mathbf{j} \odot \mathbf{k} = -\mathbf{k} \odot \mathbf{j} &= \mathbf{i} \\
\mathbf{k} \odot \mathbf{i} = -\mathbf{i} \odot \mathbf{k} &= \mathbf{j}
\end{aligned} \tag{2.39}$$

Given these identities, the product of two quaternions $\mathbf{q}$ and $\mathbf{p}$ is constructed through distribution as follows:

$$\begin{aligned}
\mathbf{q} \odot \mathbf{p} = &\left(q_{(0)} + q_{(1)}\mathbf{i} + q_{(2)}\mathbf{j} + q_{(3)}\mathbf{k}\right) \odot \left(p_{(0)} + p_{(1)}\mathbf{i} + p_{(2)}\mathbf{j} + p_{(3)}\mathbf{k}\right) \\
&= q_{(0)}p_{(0)} + q_{(0)}p_{(1)}\mathbf{i} + q_{(0)}p_{(2)}\mathbf{j} + q_{(0)}p_{(3)}\mathbf{k} \\
&+ q_{(1)}p_{(0)}\mathbf{i} + q_{(1)}p_{(1)}\mathbf{i} \odot \mathbf{i} + q_{(1)}p_{(2)}\mathbf{i} \odot \mathbf{j} + q_{(1)}p_{(3)}\mathbf{i} \odot \mathbf{k} \\
&+ q_{(2)}p_{(0)}\mathbf{j} + q_{(2)}p_{(1)}\mathbf{j} \odot \mathbf{i} + q_{(2)}p_{(2)}\mathbf{j} \odot \mathbf{j} + q_{(2)}p_{(3)}\mathbf{j} \odot \mathbf{k} \\
&+ q_{(3)}p_{(0)}\mathbf{k} + q_{(3)}p_{(1)}\mathbf{k} \odot \mathbf{i} + q_{(3)}p_{(2)}\mathbf{k} \odot \mathbf{j} + q_{(3)}p_{(3)}\mathbf{k} \odot \mathbf{k}
\end{aligned} \tag{2.40}$$

Applying identities (2.39) in (2.40) and rearranging yields:

$$\begin{aligned}
\mathbf{q} \odot \mathbf{p} = q_{(0)}p_{(0)} &- \left(q_{(1)}p_{(1)} + q_{(2)}p_{(2)} + q_{(3)}p_{(3)}\right) \\
&+ q_{(0)}\left(p_{(1)}\mathbf{i} + p_{(2)}\mathbf{j} + p_{(3)}\mathbf{k}\right) \\
&+ p_{(0)}\left(q_{(1)}\mathbf{i} + q_{(2)}\mathbf{j} + q_{(3)}\mathbf{k}\right) \\
&+ \left(q_{(2)}p_{(3)} - q_{(3)}p_{(2)}\right)\mathbf{i} \\
&+ \left(q_{(3)}p_{(1)} - q_{(1)}p_{(3)}\right)\mathbf{j} \\
&+ \left(q_{(1)}p_{(2)} - q_{(2)}p_{(1)}\right)\mathbf{k}
\end{aligned} \tag{2.41}$$

We now introduce the following notation:

$$\mathbf{q} = \begin{bmatrix} q_{(0)} \\ q_{(1)} \\ q_{(2)} \\ q_{(3)} \end{bmatrix} = \begin{bmatrix} q_{(0)} \\ \boldsymbol{q} \end{bmatrix} \tag{2.42}$$

Where $\boldsymbol{q}$ is a column matrix holding the imaginary coefficients of $\mathbf{q}$:

$$\boldsymbol{q} = \begin{pmatrix} q_{(1)} \\ q_{(2)} \\ q_{(3)} \end{pmatrix}$$

If we interpret these coefficients as the coordinates of an Euclidean vector $\vec{\boldsymbol{q}}$ in some orthonormal basis, then $\boldsymbol{q}$ is the column matrix representation of $\vec{\boldsymbol{q}}$, and expressions (1.57) to (1.60) become applicable. Combined with notation (2.42), this enables us to rewrite the quaternion product (2.41) more compactly as:

$$\mathbf{q} \odot \mathbf{p} = \begin{bmatrix} q_{(0)}p_{(0)} - \boldsymbol{q} \cdot \boldsymbol{p} \\ q_{(0)}\boldsymbol{p} + p_{(0)}\boldsymbol{q} + \boldsymbol{q} \times \boldsymbol{p} \end{bmatrix} \tag{2.43}$$

The cross product term in (2.43) shows that quaternion product is generally *non-commutative*. And, since this non-commutativity arises from the imaginary parts of the operands, it follows that the product by a real quaternion is always commutative. This is consistent with the interpretation of real quaternions as real numbers.

It is not difficult to verify that the quaternion product is, however, *associative* and *distributive* over quaternion addition:

$$(\mathbf{q} \odot \mathbf{p}) \odot \mathbf{r} = \mathbf{q} \odot (\mathbf{p} \odot \mathbf{r}) = \mathbf{q} \odot \mathbf{p} \odot \mathbf{r} \tag{2.44a}$$

$$(\mathbf{q} + \mathbf{p}) \odot \mathbf{r} = \mathbf{q} \odot \mathbf{r} + \mathbf{p} \odot \mathbf{r} \tag{2.44b}$$

The *identity element* for the quaternion product is the real quaternion 1.

The *conjugate* of a quaternion $\mathbf{q}$ is obtained by negating its imaginary part:

$$\mathbf{q}^* = \begin{bmatrix} q_{(0)} \\ -\boldsymbol{q} \end{bmatrix} \tag{2.45}$$

Obviously:

$$(\mathbf{q}^*)^* = \mathbf{q} \tag{2.46}$$

The conjugate of a quaternion product satisfies:

$$\begin{aligned} (\mathbf{q} \odot \mathbf{p})^* &= \begin{bmatrix} q_{(0)}p_{(0)} - \boldsymbol{q} \cdot \boldsymbol{p} \\ -\left(q_{(0)}\boldsymbol{p} + p_{(0)}\boldsymbol{q} + \boldsymbol{q} \times \boldsymbol{p}\right) \end{bmatrix} \\ &= \begin{bmatrix} p_{(0)}q_{(0)} - (-\boldsymbol{p}) \cdot (-\boldsymbol{q}) \\ p_{(0)}\,(-\boldsymbol{q}) + q_{(0)}\,(-\boldsymbol{p}) + (-\boldsymbol{p}) \times (-\boldsymbol{q}) \end{bmatrix} = \mathbf{p}^* \odot \mathbf{q}^* \end{aligned} \tag{2.47}$$

The *norm* of a quaternion $\mathbf{q}$ is the real number defined as:

$$\begin{aligned} \|\mathbf{q}\|^2 = \mathbf{q} \odot \mathbf{q}^* = \mathbf{q}^* \odot \mathbf{q} &= \begin{bmatrix} q_{(0)}q_{(0)} + \boldsymbol{q} \cdot \boldsymbol{q} \\ q_{(0)}\boldsymbol{q} - q_{(0)}\boldsymbol{q} - \boldsymbol{q} \times \boldsymbol{q} \end{bmatrix} \\ &= \begin{bmatrix} {q_{(0)}}^2 + |\boldsymbol{q}|^2 \\ \mathbf{0} \end{bmatrix} = q_{(0)}^2 + |\boldsymbol{q}|^2 = \sum_{i=0}^{i=3} {q_{(i)}}^2 \end{aligned} \tag{2.48}$$

The *inverse* of a quaternion $\mathbf{q}$ is another quaternion $\mathbf{q}^{-1}$ such that:

$$\mathbf{q} \odot \mathbf{q}^{-1} = \mathbf{q}^{-1} \odot \mathbf{q} = 1 \tag{2.49}$$

It is straightforward to verify that:

$$\mathbf{q}^{-1} = \frac{\mathbf{q}^*}{\|\mathbf{q}\|^2} \tag{2.50}$$

Note that $\forall \mathbf{q} \neq 0, \|\mathbf{q}\| > 0$, so the inverse is well-defined except for the null quaternion.

The derivative of a quaternion $\mathbf{q}$ with respect to a scalar variable x is defined as:

$$\frac{d\mathbf{q}}{dx} = \begin{bmatrix} \frac{dq_{(0)}}{dx} \\ \frac{d\boldsymbol{q}}{dx} \end{bmatrix} \tag{2.51}$$

It is easy to see that:

$$\frac{d\mathbf{q}^*}{dx} = \left(\frac{d\mathbf{q}}{dx}\right)^* \tag{2.52}$$

And, since the quaternion product is a linear operation, its derivative satisfies:

$$\frac{d}{dx}(\mathbf{q} \odot \mathbf{p}) = \frac{d\mathbf{q}}{dx} \odot \mathbf{p} + \mathbf{q} \odot \frac{d\mathbf{p}}{dx} \tag{2.53}$$

We have seen thus far that quaternions are a vector space equipped with an associative, non-commutative bilinear product operation, and a corresponding multiplicative inverse. This characterizes their mathematical structure as an *associative, non-commutative division algebra*.

A quaternion $\mathbf{q}$ with $\|\mathbf{q}\| = 1$ is called a *unit quaternion*. For a unit quaternion, (2.50) becomes:

$$\mathbf{q}^{-1} = \mathbf{q}^* \tag{2.54}$$

The set of unit quaternions, together with the quaternion product, jointly satisfy the following properties:

a) Closure: The product of two unit quaternions $\mathbf{q}$ and $\mathbf{p}$ is also a unit quaternion:

$$\begin{aligned}\|\mathbf{q} \odot \mathbf{p}\|^2 = (\mathbf{q} \odot \mathbf{p}) \odot (\mathbf{q} \odot \mathbf{p})^* &= (\mathbf{q} \odot \mathbf{p}) \odot (\mathbf{p}^* \odot \mathbf{q}^*) \\ &= \mathbf{q} \odot (\mathbf{p} \odot \mathbf{p}^*) \odot \mathbf{q}^* = \mathbf{q} \odot \mathbf{q}^* = 1\end{aligned}$$

b) Identity: The real quaternion 1, which is the identity element for the quaternion product, is also a unit quaternion.

c) Associativity: This follows directly from the associativity of quaternion product in general.

d) Invertibility: The inverse of a quaternion exists as long as $\|\mathbf{q}\| \neq 0$, so every unit quaternion has an inverse, given by (2.54).

These properties characterize the set of unit quaternions as a *group*.

2.2.2.2 Unit Quaternions As an Attitude Descriptor

Our goal here is to uncover the connection between unit quaternions and rotation matrices, and to show how unit quaternions can be used as an alternative attitude descriptor.

Let us begin by defining the following quaternion parameterization, where $\boldsymbol{n}$ is the column matrix representation of some Euclidean unit vector $\vec{\boldsymbol{n}}$ and $\vartheta \in [-\pi, \pi)$ is an arbitrary angle:

$$\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta) = \begin{bmatrix} \cos(\vartheta/2) \\ \boldsymbol{n}\sin(\vartheta/2) \end{bmatrix} \tag{2.55}$$

Note that, if $\vec{\boldsymbol{n}}$ is indeed a unit vector, $\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta)$ always yields a unit quaternion:

$$\begin{aligned} \|\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta)\|^2 &= (\cos(\vartheta/2))^2 + |\boldsymbol{n}|^2 (\sin(\vartheta/2))^2 \\ &= (\cos(\vartheta/2))^2 + (\sin(\vartheta/2))^2 = 1 \end{aligned} \tag{2.56}$$

Since a proper rotation in Euclidean space is fully defined by a unit vector and an angle, it follows that any such rotation can be represented by a suitable unit quaternion, given by (2.55). Conversely, any unit quaternion $\mathbf{r}$ may be interpreted as a rotation in Euclidean space. Note, however, that:

$$\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta + 2\pi) = \begin{bmatrix} \cos(\vartheta/2 + \pi) \\ \boldsymbol{n}\sin(\vartheta/2 + \pi) \end{bmatrix} = \begin{bmatrix} -\cos(\vartheta/2) \\ -\boldsymbol{n}\sin(\vartheta/2) \end{bmatrix} = -\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta) \tag{2.57}$$

Because a rotation by an angle $\vartheta + 2\pi$ yields the same end result as a rotation by ϑ, we must conclude that any two unit quaternions $\mathbf{r}$ and $-\mathbf{r}$ represent the same actual rotation.

We now define the following matrix parameterization, where $\mathbf{r}$ denotes an arbitrary unit quaternion:

$$\boldsymbol{R}_{\mathbf{r}}(\mathbf{r}) = \boldsymbol{I} + 2r_{(0)}[\boldsymbol{r}^{\times}] + 2[\boldsymbol{r}^{\times}]^2 \tag{2.58}$$

Carrying out the matrix operations in (2.58) yields:

$$\boldsymbol{R}_{\mathbf{r}}(\mathbf{r}) = \begin{pmatrix} 1 - 2r_{(2)}^2 - 2r_{(3)}^2 & 2r_{(1)}r_{(2)} - 2r_{(0)}r_{(3)} & 2r_{(1)}r_{(3)} + 2r_{(0)}r_{(2)} \\ 2r_{(1)}r_{(2)} + 2r_{(0)}r_{(3)} & 1 - 2r_{(1)}^2 - 2r_{(3)}^2 & 2r_{(2)}r_{(3)} - 2r_{(0)}r_{(1)} \\ 2r_{(1)}r_{(3)} - 2r_{(0)}r_{(2)} & 2r_{(2)}r_{(3)} + 2r_{(0)}r_{(1)} & 1 - 2r_{(1)}^2 - 2r_{(2)}^2 \end{pmatrix} \tag{2.59}$$

Now, setting $\mathbf{r} = \mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta)$ in (2.58) leads to:

$$\begin{aligned} \boldsymbol{R}_{\mathbf{r}}(\mathbf{r}_{n\vartheta}(\boldsymbol{n}, \vartheta)) &= \boldsymbol{I} + 2\cos(\vartheta/2)\left[(\boldsymbol{n}\sin(\vartheta/2))^{\times}\right] + 2\left[(\boldsymbol{n}\sin(\vartheta/2))^{\times}\right]^2 \\ &= \boldsymbol{I} + 2\cos(\vartheta/2)\sin(\vartheta/2)[\boldsymbol{n}^{\times}] + 2(\sin(\vartheta/2))^2[\boldsymbol{n}^{\times}]^2 \\ &= \boldsymbol{I} + \sin\vartheta[\boldsymbol{n}^{\times}] + (1 - \cos\vartheta)[\boldsymbol{n}^{\times}]^2 \end{aligned} \tag{2.60}$$

Where we have made use of property (1.63a) and the trigonometric identities:

$$\sin\vartheta = 2\sin\frac{\vartheta}{2}\cos\frac{\vartheta}{2}$$
$$\left(\sin\frac{\vartheta}{2}\right)^2 = \frac{1}{2}(1-\cos\vartheta)$$

By comparison with (2.35), we see that (2.60) is precisely the axis angle parameterization of the rotation matrix:

$$\boldsymbol{R}_{\mathbf{r}}\left(\mathbf{r}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right)\right) = \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right)$$

Turning to (2.58), it is easy to see that $\boldsymbol{R}_{\mathbf{r}}(\mathbf{r}) = \boldsymbol{R}_{\mathbf{r}}(-\mathbf{r})$, that is, both $\mathbf{r}$ and $-\mathbf{r}$ map into the same rotation matrix. Unit quaternions are thus said to provide a *double cover* of the special orthogonal group $SO(3)$.

Now, let $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ be two orthonormal right-handed bases for the Euclidean vector space, with unit vector $\vec{\boldsymbol{n}}_{\alpha\beta}$ and angle $\vartheta_{\alpha\beta}$ defining the proper rotation that maps $\mathcal{E}_\alpha$ into $\mathcal{E}_\beta$. Recall that the components of $\vec{\boldsymbol{n}}_{\alpha\beta}$ are the same in both bases and they are denoted simply by $\boldsymbol{n}_{\alpha\beta}$. For convenience, we introduce the following notational convention:

$$\mathbf{r}^\alpha_\beta = \mathbf{r}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta},\vartheta_{\alpha\beta}\right) = \begin{bmatrix}\cos(\vartheta_{\alpha\beta}/2)\\ \sin(\vartheta_{\alpha\beta}/2)\boldsymbol{n}_{\alpha\beta}\end{bmatrix} \tag{2.61}$$

So that:

$$\boldsymbol{R}^\alpha_\beta = \boldsymbol{R}_{\mathbf{r}}\left(\mathbf{r}^\alpha_\beta\right) = \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta},\vartheta_{\alpha\beta}\right) \tag{2.62}$$

At this point we have yet to see how this seemingly arbitrary quaternion representation of rotations may be useful at all. To find out, we begin by computing the following quaternion product, where $\mathbf{r}$ is a unit quaternion and $\mathbf{p}$ is a pure quaternion:

$$\mathbf{r}\odot\mathbf{p}\odot\mathbf{r}^* = (\mathbf{r}\odot\mathbf{p})\odot\mathbf{r}^* = \begin{bmatrix}-\boldsymbol{r}\cdot\boldsymbol{p}\\ r_{(0)}\boldsymbol{p}+\boldsymbol{r}\times\boldsymbol{p}\end{bmatrix}\odot\begin{bmatrix}r_{(0)}\\ -\boldsymbol{r}\end{bmatrix} \tag{2.63}$$

Explicit computation of the rightmost term yields, after applying scalar and cross product properties:

$$\mathbf{r}\odot\mathbf{p}\odot\mathbf{r}^* = \begin{bmatrix}0\\ r_{(0)}^2\boldsymbol{p}+\boldsymbol{r}(\boldsymbol{r}\cdot\boldsymbol{p})+2r_{(0)}\boldsymbol{r}\times\boldsymbol{p}+\boldsymbol{r}\times(\boldsymbol{r}\times\boldsymbol{p})\end{bmatrix} \tag{2.64}$$

Using (1.59) and (1.60) we can rewrite the scalar and cross products in (2.64) as matrix operations:

$$\mathbf{r}\odot\mathbf{p}\odot\mathbf{r}^* = \begin{bmatrix}0\\ r_{(0)}^2\boldsymbol{p}+\boldsymbol{r}(\boldsymbol{r}^T\boldsymbol{p})+2r_{(0)}\left[\boldsymbol{r}^\times\right]\boldsymbol{p}+\left[\boldsymbol{r}^\times\right]\left(\left[\boldsymbol{r}^\times\right]\boldsymbol{p}\right)\end{bmatrix} \tag{2.65}$$

Exploiting the associativity property of the matrix product:

$$\mathbf{r}\odot\mathbf{p}\odot\mathbf{r}^* = \begin{bmatrix}0\\ \left(r_{(0)}^2\boldsymbol{I}+\boldsymbol{r}\boldsymbol{r}^T+2r_{(0)}\left[\boldsymbol{r}^\times\right]+\left[\boldsymbol{r}^\times\right]^2\right)\boldsymbol{p}\end{bmatrix} \tag{2.66}$$

Turning now to (1.63e) and setting $\boldsymbol{v} = \boldsymbol{w} = \boldsymbol{r}$ gives:

$$\boldsymbol{r}\boldsymbol{r}^T = [r^\times]^2 + (\boldsymbol{r}^T\boldsymbol{r})\boldsymbol{I} = [\boldsymbol{r}^\times]^2 + |\boldsymbol{r}|^2\,\boldsymbol{I} \tag{2.67}$$

Inserting (2.67) into (2.66) we arrive at a remarkable result:

$$\begin{aligned}
\mathbf{r} \odot \mathbf{p} \odot \mathbf{r}^* &= \begin{bmatrix} 0 \\ \left(r_{(0)}^2\boldsymbol{I} + |\boldsymbol{r}|^2\,\boldsymbol{I} + [\boldsymbol{r}^\times]^2 + 2r_{(0)}\,[\boldsymbol{r}^\times] + [\boldsymbol{r}^\times]^2\right)\boldsymbol{p} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ \left(\|\mathbf{r}\|^2\,\boldsymbol{I} + 2r_{(0)}\,[\boldsymbol{r}^\times] + 2\,[\boldsymbol{r}^\times]^2\right)\boldsymbol{p} \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ \left(\boldsymbol{I} + 2r_{(0)}\,[\boldsymbol{r}^\times] + 2\,[\boldsymbol{r}^\times]^2\right)\boldsymbol{p} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{R}_{\mathbf{r}}\,(\mathbf{r})\,\boldsymbol{p} \end{bmatrix}
\end{aligned} \tag{2.68}$$

Now let $\vec{\boldsymbol{x}}$ be an arbitrary Euclidean vector with components $\boldsymbol{x}^\alpha$ in $\mathcal{E}_\alpha$ and $\boldsymbol{x}^\beta$ in $\mathcal{E}_\beta$, and let us define the pure quaternions:

$$\mathbf{x}^\alpha = \begin{bmatrix} 0 \\ \boldsymbol{x}^\alpha \end{bmatrix} \tag{2.69a}$$

$$\mathbf{x}^\beta = \begin{bmatrix} 0 \\ \boldsymbol{x}^\beta \end{bmatrix} \tag{2.69b}$$

Setting $\mathbf{r} = \mathbf{r}_\beta^\alpha$ and $\mathbf{p} = \mathbf{x}^\beta$ in (2.68) yields:

$$\mathbf{r}_\beta^\alpha \odot \mathbf{x}^\beta \odot \left(\mathbf{r}_\beta^\alpha\right)^* = \begin{bmatrix} 0 \\ \boldsymbol{R}_{\mathbf{r}}\left(\mathbf{r}_\beta^\alpha\right)\boldsymbol{x}^\beta \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{R}_\beta^\alpha\boldsymbol{x}^\beta \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{x}^\alpha \end{bmatrix} = \mathbf{x}^\alpha \tag{2.70}$$

Equation (2.70) shows how a change of basis can be carried out within the quaternion representation. It is analogous to (2.7).

If we now consider a second proper rotation mapping the transformed basis $\mathcal{E}_\beta$ into another $\mathcal{E}_\delta$ and define the unit quaternion $\mathbf{r}_\delta^\beta = \mathbf{r}_{n\vartheta}\left(\boldsymbol{n}_{\beta\delta}, \vartheta_{\beta\delta}\right)$, by analogy with (2.70) we can write:

$$\mathbf{x}^\beta = \mathbf{r}_\delta^\beta \odot \mathbf{x}^\delta \odot \left(\mathbf{r}_\delta^\beta\right)^* \tag{2.71}$$

Substituting (2.71) into (2.70) and applying properties (2.44a) and (2.47):

$$\begin{aligned}
\mathbf{x}^\alpha &= \mathbf{r}_\beta^\alpha \odot \mathbf{x}^\beta \odot \left(\mathbf{r}_\beta^\alpha\right)^* = \mathbf{r}_\beta^\alpha \odot \left(\mathbf{r}_\delta^\beta \odot \mathbf{x}^\delta \odot \left(\mathbf{r}_\delta^\beta\right)^*\right) \odot \left(\mathbf{r}_\beta^\alpha\right)^* \\
&= \left(\mathbf{r}_\beta^\alpha \odot \mathbf{r}_\delta^\beta\right) \odot \mathbf{x}^\delta \odot \left(\left(\mathbf{r}_\delta^\beta\right)^* \odot \left(\mathbf{r}_\beta^\alpha\right)^*\right) \\
&= \left(\mathbf{r}_\beta^\alpha \odot \mathbf{r}_\delta^\beta\right) \odot \mathbf{x}^\delta \odot \left(\mathbf{r}_\beta^\alpha \odot \mathbf{r}_\delta^\beta\right)^* = \mathbf{r}_\delta^\alpha \odot \mathbf{x}^\delta \odot \left(\mathbf{r}_\delta^\alpha\right)^*
\end{aligned} \tag{2.72}$$

Therefore, we have that:

$$\mathbf{r}_\delta^\alpha = \mathbf{r}_\beta^\alpha \odot \mathbf{r}_\delta^\beta \tag{2.73}$$

This result is analogous to (2.4). It shows how rotation composition is accomplished within the quaternion representation. Here, the closure property of the unit quaternion group guarantees that $\mathbf{r}^{\alpha}_{\delta}$ is also a unit quaternion. This is to be expected, since the composition of proper rotations yields another proper rotation.

Finally, multiplying (2.70) by $\left(\mathbf{r}^{\alpha}_{\beta}\right)^*$ on the left and by $\mathbf{r}^{\alpha}_{\beta}$ on the right gives the reciprocal change of basis:

$$\mathbf{x}^{\beta} = \left(\mathbf{r}^{\alpha}_{\beta}\right)^* \odot \mathbf{x}^{\alpha} \odot \mathbf{r}^{\alpha}_{\beta} \tag{2.74}$$

From which we find the quaternion equivalent of (2.2) to be:

$$\mathbf{r}^{\beta}_{\alpha} = \left(\mathbf{r}^{\alpha}_{\beta}\right)^* \tag{2.75}$$

Equations (2.70), (2.73) and (2.75) show that unit quaternions are indeed a valid alternative to rotation matrices for attitude representation. Within the quaternion representation, basis changes and rotation composition are realized through quaternion multiplication, instead of matrix multiplication. And reciprocal rotations correspond to quaternion conjugation, instead of matrix transposition. Like rotation matrices, unit quaternions are a non-minimal attitude descriptor, with their four coefficients bound by a unit norm constraint.

2.2.3 Rotation Vector

For a rotation described by unit vector $\vec{\boldsymbol{n}}_{\alpha\beta}$ and angle $\vartheta_{\alpha\beta}$, the rotation vector is defined as:

$$\vec{\boldsymbol{\rho}}_{\alpha\beta} = \vartheta_{\alpha\beta}\vec{\boldsymbol{n}}_{\alpha\beta} \tag{2.76}$$

From (2.11) we can write:

$$\boldsymbol{\rho}^{\alpha}_{\alpha\beta} = \boldsymbol{\rho}^{\beta}_{\alpha\beta} = \boldsymbol{\rho}_{\alpha\beta} = \vartheta_{\alpha\beta}\boldsymbol{n}_{\alpha\beta} \tag{2.77}$$

Now, taking the norm of (2.76):

$$\left|\vec{\boldsymbol{\rho}}_{\alpha\beta}\right| = \left|\vartheta_{\alpha\beta}\vec{\boldsymbol{n}}_{\alpha\beta}\right| = \left|\vartheta_{\alpha\beta}\right|\left|\vec{\boldsymbol{n}}_{\alpha\beta}\right| = \left|\vartheta_{\alpha\beta}\right| = \vartheta_{\alpha\beta}\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right) \tag{2.78}$$

Thus, the axis and angle can be recovered from the rotation vector as:

$$\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{\rho}_{\alpha\beta}}{\vartheta_{\alpha\beta}} = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\frac{\boldsymbol{\rho}_{\alpha\beta}}{\left|\boldsymbol{\rho}_{\alpha\beta}\right|} \tag{2.79a}$$

$$\vartheta_{\alpha\beta} = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\left|\boldsymbol{\rho}_{\alpha\beta}\right| \tag{2.79b}$$

Note that, on account of (2.36b), the choice of sign is irrelevant as long as it is consistent between (2.79a) and (2.79b). We shall choose $\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right) = 1$:

$$\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{\rho}_{\alpha\beta}}{\left|\boldsymbol{\rho}_{\alpha\beta}\right|} \tag{2.80a}$$

$$\vartheta_{\alpha\beta} = \left|\boldsymbol{\rho}_{\alpha\beta}\right| \tag{2.80b}$$

To find the connection between rotation vector and rotation matrix, we substitute $\sin\vartheta$ and $\cos\vartheta$ in (2.35) with their respective Taylor series:

$$\begin{aligned}\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right) &= \boldsymbol{I} + \left[\boldsymbol{n}^{\times}\right]\sum_{k=0}^{\infty}\frac{(-1)^k\vartheta^{2k+1}}{(2k+1)!} + \left[\boldsymbol{n}^{\times}\right]^2\left(1-\sum_{k=0}^{\infty}\frac{(-1)^k\vartheta^{2k}}{(2k)!}\right) \\ &= \boldsymbol{I} + \sum_{k=0}^{\infty}\left[\boldsymbol{n}^{\times}\right]\frac{(-1)^k\vartheta^{2k+1}}{(2k+1)!} - \sum_{k=1}^{\infty}\left[\boldsymbol{n}^{\times}\right]^2\frac{(-1)^k\vartheta^{2k}}{(2k)!}\end{aligned} \tag{2.81}$$

Applying (1.64a) and (1.64b) with $|\boldsymbol{n}| = 1$ yields:

$$\left[\boldsymbol{n}^{\times}\right]^{2k} = -(-1)^k\left[\boldsymbol{n}^{\times}\right]^2, \forall k = 1, 2, ... \tag{2.82a}$$

$$\left[\boldsymbol{n}^{\times}\right]^{2k+1} = (-1)^k\left[\boldsymbol{n}^{\times}\right], \forall k = 0, 1, ... \tag{2.82b}$$

Substituting (2.82a) and (2.82b) into (2.81) yields:

$$\begin{aligned}\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n},\vartheta\right) &= \boldsymbol{I} + \sum_{k=0}^{\infty}\left[\boldsymbol{n}^{\times}\right]^{2k+1}\frac{\vartheta^{2k+1}}{(2k+1)!} + \sum_{k=1}^{\infty}\left[\boldsymbol{n}^{\times}\right]^{2k}\frac{\vartheta^{2k}}{(2k)!} \\ &= \boldsymbol{I} + \sum_{k=0}^{\infty}\frac{\left[(\vartheta\boldsymbol{n})^{\times}\right]^{2k+1}}{(2k+1)!} + \sum_{k=1}^{\infty}\frac{\left[(\vartheta\boldsymbol{n})^{\times}\right]^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty}\frac{\left[(\vartheta\boldsymbol{n})^{\times}\right]^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty}\frac{\left[(\vartheta\boldsymbol{n})^{\times}\right]^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty}\frac{\left[(\vartheta\boldsymbol{n})^{\times}\right]^{k}}{k!} = \exp\left[(\vartheta\boldsymbol{n})^{\times}\right]\end{aligned}$$

Therefore:

$$\boldsymbol{R}^{\alpha}_{\beta} = \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta},\vartheta_{\alpha\beta}\right) = \exp\left[\left(\vartheta_{\alpha\beta}\boldsymbol{n}_{\alpha\beta}\right)^{\times}\right] = \exp\left[\boldsymbol{\rho}_{\alpha\beta}{}^{\times}\right] = \boldsymbol{R}_{\rho}\left(\boldsymbol{\rho}_{\alpha\beta}\right) \tag{2.83}$$

Where we have defined the following operator, which maps the rotation vector representation onto the special orthogonal group $SO(3)$:

$$\boldsymbol{R}_{\rho}\left(\boldsymbol{\rho}\right) = \exp\left[\boldsymbol{\rho}^{\times}\right] = \sum_{k=0}^{\infty}\frac{1}{k!}\left[\boldsymbol{\rho}^{\times}\right]^k \tag{2.84}$$

Note the following properties:

$$\boldsymbol{R}_{\rho}\left(\boldsymbol{0}\right) = \boldsymbol{I}$$

$$\boldsymbol{R}_{\rho}\left(-\boldsymbol{\rho}\right) = \left(\boldsymbol{R}_{\rho}\left(\boldsymbol{\rho}\right)\right)^T$$

A rotation vector parameterization of quaternion $\mathbf{r}_\beta^\alpha$ is readily obtained from (2.61), (2.80b) and (2.80a):

$$\begin{aligned}\mathbf{r}_\beta^\alpha &= \mathbf{r}_{n\vartheta}\left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\frac{\boldsymbol{\rho}_{\alpha\beta}}{\left|\boldsymbol{\rho}_{\alpha\beta}\right|}, \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\left|\boldsymbol{\rho}_{\alpha\beta}\right|\right) \\ &= \begin{bmatrix}\cos\left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\left|\boldsymbol{\rho}_{\alpha\beta}\right|/2\right) \\ \left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\boldsymbol{\rho}_{\alpha\beta}/\left|\boldsymbol{\rho}_{\alpha\beta}\right|\right)\sin\left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\left|\boldsymbol{\rho}_{\alpha\beta}\right|/2\right)\end{bmatrix} \\ &= \begin{bmatrix}\cos\left(\left|\boldsymbol{\rho}_{\alpha\beta}\right|/2\right) \\ \left(\boldsymbol{\rho}_{\alpha\beta}/\left|\boldsymbol{\rho}_{\alpha\beta}\right|\right)\sin\left(\left|\boldsymbol{\rho}_{\alpha\beta}\right|/2\right)\end{bmatrix} = \mathbf{r}_\rho\left(\boldsymbol{\rho}_{\alpha\beta}\right)\end{aligned} \tag{2.85}$$

Where:

$$\mathbf{r}_\rho\left(\boldsymbol{\rho}\right) = \begin{bmatrix}\cos\left(\left|\boldsymbol{\rho}\right|/2\right) \\ \left(\boldsymbol{\rho}/\left|\boldsymbol{\rho}\right|\right)\sin\left(\left|\boldsymbol{\rho}\right|/2\right)\end{bmatrix} \tag{2.86}$$

Parameterization (2.86) is ill-defined for $|\boldsymbol{\rho}| \to 0$. An alternative form can be found by replacing $\cos\left(|\boldsymbol{\rho}|/2\right)$ and $\sin\left(|\boldsymbol{\rho}|/2\right)$ with their respective Taylor series. After some manipulation, this yields:

$$\mathbf{r}_\rho\left(\boldsymbol{\rho}\right) = \sum_{k=0}^{\infty}(-1)^k\left(\frac{|\boldsymbol{\rho}|}{2}\right)^{2k}\begin{bmatrix}\dfrac{1}{(2k)!} \\ \dfrac{1}{(2k+1)!}\dfrac{|\boldsymbol{\rho}|}{2}\end{bmatrix} \tag{2.87}$$

The rotation vector condenses the axis-angle representation into three scalar parameters. Thus, it is a minimal attitude descriptor.

It must be emphasized that, in general, the addition of rotation vectors does not correspond to the composition of rotations (in fact, it is not even physically meaningful). This should not be surprising, given that vector addition is commutative and composition of rotations in general is not.

2.2.4 Euler Angles

A rotation around one of the three coordinate axes is called a *basic* or *elemental rotation*. Elemental rotations have the following matrix and quaternion representations:

a) Rotation around the x-axis:

$$\boldsymbol{R}_{\boldsymbol{x}}(\vartheta) = \boldsymbol{R}_{n\vartheta}\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}^T, \vartheta\right) = \begin{pmatrix}1 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta \\ 0 & \sin\vartheta & \cos\vartheta\end{pmatrix} \tag{2.88a}$$

$$\mathbf{r}_x(\vartheta) = \mathbf{r}_{n\vartheta}\left(\begin{pmatrix}1 & 0 & 0\end{pmatrix}^T, \vartheta\right) = \begin{bmatrix}\cos\vartheta/2 \\ \sin\vartheta/2 \\ 0 \\ 0\end{bmatrix} \tag{2.88b}$$

b) Rotation around the y-axis:

$$\boldsymbol{R}_{\boldsymbol{y}}(\vartheta) = \boldsymbol{R}_{n\vartheta}\left(\begin{pmatrix}0 & 1 & 0\end{pmatrix}^T, \vartheta\right) = \begin{pmatrix}\cos\vartheta & 0 & \sin\vartheta \\ 0 & 1 & 0 \\ -\sin\vartheta & 0 & \cos\vartheta\end{pmatrix} \tag{2.89a}$$

$$\mathbf{r}_y(\vartheta) = \mathbf{r}_{n\vartheta}\left(\begin{pmatrix}0 & 1 & 0\end{pmatrix}^T, \vartheta\right) = \begin{bmatrix}\cos\vartheta/2 \\ 0 \\ \sin\vartheta/2 \\ 0\end{bmatrix} \tag{2.89b}$$

c) Rotation around the z-axis:

$$\boldsymbol{R}_{\boldsymbol{z}}(\vartheta) = \boldsymbol{R}_{n\vartheta}\left(\begin{pmatrix}0 & 0 & 1\end{pmatrix}^T, \vartheta\right) = \begin{pmatrix}\cos\vartheta & -\sin\vartheta & 0 \\ \sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1\end{pmatrix} \tag{2.90a}$$

$$\mathbf{r}_z(\vartheta) = \mathbf{r}_{n\vartheta}\left(\begin{pmatrix}0 & 0 & 1\end{pmatrix}^T, \vartheta\right) = \begin{bmatrix}\cos\vartheta/2 \\ 0 \\ 0 \\ \sin\vartheta/2\end{bmatrix} \tag{2.90b}$$

The rotation matrices above can be obtained from (2.34), but they are more easily constructed from (2.1a) by applying simple trigonometry in the plane of rotation.

Any proper rotation in three-dimensional space can be expressed a sequence of three elemental rotations such that no consecutive rotations are about the same axis. This fact was originally shown by Euler, and the angles corresponding to these elemental rotations are commonly known as *Euler angles*.

The restriction that successive rotation axes be distinct allows for twelve possible sequences to describe an arbitrary rotation:

$$\begin{array}{cccccc} x-y-x & x-z-x & y-x-y & y-z-y & z-x-z & z-y-z \\ x-y-z & x-z-y & y-x-z & y-z-x & z-x-y & z-y-x \end{array}$$

The sequence most frequently used in aerospace is $z-y-x$, for which the steps rotating an orthonormal basis $\mathcal{E}_\alpha$ into another $\mathcal{E}_\beta$ are:

1) $\mathcal{E}_\alpha$ rotates around its z axis by an angle $\psi_{\alpha\beta}$ to yield intermediate basis $\mathcal{E}_\delta$
2) $\mathcal{E}_\delta$ rotates around its y axis by an angle $\theta_{\alpha\beta}$ to yield intermediate basis $\mathcal{E}_\gamma$
3) $\mathcal{E}_\gamma$ rotates around its x axis by an angle $\phi_{\alpha\beta}$ to yield $\mathcal{E}_\beta$

Euler angles $\psi_{\alpha\beta}$, $\theta_{\alpha\beta}$, $\psi_{\alpha\beta}$ are called respectively *azimuth*, *inclination* and *bank*.

Rotation matrix $\boldsymbol{R}_\beta^\alpha$ can be computed through composition as:

$$\boldsymbol{R}_\beta^\alpha = \boldsymbol{R}_\delta^\alpha \boldsymbol{R}_\gamma^\delta \boldsymbol{R}_\beta^\gamma = \boldsymbol{R}_{\boldsymbol{z}}(\psi_{\alpha\beta})\boldsymbol{R}_{\boldsymbol{y}}(\theta_{\alpha\beta})\boldsymbol{R}_{\boldsymbol{x}}(\phi_{\alpha\beta})$$

We now define the following parameterization:

$$\boldsymbol{R}_\Psi(\psi,\theta,\phi) = \boldsymbol{R}_{\boldsymbol{z}}(\psi)\boldsymbol{R}_{\boldsymbol{y}}(\theta)\boldsymbol{R}_{\boldsymbol{x}}(\phi) \tag{2.91}$$

So that:

$$\boldsymbol{R}^\alpha_\beta = \boldsymbol{R}_\Psi\left(\psi_{\alpha\beta},\theta_{\alpha\beta},\psi_{\alpha\beta}\right) \tag{2.92}$$

Substituting (2.88a), (2.89a) and (2.90a) into (2.91) and carrying out the matrix products yields:

$$\boldsymbol{R}_\Psi(\psi,\theta,\phi) = \begin{pmatrix} c_\psi c_\theta & c_\psi s_\theta s_\phi - s_\psi c_\phi & c_\psi s_\theta c_\phi + s_\psi s_\phi \\ s_\psi c_\theta & s_\psi s_\theta s_\phi + c_\psi c_\phi & s_\psi s_\theta c_\phi - c_\psi s_\phi \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{pmatrix} \tag{2.93}$$

Where s_ϑ and c_ϑ are shorthands for $\sin\vartheta$ and $\cos\vartheta$, respectively.

The quaternion equivalents of (2.91) and (2.92) are:

$$\mathbf{r}_\Psi(\psi,\theta,\phi) = \mathbf{r}_z(\psi)\odot\mathbf{r}_y(\theta)\odot\mathbf{r}_x(\phi) \tag{2.94}$$

$$\mathbf{r}^\alpha_\beta = \mathbf{r}_\Psi\left(\psi_{\alpha\beta},\theta_{\alpha\beta},\psi_{\alpha\beta}\right) \tag{2.95}$$

Substituting (2.88b), (2.89b) and (2.90b) into (2.94) and carrying out the quaternion products yields:

$$\mathbf{r}_\Psi(\psi,\theta,\phi) = \begin{bmatrix} \cos(\psi/2)\cos(\theta/2)\cos(\phi/2) + \sin(\psi/2)\sin(\theta/2)\sin(\phi/2) \\ \cos(\psi/2)\cos(\theta/2)\sin(\phi/2) - \sin(\psi/2)\sin(\theta/2)\cos(\phi/2) \\ \cos(\psi/2)\sin(\theta/2)\cos(\phi/2) + \sin(\psi/2)\cos(\theta/2)\sin(\phi/2) \\ \sin(\psi/2)\cos(\theta/2)\cos(\phi/2) - \cos(\psi/2)\sin(\theta/2)\sin(\phi/2) \end{bmatrix} \tag{2.96}$$

2.3 Infinitesimal Rotations

If $|\boldsymbol{\rho}| \to 0$, then $|\rho_{(i)}| \to 0, \forall i \in \{1,2,3\}$, and all powers of $[\boldsymbol{\rho}^\times]$ beyond $k=1$ in (2.84) can be neglected, which gives:

$$\boldsymbol{R}_\rho(\boldsymbol{\rho}) \approx \boldsymbol{I} + [\boldsymbol{\rho}^\times] \tag{2.97}$$

Using (1.39b) and (1.63d):

$$\left(\boldsymbol{R}_\rho(\boldsymbol{\rho})\right)^T \approx \left(\boldsymbol{I} + [\boldsymbol{\rho}^\times]\right)^T = \boldsymbol{I} - [\boldsymbol{\rho}^\times] \tag{2.98}$$

The orthogonality of $\boldsymbol{R}_\rho(\boldsymbol{\rho})$ is preserved to first order by these approximations:

$$\left(\boldsymbol{R}_\rho(\boldsymbol{\rho})\right)\left(\boldsymbol{R}_\rho(\boldsymbol{\rho})\right)^T = \left(\boldsymbol{I} + [\boldsymbol{\rho}^\times]\right)\left(\boldsymbol{I} - [\boldsymbol{\rho}^\times]\right) = \boldsymbol{I} - [\boldsymbol{\rho}^\times]^2 \approx \boldsymbol{I}$$

Now let matrices $\boldsymbol{R}_\delta^\alpha$ and $\boldsymbol{R}_\beta^\delta$ represent two infinitesimal rotations, so that $|\vartheta_{\alpha\delta}| = |\boldsymbol{\rho}_{\alpha\delta}| \to 0$ and $|\vartheta_{\delta\beta}| = |\boldsymbol{\rho}_{\delta\beta}| \to 0$. Then:

$$\boldsymbol{R}_\delta^\alpha \approx \boldsymbol{I} + [\boldsymbol{\rho}_{\alpha\delta}{}^\times]$$
$$\boldsymbol{R}_\beta^\delta \approx \boldsymbol{I} + \left[\boldsymbol{\rho}_{\alpha\beta}{}^\times\right]$$

Applying (2.4) and neglecting higher order terms:

$$\begin{aligned}\boldsymbol{R}_\beta^\alpha = \boldsymbol{R}_\delta^\alpha \boldsymbol{R}_\beta^\delta &= (\boldsymbol{I} + [\boldsymbol{\rho}_{\alpha\delta}{}^\times]) \left(\boldsymbol{I} + \left[\boldsymbol{\rho}_{\delta\beta}{}^\times\right]\right) \\ &= \boldsymbol{I} + [\boldsymbol{\rho}_{\alpha\delta}{}^\times] + \left[\boldsymbol{\rho}_{\delta\beta}{}^\times\right] + [\boldsymbol{\rho}_{\alpha\delta}{}^\times] \left[\boldsymbol{\rho}_{\delta\beta}{}^\times\right] \approx \boldsymbol{I} + [\boldsymbol{\rho}_{\alpha\delta}{}^\times] + \left[\boldsymbol{\rho}_{\delta\beta}{}^\times\right]\end{aligned} \tag{2.99}$$

Since the rotation represented by $\boldsymbol{R}_\beta^\alpha$ is a composition of infinitesimal rotations, it must be an infinitesimal rotation itself. Therefore, we can write:

$$\boldsymbol{R}_\beta^\alpha \approx \boldsymbol{I} + \left[\boldsymbol{\rho}_{\alpha\beta}{}^\times\right] \tag{2.100}$$

Comparing (2.100) and (2.99) we see that:

$$\boldsymbol{\rho}_{\alpha\beta} = \boldsymbol{\rho}_{\alpha\delta} + \boldsymbol{\rho}_{\delta\beta}$$

Thus, while rotations in general do not commute, *infinitesimal rotations commute*, and they can be composed through addition of their respective rotation vectors.

For $|\boldsymbol{\rho}| \to 0$, retaining the first order term in (2.87) yields the quaternion equivalents of (2.97) and (2.98):

$$\mathbf{r}_\rho(\boldsymbol{\rho}) = \begin{bmatrix} 1 \\ \boldsymbol{\rho}/2 \end{bmatrix} \tag{2.101}$$
$$\left(\mathbf{r}_\rho(\boldsymbol{\rho})\right)^* = \begin{bmatrix} 1 \\ -\boldsymbol{\rho}/2 \end{bmatrix} \tag{2.102}$$

The linearity of the above approximations presents a significant advantage in many scenarios, and they are often applied in the context of small (rather than strictly infinitesimal) rotations.

2.4 Conversions

2.4.1 Rotation Matrix to Quaternion

From (2.59), the following identities can be obtained:

$$\begin{bmatrix} 1+\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} \\ R^{\alpha}_{\beta(3,2)}-R^{\alpha}_{\beta(2,3)} \\ R^{\alpha}_{\beta(1,3)}-R^{\alpha}_{\beta(3,1)} \\ R^{\alpha}_{\beta(2,1)}-R^{\alpha}_{\beta(1,2)} \end{bmatrix} = 4r^{\alpha}_{\beta(0)}\mathbf{r}^{\alpha}_{\beta} \qquad \begin{bmatrix} R^{\alpha}_{\beta(3,2)}-R^{\alpha}_{\beta(2,3)} \\ 1+2R^{\alpha}_{\beta(1,1)}-\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} \\ R^{\alpha}_{\beta(1,2)}+R^{\alpha}_{\beta(2,1)} \\ R^{\alpha}_{\beta(1,3)}+R^{\alpha}_{\beta(3,1)} \end{bmatrix} = 4r^{\alpha}_{\beta(1)}\mathbf{r}^{\alpha}_{\beta}$$

$$\begin{bmatrix} R^{\alpha}_{\beta(1,3)}-R^{\alpha}_{\beta(3,1)} \\ R^{\alpha}_{\beta(1,2)}+R^{\alpha}_{\beta(2,1)} \\ 1+2R^{\alpha}_{\beta(2,2)}-\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} \\ R^{\alpha}_{\beta(2,3)}+R^{\alpha}_{\beta(3,2)} \end{bmatrix} = 4r^{\alpha}_{\beta(2)}\mathbf{r}^{\alpha}_{\beta} \qquad \begin{bmatrix} R^{\alpha}_{\beta(2,1)}-R^{\alpha}_{\beta(1,2)} \\ R^{\alpha}_{\beta(1,3)}+R^{\alpha}_{\beta(3,1)} \\ R^{\alpha}_{\beta(2,3)}+R^{\alpha}_{\beta(3,2)} \\ 1+2R^{\alpha}_{\beta(3,3)}-\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} \end{bmatrix} = 4r^{\alpha}_{\beta(3)}\mathbf{r}^{\alpha}_{\beta}$$

Given $\boldsymbol{R}^{\alpha}_{\beta}$, $\mathbf{r}^{\alpha}_{\beta}$ is found by computing and normalizing any one of the above quaternions. Numerical errors are minimized by choosing the one with the greatest norm, that is, the one for which $|r^{\alpha}_{\beta(i)}|$ is largest. The best candidate can be determined beforehand from $\boldsymbol{R}^{\alpha}_{\beta}$ as follows:

$$\begin{aligned} \max\left\{\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta}, R^{\alpha}_{\beta(1,1)}, R^{\alpha}_{\beta(2,2)}, R^{\alpha}_{\beta(3,3)}\right\} = \operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} &\implies |r^{\alpha}_{\beta(i)}|_{max} = |r^{\alpha}_{\beta(0)}| \\ \max\left\{\operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta}, R^{\alpha}_{\beta(1,1)}, R^{\alpha}_{\beta(2,2)}, R^{\alpha}_{\beta(3,3)}\right\} = R^{\alpha}_{\beta(i,i)} &\implies |r^{\alpha}_{\beta(i)}|_{max} = |r^{\alpha}_{\beta(i)}|, \forall i \in \{1,2,3\} \end{aligned}$$

The above criterion is easily derived from the following set of equalities:

$$\begin{aligned} \operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} - R^{\alpha}_{\beta(i,i)} &= 2\left({r^{\alpha}_{\beta(0)}}^2 - {r^{\alpha}_{\beta(i)}}^2\right), \forall i \in \{1,2,3\} \\ R^{\alpha}_{\beta(1,1)} - R^{\alpha}_{\beta(2,2)} &= 2\left({r^{\alpha}_{\beta(1)}}^2 - {r^{\alpha}_{\beta(2)}}^2\right) \\ R^{\alpha}_{\beta(1,1)} - R^{\alpha}_{\beta(3,3)} &= 2\left({r^{\alpha}_{\beta(1)}}^2 - {r^{\alpha}_{\beta(2)}}^2\right) \\ R^{\alpha}_{\beta(2,2)} - R^{\alpha}_{\beta(3,3)} &= 2\left({r^{\alpha}_{\beta(2)}}^2 - {r^{\alpha}_{\beta(3)}}^2\right) \end{aligned}$$

2.4.2 Rotation Matrix from Euler Angles

Given $\psi_{\alpha\beta}$, $\theta_{\alpha\beta}$, $\phi_{\alpha\beta}$, $\boldsymbol{R}^{\alpha}_{\beta}$ can be computed directly from (2.92) and (2.93).

2.4.3 Rotation Matrix to Euler Angles

Given $\boldsymbol{R}^{\alpha}_{\beta}$, $\theta_{\alpha\beta}$ is first extracted from (2.93) as:

$$\theta_{\alpha\beta} = \operatorname{atan2}\left(-R^{\alpha}_{\beta(3,1)}, \sqrt{{R^{\alpha}_{\beta(3,2)}}^2 + {R^{\alpha}_{\beta(3,3)}}^2}\right)$$

If $|\theta_{\alpha\beta}| \neq \pi/2$, $\psi_{\alpha\beta}$ and $\phi_{\alpha\beta}$ are given by:

$$\begin{aligned} \psi_{\alpha\beta} &= \operatorname{atan2}\left(R^{\alpha}_{\beta(2,1)}, R^{\alpha}_{\beta(1,1)}\right) \\ \phi_{\alpha\beta} &= \operatorname{atan2}\left(R^{\alpha}_{\beta(3,2)}, R^{\alpha}_{\beta(3,3)}\right) \end{aligned}$$

Otherwise, $\psi_{\alpha\beta}$ and $\phi_{\alpha\beta}$ cannot be independently determined:

- For $\theta = \pi/2$, (2.93) becomes:

$$\boldsymbol{R}_{\Psi}\left(\psi, \pi/2, \phi\right) = \begin{pmatrix} 0 & c_\psi s_\phi - s_\psi c_\phi & c_\psi c_\phi + s_\psi s_\phi \\ 0 & s_\psi s_\phi + c_\psi c_\phi & s_\psi c_\phi - c_\psi s_\phi \\ -1 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & -\sin(\psi - \phi) & \cos(\psi - \phi) \\ 0 & \cos(\psi - \phi) & \sin(\psi - \phi) \\ -1 & 0 & 0 \end{pmatrix}$$

  In this case, $\psi_{\alpha\beta} - \psi_{\alpha\beta}$ can be extracted as:

$$\psi_{\alpha\beta} - \phi_{\alpha\beta} = \operatorname{atan2}\left(R^{\alpha}_{\beta(2,3)} R^{\alpha}_{\beta(2,2)}\right)$$

- For $\theta = -\pi/2$, (2.93) becomes:

$$\boldsymbol{R}_{\Psi}\left(\psi, -\pi/2, \phi\right) = \begin{pmatrix} 0 & -c_\psi s_\phi - s_\psi c_\phi & -c_\psi c_\phi + s_\psi s_\phi \\ 0 & -s_\psi s_\phi + c_\psi c_\phi & -s_\psi c_\phi - c_\psi s_\phi \\ 1 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & -\sin(\psi + \phi) & -\cos(\psi + \phi) \\ 0 & \cos(\psi + \phi) & -\sin(\psi + \phi) \\ 1 & 0 & 0 \end{pmatrix}$$

  In this case, $\psi_{\alpha\beta} + \phi_{\alpha\beta}$ can be extracted as:

$$\psi_{\alpha\beta} + \phi_{\alpha\beta} = \operatorname{atan2}\left(R^{\alpha}_{\beta(1,2)} R^{\alpha}_{\beta(1,3)}\right)$$

2.4.4 Rotation Matrix from Axis-Angle

Given $\boldsymbol{n}_{\alpha\beta}$ and $\vartheta_{\alpha\beta}$, $\boldsymbol{R}^{\alpha}_{\beta}$ can be computed directly from (2.34).

2.4.5 Rotation Matrix to Axis-Angle

First, let us define:

$$\boldsymbol{m}_{\alpha\beta} = 2 \sin \vartheta_{\alpha\beta} \boldsymbol{n}_{\alpha\beta} \tag{2.103}$$

So that:

$$\left|\boldsymbol{m}_{\alpha\beta}\right| = 2 \left|\sin \vartheta_{\alpha\beta}\right| = 2 \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right) \sin \vartheta_{\alpha\beta} \tag{2.104}$$

Turning now to (2.34) and applying property (2.36d), we can derive the following equality:

$$\begin{aligned} \boldsymbol{R}^{\alpha}_{\beta} - \left(\boldsymbol{R}^{\alpha}_{\beta}\right)^T &= \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta}, \vartheta_{\alpha\beta}\right) - \left(\boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta}, \vartheta_{\alpha\beta}\right)\right)^T \\ &= \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta}, \vartheta_{\alpha\beta}\right) - \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta}, -\vartheta_{\alpha\beta}\right) = 2 \sin \vartheta_{\alpha\beta} \left[{\boldsymbol{n}_{\alpha\beta}}^{\times}\right] = \left[{\boldsymbol{m}_{\alpha\beta}}^{\times}\right] \end{aligned}$$

Using (1.61), the above can be written element-wise as:

$$
\begin{aligned}
\left[\boldsymbol{m}_{\alpha\beta}{}^{\times}\right] &= \begin{pmatrix} 0 & -m_{\alpha\beta(3)} & m_{\alpha\beta(2)} \\ m_{\alpha\beta(3)} & 0 & -m_{\alpha\beta(1)} \\ -m_{\alpha\beta(2)} & m_{\alpha\beta(1)} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & R^{\alpha}_{\beta(1,2)} - R^{\alpha}_{\beta(2,1)} & R^{\alpha}_{\beta(1,3)} - R^{\alpha}_{\beta(3,1)} \\ R^{\alpha}_{\beta(2,1)} - R^{\alpha}_{\beta(1,2)} & 0 & R^{\alpha}_{\beta(2,3)} - R^{\alpha}_{\beta(3,2)} \\ R^{\alpha}_{\beta(3,1)} - R^{\alpha}_{\beta(1,3)} & R^{\alpha}_{\beta(3,2)} - R^{\alpha}_{\beta(2,3)} & 0 \end{pmatrix}
\end{aligned}
$$

Therefore, given $\boldsymbol{R}^{\alpha}_{\beta}$, we can compute $\boldsymbol{m}_{\alpha\beta}$ as:

$$
\boldsymbol{m}_{\alpha\beta} = \begin{pmatrix} m_{\alpha\beta(1)} \\ m_{\alpha\beta(2)} \\ m_{\alpha\beta(3)} \end{pmatrix} = \begin{pmatrix} R^{\alpha}_{\beta(3,2)} - R^{\alpha}_{\beta(3,2)} \\ R^{\alpha}_{\beta(1,3)} - R^{\alpha}_{\beta(3,1)} \\ R^{\alpha}_{\beta(2,1)} - R^{\alpha}_{\beta(1,2)} \end{pmatrix} \tag{2.105}
$$

With $\boldsymbol{m}_{\alpha\beta}$, from (2.103) and (2.104):

$$
\begin{aligned}
\vartheta_{\alpha\beta} &= \arcsin\left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\frac{\left|\boldsymbol{m}_{\alpha\beta}\right|}{2}\right) = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\arcsin\left(\frac{\left|\boldsymbol{m}_{\alpha\beta}\right|}{2}\right) \\
\boldsymbol{n}_{\alpha\beta} &= \frac{\boldsymbol{m}_{\alpha\beta}}{2\sin\vartheta_{\alpha\beta}} = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\frac{\boldsymbol{m}_{\alpha\beta}}{\left|\boldsymbol{m}_{\alpha\beta}\right|}
\end{aligned}
$$

Since $\boldsymbol{R}_{n\vartheta}\left(-\boldsymbol{n}, -\vartheta\right) = \boldsymbol{R}_{n\vartheta}\left(\boldsymbol{n}, \vartheta\right)$, $\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)$ is irrelevant and it can be chosen arbitrarily. Let $\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right) = 1$, so that:

$$
2\sin\vartheta_{\alpha\beta} = \left|\boldsymbol{m}_{\alpha\beta}\right| \tag{2.106a}
$$

$$
\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{m}_{\alpha\beta}}{\left|\boldsymbol{m}_{\alpha\beta}\right|} \tag{2.106b}
$$

From (2.37) we also have:

$$
2\cos\vartheta_{\alpha\beta} = \operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta} - 1 \tag{2.107}
$$

With (2.106a) and (2.107), the rotation angle can be computed robustly as:

$$
\vartheta_{\alpha\beta} = \operatorname{atan2}(2\sin\vartheta_{\alpha\beta}, 2\cos\vartheta_{\alpha\beta}) = \operatorname{atan2}\left(\left|\boldsymbol{m}_{\alpha\beta}\right|, \operatorname{tr}\boldsymbol{R}^{\alpha}_{\beta}\right) \tag{2.108}
$$

If $\left|\boldsymbol{m}_{\alpha\beta}\right| \neq 0$, then $\boldsymbol{n}_{\alpha\beta}$ is found from (2.106b). Otherwise it is a null rotation, and the axis is undefined.

2.4.6 Rotation Matrix from Rotation Vector

To compute $\boldsymbol{R}^{\alpha}_{\beta}$ from $\boldsymbol{\rho}_{\alpha\beta}$, we first extract $\vartheta_{\alpha\beta}$ from (2.80b).

If $\left|\boldsymbol{\rho}_{\alpha\beta}\right| > \varepsilon$, $\boldsymbol{n}_{\alpha\beta}$ can be safely computed from (2.80a). $\boldsymbol{R}^{\alpha}_{\beta}$ is then found from (2.34).

If $\left|\boldsymbol{\rho}_{\alpha\beta}\right| < \varepsilon$, to avoid numerical issues, exponential series (2.97) should be used instead, retaining the desired number of terms.

2.4.7 Rotation Matrix to Rotation Vector

To compute $\boldsymbol{\rho}_{\alpha\beta}$ from $\boldsymbol{R}_{\beta}^{\alpha}$, we first extract $\vartheta_{\alpha\beta}$ using (2.108).

If $|\vartheta_{\alpha\beta}| > \varepsilon$, then $\boldsymbol{n}_{\alpha\beta}$ can be safely computed from (2.106b) and $\boldsymbol{R}_{\beta}^{\alpha}$ can then be found from (2.77).

If $|\vartheta_{\alpha\beta}| < \varepsilon$, to avoid numerical issues we should turn to first order approximation (2.97), from which $\boldsymbol{\rho}_{\alpha\beta}$ can be extracted using (1.61):

$$\left[\boldsymbol{\rho}_{\alpha\beta}{}^{\times}\right] = \boldsymbol{R}_{\beta}^{\alpha} - \boldsymbol{I}$$

2.4.8 Quaternion to Rotation Matrix

Given $\mathbf{r}_{\beta}^{\alpha}$, $\boldsymbol{R}_{\beta}^{\alpha}$ can be computed directly from (2.62) and (2.59).

2.4.9 Quaternion from Euler Angles

Given $\psi_{\alpha\beta}$, $\theta_{\alpha\beta}$, $\phi_{\alpha\beta}$, $\mathbf{r}_{\beta}^{\alpha}$ can be computed directly from (2.95) and (2.94).

2.4.10 Quaternion to Euler Angles

The procedure outlined in section 2.4.3 can also be applied to extract $\psi_{\alpha\beta}$, $\theta_{\alpha\beta}$ and $\phi_{\alpha\beta}$ from $\mathbf{r}_{\beta}^{\alpha}$ by using (2.59) to compute the required elements of $\boldsymbol{R}_{\beta}^{\alpha}$.

2.4.11 Quaternion from Axis-Angle

Given $\boldsymbol{n}_{\alpha\beta}$ and $\vartheta_{\alpha\beta}$, $\mathbf{r}_{\beta}^{\alpha}$ can be computed directly from (2.61).

2.4.12 Quaternion to Axis-Angle

From (2.61) we have:

$$\left|\boldsymbol{r}_{\beta}^{\alpha}\right| = \left|\sin(\vartheta_{\alpha\beta}/2)\boldsymbol{n}_{\alpha\beta}\right| = \left|\sin(\vartheta_{\alpha\beta}/2)\right|\left|\boldsymbol{n}_{\alpha\beta}\right| = \left|\sin(\vartheta_{\alpha\beta}/2)\right| = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\sin(\vartheta_{\alpha\beta}/2)$$

Therefore:

$$\begin{aligned}\vartheta_{\alpha\beta} &= 2\arcsin\left(\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\left|\boldsymbol{r}_{\beta}^{\alpha}\right|\right) = 2\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\arcsin\left(\left|\boldsymbol{r}_{\beta}^{\alpha}\right|\right) \\ \boldsymbol{n}_{\alpha\beta} &= \frac{\boldsymbol{r}_{\beta}^{\alpha}}{\sin(\vartheta_{\alpha\beta}/2)} = \operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)\frac{\boldsymbol{r}_{\beta}^{\alpha}}{\left|\boldsymbol{r}_{\beta}^{\alpha}\right|}\end{aligned}$$

Since $\mathbf{r}_{n\vartheta}\left(-\boldsymbol{n}_{\alpha\beta}, -\vartheta_{\alpha\beta}\right) = \mathbf{r}_{n\vartheta}\left(\boldsymbol{n}_{\alpha\beta}, \vartheta_{\alpha\beta}\right)$, $\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right)$ is irrelevant and can be chosen arbitrarily. Let $\operatorname{sgn}\left(\vartheta_{\alpha\beta}\right) = 1$, so that:

$$\sin(\vartheta_{\alpha\beta}/2) = \left|\boldsymbol{r}_{\beta}^{\alpha}\right| \tag{2.109a}$$

$$\boldsymbol{n}_{\alpha\beta} = \frac{\boldsymbol{r}_{\beta}^{\alpha}}{\left|\boldsymbol{r}_{\beta}^{\alpha}\right|} \tag{2.109b}$$

The rotation angle can always be computed robustly as:

$$\vartheta_{\alpha\beta} = 2\,\mathrm{atan2}(\sin(\vartheta_{\alpha\beta}/2), \cos(\vartheta_{\alpha\beta}/2)) = 2\,\mathrm{atan2}\left(\left|\boldsymbol{r}^{\alpha}_{\beta}\right|, r^{\alpha}_{\beta(0)}\right) \tag{2.110}$$

If $\left|\boldsymbol{r}^{\alpha}_{\beta}\right| \neq 0$, then $\boldsymbol{n}_{\alpha\beta}$ is found from (2.109b). Otherwise it is a null rotation, and the axis is undefined.

2.4.13 Quaternion from Rotation Vector

Given $\boldsymbol{\rho}_{\alpha\beta}$, $\mathbf{r}^{\alpha}_{\beta}$ may be found either from (2.86) or (2.87).

If $|\boldsymbol{\rho}_{\alpha\beta}| > \varepsilon$, where ε is some (very small) threshold, then $\mathbf{r}^{\alpha}_{\beta}$ can be computed exactly from (2.86).

If $|\boldsymbol{\rho}_{\alpha\beta}| < \varepsilon$, to avoid numerical issues, alternative parameterization (2.87) should be used instead, retaining the desired number of terms.

2.4.14 Quaternion to Rotation Vector

To compute $\boldsymbol{\rho}_{\alpha\beta}$ from $\mathbf{r}^{\alpha}_{\beta}$, first extract $\vartheta_{\alpha\beta}$ from (2.110).

If $|\vartheta_{\alpha\beta}| > \varepsilon$, then $\boldsymbol{n}_{\alpha\beta}$ can be safely extracted from (2.109b) and $\boldsymbol{\rho}_{\alpha\beta}$ is found directly from (2.77).

If $|\vartheta_{\alpha\beta}| > \varepsilon$, to avoid numerical issues, we can proceed as follows. From (2.77), (2.109a) and (2.109b):

$$\boldsymbol{\rho}_{\alpha\beta} = 2\arcsin\left|\boldsymbol{r}^{\alpha}_{\beta}\right| \frac{\boldsymbol{r}^{\alpha}_{\beta}}{\left|\boldsymbol{r}^{\alpha}_{\beta}\right|}$$

Substituting $\arcsin\left|\boldsymbol{r}^{\alpha}_{\beta}\right|$ with its Taylor series yields, after some manipulation:

$$\boldsymbol{\rho}_{\alpha\beta} = 2\boldsymbol{r}^{\alpha}_{\beta} \sum_{k=0}^{\infty} \frac{(2k)!\left|\boldsymbol{r}^{\alpha}_{\beta}\right|^{2k}}{4^k (k!)^2 (2k+1)} \tag{2.111}$$

From which $\boldsymbol{\rho}_{\alpha\beta}$ can be computed by retaining the desired number of terms.

Chapter 3

Rotational Motion

3.1 Angular Velocity

3.1.1 Definition

Let us consider two orthonormal bases $\mathcal{E}_\alpha$ and $\mathcal{E}_\beta$ in rotational motion with respect to each other, so that rotation matrix $\boldsymbol{R}_\beta^\alpha$ is a function of time.

The orthogonality condition for $\boldsymbol{R}_\beta^\alpha$ can be written:

$$\left(\boldsymbol{R}_\beta^\alpha\right)^T \boldsymbol{R}_\beta^\alpha = \boldsymbol{I} \tag{3.1}$$

Transposing (3.1) yields the equivalent:

$$\boldsymbol{R}_\beta^\alpha \left(\boldsymbol{R}_\beta^\alpha\right)^T = \boldsymbol{I} \tag{3.2}$$

Taking the time derivative in these two expressions and using properties (1.56a) and (1.56b) gives:

$$\left(\dot{\boldsymbol{R}}_\beta^\alpha\right)^T \boldsymbol{R}_\beta^\alpha + \left(\boldsymbol{R}_\beta^\alpha\right)^T \dot{\boldsymbol{R}}_\beta^\alpha = \boldsymbol{0} \tag{3.3}$$

$$\dot{\boldsymbol{R}}_\beta^\alpha \left(\boldsymbol{R}_\beta^\alpha\right)^T + \boldsymbol{R}_\beta^\alpha \left(\dot{\boldsymbol{R}}_\beta^\alpha\right)^T = \boldsymbol{0} \tag{3.4}$$

Where we have introduced the shorthand notation:

$$\dot{\boldsymbol{R}}_\beta^\alpha = \frac{d\boldsymbol{R}_\beta^\alpha}{dt}$$

Using (1.39a) and (1.39c), (3.3) and (3.4) can be written as:

$$\left(\left(\boldsymbol{R}_\beta^\alpha\right)^T \dot{\boldsymbol{R}}_\beta^\alpha\right)^T + \left(\boldsymbol{R}_\beta^\alpha\right)^T \dot{\boldsymbol{R}}_\beta^\alpha = \boldsymbol{X}^T + \boldsymbol{X} = \boldsymbol{0} \tag{3.5}$$

$$\dot{\boldsymbol{R}}_\beta^\alpha \left(\boldsymbol{R}_\beta^\alpha\right)^T + \left(\dot{\boldsymbol{R}}_\beta^\alpha \left(\boldsymbol{R}_\beta^\alpha\right)^T\right)^T = \boldsymbol{Y} + \boldsymbol{Y}^T = \boldsymbol{0} \tag{3.6}$$

Where the following matrices have been defined for convenience:

$$\boldsymbol{X} = \left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \dot{\boldsymbol{R}}_{\beta}^{\alpha}$$
$$\boldsymbol{Y} = \dot{\boldsymbol{R}}_{\beta}^{\alpha}\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T}$$

Expressions (3.5) and (3.6) show that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are both skew-symmetric. Therefore, they can be interpreted as cross-product matrices, and from (1.61) we know there exist two column matrices $\boldsymbol{x}$ and $\boldsymbol{y}$ such that:

$$[\boldsymbol{x}^{\times}] = \boldsymbol{X} = \left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \dot{\boldsymbol{R}}_{\beta}^{\alpha} \tag{3.7}$$
$$[\boldsymbol{y}^{\times}] = \boldsymbol{Y} = \dot{\boldsymbol{R}}_{\beta}^{\alpha}\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \tag{3.8}$$

Solving for $\dot{\boldsymbol{R}}_{\beta}^{\alpha}$ in (3.7) and substituting into (3.8):

$$[\boldsymbol{y}^{\times}] = \boldsymbol{R}_{\beta}^{\alpha}\,[\boldsymbol{x}^{\times}]\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \tag{3.9}$$

Now, turning to property (1.65) and setting $\boldsymbol{A} = \boldsymbol{R}_{\beta}^{\alpha}$, we obtain the following equality:

$$\left[\left(\boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{x}\right)^{\times}\right] = \boldsymbol{R}_{\beta}^{\alpha}\,[\boldsymbol{x}^{\times}]\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \tag{3.10}$$

Comparing (3.9) and (3.10) we see that the relation between $\boldsymbol{x}$ and $\boldsymbol{y}$ is:

$$\boldsymbol{y} = \boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{x}$$

In its passive interpretation, this transformation represents a coordinate change from basis $\mathcal{E}_{\beta}$ to basis $\mathcal{E}_{\alpha}$ of some time-varying Euclidean vector whose $\mathcal{E}_{\beta}$ and $\mathcal{E}_{\alpha}$ components are given respectively by column matrices $\boldsymbol{x}$ and $\boldsymbol{y}$, which are defined implicitly through (3.7) and (3.8).

Letting $\vec{\boldsymbol{\omega}}_{\alpha\beta}$ denote this Euclidean vector, we have:

$$\left[{\boldsymbol{\omega}_{\alpha\beta}^{\beta}}^{\times}\right] = \left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \dot{\boldsymbol{R}}_{\beta}^{\alpha} \tag{3.11}$$
$$\left[{\boldsymbol{\omega}_{\alpha\beta}^{\alpha}}^{\times}\right] = \dot{\boldsymbol{R}}_{\beta}^{\alpha}\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \tag{3.12}$$

Vector $\vec{\boldsymbol{\omega}}_{\alpha\beta}$ is the *angular velocity* of $\mathcal{E}_{\beta}$ with respect to $\mathcal{E}_{\alpha}$.

We now introduce the following notation:

$$\boldsymbol{\Omega}_{\alpha\beta}^{\beta} = \left[{\boldsymbol{\omega}_{\alpha\beta}^{\beta}}^{\times}\right] \tag{3.13}$$
$$\boldsymbol{\Omega}_{\alpha\beta}^{\alpha} = \left[{\boldsymbol{\omega}_{\alpha\beta}^{\alpha}}^{\times}\right] \tag{3.14}$$

With this, (3.11) and (3.12) become:

$$\dot{\boldsymbol{R}}_{\beta}^{\alpha} = \boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{\Omega}_{\alpha\beta}^{\beta} \tag{3.15}$$

$$\dot{\boldsymbol{R}}_{\beta}^{\alpha} = \boldsymbol{\Omega}_{\alpha\beta}^{\alpha}\boldsymbol{R}_{\beta}^{\alpha} \tag{3.16}$$

We also have, from (3.9):

$$\boldsymbol{\Omega}_{\alpha\beta}^{\alpha} = \boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{\Omega}_{\alpha\beta}^{\beta}\left(\boldsymbol{R}_{\beta}^{\alpha}\right)^{T} \tag{3.17}$$

Taking the time derivative of (2.7) and applying (3.15) yields the following relation between the time derivatives of the $\mathcal{E}_{\beta}$ and $\mathcal{E}_{\alpha}$ components of an arbitrary Euclidean vector:

$$\begin{aligned}\dot{\boldsymbol{x}}^{\alpha} &= \boldsymbol{R}_{\beta}^{\alpha}\dot{\boldsymbol{x}}^{\beta} + \dot{\boldsymbol{R}}_{\beta}^{\alpha}\boldsymbol{x}^{\beta} \\ &= \boldsymbol{R}_{\beta}^{\alpha}\dot{\boldsymbol{x}}^{\beta} + \boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{\Omega}_{\alpha\beta}^{\beta}\boldsymbol{x}^{\beta} \\ &= \boldsymbol{R}_{\beta}^{\alpha}\left(\dot{\boldsymbol{x}}^{\beta} + \boldsymbol{\Omega}_{\alpha\beta}^{\beta}\boldsymbol{x}^{\beta}\right)\end{aligned} \tag{3.18}$$

$$\dot{\boldsymbol{x}}^{\alpha} = \boldsymbol{R}_{\beta}^{\alpha}\dot{\boldsymbol{x}}^{\beta} + \dot{\boldsymbol{R}}_{\beta}^{\alpha}\boldsymbol{x}^{\beta} = \boldsymbol{R}_{\beta}^{\alpha}\dot{\boldsymbol{x}}^{\beta} + \boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{\Omega}_{\alpha\beta}^{\beta}\boldsymbol{x}^{\beta} = \boldsymbol{R}_{\beta}^{\alpha}\left(\dot{\boldsymbol{x}}^{\beta} + \boldsymbol{\Omega}_{\alpha\beta}^{\beta}\boldsymbol{x}^{\beta}\right)$$

This result is sometimes known as *Coriolis' Theorem*. It shows that, in general, $\dot{\boldsymbol{x}}^{\beta}$ is not simply a rotation of $\dot{\boldsymbol{x}}^{\alpha}$. The additional term due to the angular velocity is called the *transport rate*.

3.1.2 Interpretation

The previous definition of the angular velocity vector provides little insight about its physical interpretation. To uncover it, we shall take a different approach.

We start from the formal definition of $\dot{\boldsymbol{R}}_{\beta}^{\alpha}$:

$$\dot{\boldsymbol{R}}_{\beta}^{\alpha} = \lim_{\Delta t \to 0} \frac{\boldsymbol{R}_{\beta}^{\alpha}(t+\Delta t) - \boldsymbol{R}_{\beta}^{\alpha}(t)}{\Delta t}$$

This can be rewritten as:

$$\begin{aligned}\dot{\boldsymbol{R}}_{\beta}^{\alpha} &= \boldsymbol{R}_{\beta}^{\alpha}(t)\left(\boldsymbol{R}_{\beta}^{\alpha}(t)\right)^{T} \lim_{\Delta t \to 0} \frac{\boldsymbol{R}_{\beta}^{\alpha}(t+\Delta t) - \boldsymbol{R}_{\beta}^{\alpha}(t)}{\Delta t} \\ &= \boldsymbol{R}_{\beta}^{\alpha}(t) \lim_{\Delta t \to 0} \frac{\left(\boldsymbol{R}_{\beta}^{\alpha}(t)\right)^{T}\boldsymbol{R}_{\beta}^{\alpha}(t+\Delta t) - \boldsymbol{I}}{\Delta t}\end{aligned} \tag{3.19}$$

Being a product of rotation matrices, the first term in the numerator of (3.19) must also be a rotation matrix, but its meaning is not immediately clear.

To make sense of it, consider a Euclidean vector $\vec{\boldsymbol{x}}$ with constant $\mathcal{E}_{\alpha}$ components. The $\mathcal{E}_{\beta}$ components of $\vec{\boldsymbol{x}}$ at times t and $t+\Delta t$ are given respectively by:

$$\boldsymbol{x}^{\beta}(t) = \left(\boldsymbol{R}_{\beta}^{\alpha}(t)\right)^{T}\boldsymbol{x}^{\alpha} \tag{3.20a}$$

$$\boldsymbol{x}^{\beta}(t+\Delta t) = \left(\boldsymbol{R}_{\beta}^{\alpha}(t+\Delta t)\right)^{T}\boldsymbol{x}^{\alpha} \tag{3.20b}$$

Solving for $\boldsymbol{x}^\alpha$ in (3.20b) and substituting in (3.20a) yields:

$$\boldsymbol{x}^\beta(t) = \left(\boldsymbol{R}^\alpha_\beta(t)\right)^T \boldsymbol{R}^\alpha_\beta(t+\Delta t)\boldsymbol{x}^\beta(t+\Delta t)$$

This shows that the rotation matrix in the numerator of (3.19) relates the $\mathcal{E}_\beta$ components of $\vec{\boldsymbol{x}}$ at times t and $t+\Delta t$. Thus, abusing our notation somewhat, we may represent it as $\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)}$. With this notational convention, (3.19) becomes:

$$\dot{\boldsymbol{R}}^\alpha_\beta = \boldsymbol{R}^\alpha_\beta(t) \lim_{\Delta t\to 0} \frac{1}{\Delta t}\left(\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)} - \boldsymbol{I}\right) \tag{3.21}$$

Note that this interpretation of $\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)}$ requires that the $\mathcal{E}_\alpha$ components of $\vec{\boldsymbol{x}}$ be constant, and thus it is only meaningful in the context of the rotational motion of $\mathcal{E}_\beta$ *with respect to* $\mathcal{E}_\alpha$. This essential detail is not made explicit by the notation.

From Euler's theorem, we know that the rotation represented by $\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)}$ can also be described by a three-component rotation vector, which we will denote by $\vec{\boldsymbol{\rho}}_{\beta(t)\beta(t+\Delta t)}$. The relation between $\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)}$ and $\vec{\boldsymbol{\rho}}_{\beta(t)\beta(t+\Delta t)}$ is given by the exponential map (2.84):

$$\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)} = \boldsymbol{R}_\rho\left(\boldsymbol{\rho}_{\beta(t)\beta(t+\Delta t)}\right) \tag{3.22}$$

Substituting in (3.21):

$$\dot{\boldsymbol{R}}^\alpha_\beta = \boldsymbol{R}^\alpha_\beta(t) \lim_{\Delta t\to 0} \frac{1}{\Delta t}\left(\boldsymbol{R}_\rho\left(\boldsymbol{\rho}_{\beta(t)\beta(t+\Delta t)}\right) - \boldsymbol{I}\right)$$

In the limit, the rotation represented by $\boldsymbol{R}^{\beta(t)}_{\beta(t+\Delta t)}$ becomes infinitesimal, so (2.97) can be applied to yield:

$$\dot{\boldsymbol{R}}^\alpha_\beta = \boldsymbol{R}^\alpha_\beta(t) \lim_{\Delta t\to 0} \left[\left(\frac{\boldsymbol{\rho}_{\beta(t)\beta(t+\Delta t)}}{\Delta t}\right)^\times\right] \tag{3.23}$$

Recall from (2.77) that the missing superscript in the rotation vector conveys the fact that it can be expressed indistinctly in any of the two bases involved in the rotation, in this case $\mathcal{E}_{\beta(t)}$ and $\mathcal{E}_{\beta(t+\Delta t)}$. Thus, we may rewrite (3.23) as:

$$\dot{\boldsymbol{R}}^\alpha_\beta = \boldsymbol{R}^\alpha_\beta(t) \lim_{\Delta t\to 0} \left[\left(\frac{\boldsymbol{\rho}^{\beta(t)}_{\beta(t)\beta(t+\Delta t)}}{\Delta t}\right)^\times\right] \tag{3.24}$$

Comparing (3.24) and (3.15) shows that:

$$\boldsymbol{\omega}^\beta_{\alpha\beta} = \lim_{\Delta t\to 0} \frac{\boldsymbol{\rho}^{\beta(t)}_{\beta(t)\beta(t+\Delta t)}}{\Delta t}$$

More generally:

$$\vec{\boldsymbol{\omega}}_{\alpha\beta} = \lim_{\Delta t\to 0} \frac{\vec{\boldsymbol{\rho}}_{\beta(t)\beta(t+\Delta t)}}{\Delta t} \tag{3.25}$$

This expression, together with our previous definition of the incremental rotation vector $\vec{\boldsymbol{\rho}}_{\beta(t)\beta(t+\Delta t)}$, embodies a more physically meaningful interpretation of the angular velocity vector than (3.11) and (3.12).

Note that, despite the resemblance of (3.25) to a derivative, it cannot be written as such, because there is no absolute rotation vector of which $\vec{\boldsymbol{\rho}}_{\beta(t)\beta(t+\Delta t)}$ is an increment.

3.1.3 Composition

We now write $\boldsymbol{R}_{\beta}^{\alpha}$ in terms of an intermediate basis γ:

$$\boldsymbol{R}_{\beta}^{\alpha} = \boldsymbol{R}_{\gamma}^{\alpha}\boldsymbol{R}_{\beta}^{\gamma}$$

Taking the time derivative and applying (1.56b):

$$\dot{\boldsymbol{R}}_{\beta}^{\alpha} = \dot{\boldsymbol{R}}_{\gamma}^{\alpha}\boldsymbol{R}_{\beta}^{\gamma} + \boldsymbol{R}_{\gamma}^{\alpha}\dot{\boldsymbol{R}}_{\beta}^{\gamma} \tag{3.26}$$

Using (3.15) and (3.16), and recalling the notation from (3.13):

$$\begin{aligned}
\boldsymbol{R}_{\beta}^{\alpha}\boldsymbol{\Omega}_{\alpha\beta}^{\beta} &= \boldsymbol{R}_{\gamma}^{\alpha}\left(\boldsymbol{\Omega}_{\alpha\gamma}^{\gamma} + \boldsymbol{\Omega}_{\gamma\beta}^{\gamma}\right)\boldsymbol{R}_{\beta}^{\gamma} \\
\boldsymbol{\Omega}_{\alpha\beta}^{\beta} &= \boldsymbol{\Omega}_{\alpha\gamma}^{\beta} + \boldsymbol{\Omega}_{\gamma\beta}^{\beta} \\
\boldsymbol{\omega}_{\alpha\beta}^{\beta} &= \boldsymbol{\omega}_{\alpha\gamma}^{\beta} + \boldsymbol{\omega}_{\gamma\beta}^{\beta}
\end{aligned} \tag{3.27}$$

Premultiplying (3.27) by $\boldsymbol{R}_{\beta}^{\delta}$, where $\mathcal{E}_{\delta}$ is an arbitrary orthonormal basis:

$$\boldsymbol{\omega}_{\alpha\beta}^{\delta} = \boldsymbol{\omega}_{\alpha\gamma}^{\delta} + \boldsymbol{\omega}_{\gamma\beta}^{\delta} \tag{3.28}$$

Since (3.28) holds for an arbitrary $\mathcal{E}_{\delta}$, it follows that composition of angular velocities corresponds to simple Euclidean vector addition:

$$\vec{\boldsymbol{\omega}}_{\alpha\beta} = \vec{\boldsymbol{\omega}}_{\alpha\gamma} + \vec{\boldsymbol{\omega}}_{\gamma\beta} \tag{3.29}$$

This result is consistent with the interpretation of angular velocity as an infinitesimal rotation vector per unit time, embodied by (3.25): as shown in section 2.3, infinitesimal rotations commute and they compose through rotation vector addition.

Setting $\gamma = \alpha$ in (3.29) and noting that $\vec{\boldsymbol{\omega}}_{\alpha\alpha} = \vec{\mathbf{0}}$ yields the relation between reciprocal angular velocities:

$$\vec{\boldsymbol{\omega}}_{\alpha\beta} = -\vec{\boldsymbol{\omega}}_{\beta\alpha} \tag{3.30}$$

3.1.4 Rotation Around a Fixed Axis

In general, for two bases $\mathcal{E}_{\alpha}$ and $\mathcal{E}_{\beta}$ in rotational motion with respect to each other, both the rotation angle $\vartheta_{\alpha\beta}$ and the rotation axis, defined by unit vector $\vec{\boldsymbol{n}}_{\alpha\beta}$, change over time. In the particular case where the $\mathcal{E}_{\alpha}$ and $\mathcal{E}_{\beta}$ components of $\vec{\boldsymbol{n}}_{\alpha\beta}$ (which we know from (2.11) to be equal and therefore we denote simply by $\boldsymbol{n}_{\alpha\beta}$) remain constant, the rotational motion occurs around a fixed axis. As we now show, in this case the angular velocity vector takes on a particularly simple form.

We start by transposing (2.34) and using (1.39b), (1.39c) and (1.63d) to yield:

$$\begin{aligned}\left(\boldsymbol{R}_\beta^\alpha\right)^T &= \boldsymbol{I} + \left(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\right)^T \sin\vartheta_{\alpha\beta} + \left(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\right)^T \left(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\right)^T (1-\cos\vartheta_{\alpha\beta}) \\ &= \boldsymbol{I} - \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\sin\vartheta_{\alpha\beta} + \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]^2 (1-\cos\vartheta_{\alpha\beta})\end{aligned} \tag{3.31}$$

Then we take the time derivative of (2.34) with constant $\boldsymbol{n}_{\alpha\beta}$:

$$\dot{\boldsymbol{R}}_\beta^\alpha = \dot{\vartheta}_{\alpha\beta}\left(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\cos\vartheta_{\alpha\beta} + \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]^2 \sin\vartheta_{\alpha\beta}\right) \tag{3.32}$$

Solving for $\boldsymbol{\Omega}_{\alpha\beta}^\beta$ in (3.15) and inserting (3.31) and (3.32):

$$\begin{aligned}\boldsymbol{\Omega}_{\alpha\beta}^\beta = \left(\boldsymbol{R}_\beta^\alpha\right)^T \dot{\boldsymbol{R}}_\beta^\alpha = \dot{\vartheta}_{\alpha\beta}\Big(\boldsymbol{I} - \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\sin\vartheta_{\alpha\beta} + \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]^2 (1-\cos\vartheta_{\alpha\beta})\Big) \\ \Big(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\cos\vartheta_{\alpha\beta} + \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]^2 \sin\vartheta_{\alpha\beta}\Big)\end{aligned}$$

Multiplying out and applying (2.82a) and (2.82b) yields, after a bit of algebra:

$$\boldsymbol{\Omega}_{\alpha\beta}^\beta = \dot{\vartheta}_{\alpha\beta}\left(\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\cos^2\vartheta_{\alpha\beta} + \left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]\sin^2\vartheta_{\alpha\beta}\right) = \dot{\vartheta}_{\alpha\beta}\left[\boldsymbol{n}_{\alpha\beta}{}^\times\right]$$

Thus, the angular velocity is simply:

$$\boldsymbol{\omega}_{\alpha\beta}^\beta = \boldsymbol{\omega}_{\alpha\beta}^\alpha = \boldsymbol{\omega}_{\alpha\beta} = \dot{\vartheta}_{\alpha\beta}\boldsymbol{n}_{\alpha\beta} \tag{3.33}$$

And from (2.77) we have:

$$\dot{\boldsymbol{\rho}}_{\alpha\beta} = \dot{\vartheta}_{\alpha\beta}\boldsymbol{n}_{\alpha\beta} = \boldsymbol{\omega}_{\alpha\beta} \tag{3.34}$$

Note that the constancy of $\boldsymbol{n}_{\alpha\beta}$ does not translate in general to the components of $\vec{\boldsymbol{n}}_{\alpha\beta}$ in a third basis $\mathcal{E}_\gamma$. Indeed, unless $\mathcal{E}_\gamma$ is fixed to either $\mathcal{E}_\alpha$ or $\mathcal{E}_\beta$, $\boldsymbol{n}_{\alpha\beta}^\gamma$ will in fact be time-varying.

3.2 Angular Acceleration

3.2.1 Definition

We define the angular acceleration of $\mathcal{E}_\beta$ with respect to $\mathcal{E}_\alpha$ as:

$$\boldsymbol{\alpha}_{\alpha\beta}^\beta = \dot{\boldsymbol{\omega}}_{\alpha\beta}^\beta \tag{3.35}$$

Its $\mathcal{E}_\alpha$ components are given by:

$$\boldsymbol{\alpha}_{\alpha\beta}^\alpha = \boldsymbol{R}_\beta^\alpha\boldsymbol{\alpha}_{\alpha\beta}^\beta = \boldsymbol{R}_\beta^\alpha\dot{\boldsymbol{\omega}}_{\alpha\beta}^\beta \tag{3.36}$$

3.2.2 Composition

Taking the time derivative in (3.27):

$$\begin{aligned}\dot{\boldsymbol{\omega}}^{\beta}_{\alpha\beta} &= \dot{\boldsymbol{\omega}}^{\beta}_{\alpha\gamma} + \dot{\boldsymbol{\omega}}^{\beta}_{\gamma\beta} \\ \boldsymbol{\alpha}^{\beta}_{\alpha\beta} &= \dot{\boldsymbol{\omega}}^{\beta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\beta}_{\gamma\beta}\end{aligned} \tag{3.37}$$

Applying (3.18) to $\dot{\boldsymbol{\omega}}^{\beta}_{\alpha\gamma}$:

$$\dot{\boldsymbol{\omega}}^{\beta}_{\alpha\gamma} = \boldsymbol{R}^{\beta}_{\gamma}\left(\dot{\boldsymbol{\omega}}^{\gamma}_{\alpha\gamma} + \boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\boldsymbol{\omega}^{\gamma}_{\alpha\gamma}\right) = \boldsymbol{R}^{\beta}_{\gamma}\left(\boldsymbol{\alpha}^{\gamma}_{\alpha\gamma} + \boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\boldsymbol{\omega}^{\gamma}_{\alpha\gamma}\right) = \boldsymbol{\alpha}^{\beta}_{\alpha\gamma} + \boldsymbol{R}^{\beta}_{\gamma}\boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\boldsymbol{\omega}^{\gamma}_{\alpha\gamma} \tag{3.38}$$

Substituting (3.38) into (3.37):

$$\boldsymbol{\alpha}^{\beta}_{\alpha\beta} = \boldsymbol{\alpha}^{\beta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\beta}_{\gamma\beta} + \boldsymbol{R}^{\beta}_{\gamma}\boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\boldsymbol{\omega}^{\gamma}_{\alpha\gamma} \tag{3.39}$$

Premultiplying (3.39) by $\boldsymbol{R}^{\delta}_{\beta}$ and using property (3.10):

$$\begin{aligned}\boldsymbol{\alpha}^{\delta}_{\alpha\beta} &= \boldsymbol{\alpha}^{\delta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\delta}_{\gamma\beta} + \boldsymbol{R}^{\delta}_{\gamma}\boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\boldsymbol{\omega}^{\gamma}_{\alpha\gamma} \\ &= \boldsymbol{\alpha}^{\delta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\delta}_{\gamma\beta} + \boldsymbol{R}^{\delta}_{\gamma}\boldsymbol{\Omega}^{\gamma}_{\beta\gamma}\left(\boldsymbol{R}^{\delta}_{\gamma}\right)^{T}\boldsymbol{\omega}^{\delta}_{\alpha\gamma} = \boldsymbol{\alpha}^{\delta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\delta}_{\gamma\beta} + \boldsymbol{\Omega}^{\delta}_{\beta\gamma}\boldsymbol{\omega}^{\delta}_{\alpha\gamma}\end{aligned}$$

Applying (3.30) and (1.63c) finally yields:

$$\boldsymbol{\alpha}^{\delta}_{\alpha\beta} = \boldsymbol{\alpha}^{\delta}_{\alpha\gamma} + \boldsymbol{\alpha}^{\delta}_{\gamma\beta} + \boldsymbol{\Omega}^{\delta}_{\alpha\gamma}\boldsymbol{\omega}^{\delta}_{\gamma\beta} \tag{3.40}$$

Since this holds for an arbitrary basis $\mathcal{E}_{\delta}$, we can conclude:

$$\vec{\boldsymbol{\alpha}}_{\alpha\beta} = \vec{\boldsymbol{\alpha}}_{\alpha\gamma} + \vec{\boldsymbol{\alpha}}_{\gamma\beta} + \vec{\boldsymbol{\omega}}_{\alpha\gamma} \times \vec{\boldsymbol{\omega}}_{\gamma\beta} \tag{3.41}$$

3.3 Quaternion Kinematics

Let $\mathbf{r}$ be an arbitrary unit quaternion, so that:

$$\mathbf{r}^{*} \odot \mathbf{r} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$$

Taking the time derivative in this equality and applying properties (2.53) and (2.52):

$$\mathbf{r}^{*} \odot \dot{\mathbf{r}} + \dot{\mathbf{r}}^{*} \odot \mathbf{r} = \begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix} \tag{3.42}$$

Explicit computation of the quaternion products above yields:

$$\mathbf{r}^{*} \odot \dot{\mathbf{r}} = \begin{bmatrix} r_{(0)} \\ -\boldsymbol{r} \end{bmatrix} \odot \begin{bmatrix} \dot{r}_{(0)} \\ \dot{\boldsymbol{r}} \end{bmatrix} = \begin{bmatrix} r_{(0)}\dot{r}_{(0)} + \boldsymbol{r}\cdot\dot{\boldsymbol{r}} \\ r_{(0)}\dot{\boldsymbol{r}} - \dot{r}_{(0)}\boldsymbol{r} - \boldsymbol{r}\times\dot{\boldsymbol{r}} \end{bmatrix} = \begin{bmatrix} r_{(0)}\dot{r}_{(0)} + \boldsymbol{r}\cdot\dot{\boldsymbol{r}} \\ \frac{1}{2}\boldsymbol{\omega}(\mathbf{r}) \end{bmatrix} \tag{3.43a}$$

$$\dot{\mathbf{r}}^{*} \odot \mathbf{r} = \begin{bmatrix} \dot{r}_{(0)} \\ -\dot{\boldsymbol{r}} \end{bmatrix} \odot \begin{bmatrix} r_{(0)} \\ \boldsymbol{r} \end{bmatrix} = \begin{bmatrix} \dot{r}_{(0)}r_{(0)} + \dot{\boldsymbol{r}}\cdot\boldsymbol{r} \\ \dot{r}_{(0)}\boldsymbol{r} - r_{(0)}\dot{\boldsymbol{r}} - \dot{\boldsymbol{r}}\times\boldsymbol{r} \end{bmatrix} = \begin{bmatrix} \dot{r}_{(0)}r_{(0)} + \dot{\boldsymbol{r}}\cdot\boldsymbol{r} \\ -\frac{1}{2}\boldsymbol{\omega}(\mathbf{r}) \end{bmatrix} \tag{3.43b}$$

Where operator $\boldsymbol{\omega}(\mathbf{r})$, which takes a unit quaternion as an input and produces a column matrix, is defined as:

$$\boldsymbol{\omega}(\mathbf{r}) = 2\left(r_{(0)}\dot{\boldsymbol{r}} - \dot{r}_{(0)}\boldsymbol{r} - \boldsymbol{r} \times \dot{\boldsymbol{r}}\right) \tag{3.44}$$

Inserting (3.43a) and (3.43b) into (3.42) gives:

$$\dot{r}_{(0)}r_{(0)} + \dot{\boldsymbol{r}} \cdot \boldsymbol{r} = 0$$

So (3.43a) and (3.43b) become:

$$\mathbf{r}^* \odot \dot{\mathbf{r}} = \frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{\omega}(\mathbf{r}) \end{bmatrix} \tag{3.45a}$$

$$\dot{\mathbf{r}}^* \odot \mathbf{r} = -\frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{\omega}(\mathbf{r}) \end{bmatrix} \tag{3.45b}$$

Now let us consider an orthonormal basis $\mathcal{E}_\beta$ in rotational motion with respect to another $\mathcal{E}_\alpha$, with their relative orientation described by $\mathbf{r}_\beta^\alpha$. Taking the time derivative in (2.68) and applying properties (2.44a), (2.53) and (2.52) yields:

$$\dot{\mathbf{x}}^\alpha = \mathbf{r}_\beta^\alpha \odot \dot{\mathbf{x}}^\beta \odot \left(\mathbf{r}_\beta^\alpha\right)^* + \dot{\mathbf{r}}_\beta^\alpha \odot \mathbf{x}^\beta \odot \left(\mathbf{r}_\beta^\alpha\right)^* + \mathbf{r}_\beta^\alpha \odot \mathbf{x}^\beta \odot \left(\dot{\mathbf{r}}_\beta^\alpha\right)^*$$

Factoring out $\mathbf{r}_\beta^\alpha$ and $\left(\mathbf{r}_\beta^\alpha\right)^*$, and noting that $\mathbf{r}_\beta^\alpha$ is a unit quaternion, the above expression can be rewritten as follows:

$$\dot{\mathbf{x}}^\alpha = \mathbf{r}_\beta^\alpha \odot \left(\dot{\mathbf{x}}^\beta + \left(\mathbf{r}_\beta^\alpha\right)^* \odot \dot{\mathbf{r}}_\beta^\alpha \odot \mathbf{x}^\beta + \mathbf{x}^\beta \odot \left(\dot{\mathbf{r}}_\beta^\alpha\right)^* \odot \mathbf{r}_\beta^\alpha\right) \odot \left(\mathbf{r}_\beta^\alpha\right)^*$$

Inserting (2.69a) and (2.69b), and applying (3.45a) and (3.45b) with $\mathbf{r} = \mathbf{r}_\beta^\alpha$ leads to:

$$\begin{bmatrix} 0 \\ \dot{\boldsymbol{x}}^\alpha \end{bmatrix} = \mathbf{r}_\beta^\alpha \odot \left(\begin{bmatrix} 0 \\ \dot{\boldsymbol{x}}^\beta \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right) \end{bmatrix} \odot \begin{bmatrix} 0 \\ \boldsymbol{x}^\beta \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{x}^\beta \end{bmatrix} \odot \begin{bmatrix} 0 \\ -\boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right) \end{bmatrix}\right) \odot \left(\mathbf{r}_\beta^\alpha\right)^*$$

Carrying out the inner quaternion products gives:

$$\begin{bmatrix} 0 \\ \dot{\boldsymbol{x}}^\alpha \end{bmatrix} = \mathbf{r}_\beta^\alpha \odot \begin{bmatrix} 0 \\ \dot{\boldsymbol{x}}^\beta + \boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right) \times \boldsymbol{x}^\beta \end{bmatrix} \odot \left(\mathbf{r}_\beta^\alpha\right)^* \tag{3.46}$$

From (2.70) we see that the right hand side of (3.46) represents a change of basis of vector $\dot{\boldsymbol{x}}^\beta + \boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right) \times \boldsymbol{x}^\beta$ from basis $\mathcal{E}_\beta$ to $\mathcal{E}_\alpha$. Extracting the vector part of this equality we can write:

$$\dot{\boldsymbol{x}}^\alpha = \boldsymbol{R}_\beta^\alpha\left(\dot{\boldsymbol{x}}^\beta + \boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right) \times \boldsymbol{x}^\beta\right) = \boldsymbol{R}_\beta^\alpha\left(\dot{\boldsymbol{x}}^\beta + \left[\boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right)^\times\right]\boldsymbol{x}^\beta\right) \tag{3.47}$$

Comparing (3.18) and (3.47) shows that:

$$\left[\boldsymbol{\omega}\left(\mathbf{r}_\beta^\alpha\right)^\times\right] = \boldsymbol{\Omega}_{\alpha\beta}^\beta$$

And therefore:

$$\boldsymbol{\omega}\left(\mathbf{r}_{\beta}^{\alpha}\right) = \boldsymbol{\omega}_{\alpha\beta}^{\beta}$$

We may now rewrite equation (3.45a) with $\mathbf{r} = \mathbf{r}_{\beta}^{\alpha}$:

$$\left(\mathbf{r}_{\beta}^{\alpha}\right)^{*} \odot \dot{\mathbf{r}}_{\beta}^{\alpha} = \frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{\omega}\left(\mathbf{r}_{\beta}^{\alpha}\right) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 \\ \boldsymbol{\omega}_{\alpha\beta}^{\beta} \end{bmatrix} = \frac{1}{2}\boldsymbol{\omega}_{\alpha\beta}^{\beta} \tag{3.48}$$

Where $\boldsymbol{\omega}_{\alpha\beta}^{\beta}$ denotes the vector quaternion:

$$\boldsymbol{\omega}_{\alpha\beta}^{\beta} = \begin{bmatrix} 0 \\ \boldsymbol{\omega}_{\alpha\beta}^{\beta} \end{bmatrix}$$

Solving for $\dot{\mathbf{r}}_{\beta}^{\alpha}$ in (3.48) finally yields the kinematic equation for $\mathbf{r}_{\beta}^{\alpha}$:

$$\dot{\mathbf{r}}_{\beta}^{\alpha} = \frac{1}{2}\mathbf{r}_{\beta}^{\alpha} \odot \boldsymbol{\omega}_{\alpha\beta}^{\beta} \tag{3.49}$$

From (2.74) we have:

$$\boldsymbol{\omega}_{\alpha\beta}^{\beta} = \left(\mathbf{r}_{\beta}^{\alpha}\right)^{*} \odot \boldsymbol{\omega}_{\alpha\beta}^{\alpha} \odot \mathbf{r}_{\beta}^{\alpha} \tag{3.50}$$

Inserting (3.50) into (3.49) yields the alternative form of the kinematic equation:

$$\dot{\mathbf{r}}_{\beta}^{\alpha} = \frac{1}{2}\boldsymbol{\omega}_{\alpha\beta}^{\alpha} \odot \mathbf{r}_{\beta}^{\alpha} \tag{3.51}$$