A SYSTEMATIC APPROACH TO DYNAMICS

Course notes for MECH 220, MECH 600, and MECH 642

James Richard Forbes
Associate Professor
Department of Mechanical Engineering
McGill University
817 Sherbrooke Street West
Montreal, Quebec, Canada
H3A 0C3
james.richard.forbes@mcgill.ca

August 28, 2018

Preface

Dynamics is hard. Period. In this text, dynamics is approached in a systematic way. There's no hand waving, no tricks, no relying solely on gut feel and intuition. Rather, problems are solved using a series of systematic steps using rigorous and consistent notation that, when followed diligently, will not lead the student or practitioner astray.

Now, the statement "no relying on gut feel and intuition" does not mean readers are discouraged from asking the "does this answer make sense" type of questions. On the contrary. "Sanity checks", as well as mathematical and numerical checks, are very much encouraged. However, when problems get complicated, such as kinematics and dynamics problems in three dimensions (3D), solving a problem using gut feel or intuition alone is simply not possible.

At first readers might find the notation used in this text is rather mathematical and perhaps cumbersome, and for simple problems, maybe "overkill'. However, based on past experience, without a clear, consistent, and descriptive notation, dynamics problems can quickly become overwhelming.

One might ask, "well, who else teaches dynamics this way". Many world-class engineering institutions, such as the University of Toronto, the University of Waterloo, the Technion, the University of Michigan, Purdue, and others have adopted a similar presentation of the same material, as well as similar notation. The intend of this text is to help provide a dynamics education that is not just on-par with other universities, but is better.

This text has been inspired by various other texts on dynamics. The notation used is from [1] and [2], with modifications inspired by [3][1]. In addition to [1–3], the books [4, 5] have been a great inspiration as far as examples of texts that teach dynamics in a systematic way.

Please note that this text is a work in progress that is updated, corrected, and extended quite frequently. If you notice a typo (a spelling error, a mathematical error, anything), please notify Prof. Forbes at

`james.richard.forbes@mcgill.ca`

[1]Interestingly, the primary author of [3], Dr. Dennis S. Bernstein, was originally inspired by the notation of [2]. According to Dr. Bernstein, he was given a set of notes from Dr. Jon How when Dr. How was a PhD student. The notes Dr. How gave Dr. Bernstein were co-authored by Dr. D'Eleutario from the University of Toronto where Dr. How did his undergraduate degree. Those same notes were the start of [2].

Acknowledgements

Thanks to my friends, colleagues, students, and family. In particular, thanks to Christopher J. Damaren, Gabriele D'Eleuterio, Timothy D. Barfoot, Anton H. J. de Ruiter, Dennis S. Bernstein, François Robert Hogan, Alex Walsh, Ryan James Caverly, David Evan Zlotnik, Neils Roth, Bruce Francis, Christopher John Forbes, Andrew James Taylor, and Allison Mary Gerhart.

James Richard Forbes
August 28, 2018

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Why Am I Reading This?

If you're reading this, then for one reason or another you're interested in the study of the *dynamics* of particles, rigid bodies, flexible bodies, or a combination of all three. Dynamics is a subdiscipline of *mechanics*. Mechanics is one of the oldest subjects studied by great minds such as Galileo, Euler, D'Alembert, Lagrange, Laplace, Gibbs, Leibniz, Bernoulli (well, many different Bernoullies), and, of course, Newton. Well, what is mechanics?

Definition 1.1. *Mechanics* is the study of the effect of forces acting on a body [7].

Mechanics can be further divided into three subdisiplines, one of which is dynamics: *statics*, *mechanics of materials* (also called *strength of materials*), and *dynamics*. Well, what is statics, mechanics of materials, and dynamics?

Definition 1.2. *Statics* is the study of bodies at rest while under the influence of forces [7].

Definition 1.3. *Mechanics of materials* is the study of the effects of stress and strain within a body subject to forces [8].

Definition 1.4. *Dynamics* is the study of bodies in motion under the action of forces [7] [5, pp. 1].

This text will focus on the study of dynamics. Although it's important to have a precise definition of dynamics, Definition 1.4 can be a bit misleading. What dynamics is really all about is *deriving the differential equation that describe the motion of a body* [2, 4, 5]. This is very important to understand. Dynamics is not about finding the configuration of a body at a particular instant in time, it's about deriving the set of differential equations that describe the motion of the body from one point in time to another. Given that a dynamic analysis will yield a differential equation, which is often nonlinear, it is no wonder undergraduate students also take courses focused on differential equations, linear algebra, and numerical methods.

Now, how does one go about executing a dynamics analysis? A dynamic analysis first requires a *kinematic* analysis.

Definition 1.5. *Kinematics* is the study of the geometry of motion, without concern or regard for the cause of motion [7].

Kinematics involves geometry, and geometry involves vectors (... well, physical or Gibbsian vectors, actually). As such, before considering a dynamic analysis, kinematics must be clearly understood.

Much like building a house, starting with a solid foundation is imperative. This text starts with geometry, then moves onto kinematics, then finally to dynamics. It is illogical to attempt to study dynamics without first mastering kinematics, just as it is illogical to attempt to study kinematics without mastering geometry. In Appendix A is an overview of linear algebra, differential equations, and numerical methods, tools that are extremely useful when attempting to solve dynamics problems.

Chapter 2

Physical Vectors, Tensors, and the Direction Cosine Matrix

Before discussing how points, masses, and rigid bodies move in time, and the laws that govern their motion, how to describe and parameterize their position and attitude must be discussed.

2.1 Physical Vectors

In this text, bodies that exist and move in *physical space*, the three-dimensional space in which physical bodies reside, will be studied. To describe how a body moves in physical space, *physical vectors*, also called *Gibbsian vectors*, will be used.

Definition 2.1. *Physical space*, denoted $\mathbb{P}$, is a linear vector space with elements called *physical vectors*, denoted $\underset{\rightarrow}{u}$, which have both magnitude and direction.

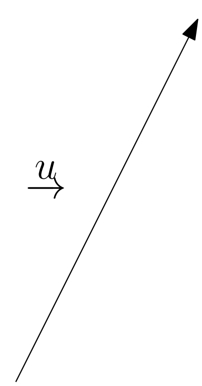


Figure 2.1: A physical vector.

As noted in Definition 2.1, a physical vector has both magnitude and direction [4, pp. 3]. Often physical vectors are drawn as arrows, as shown in Figure 2.1. The notation $\underset{\rightarrow}{u} \in \mathbb{P}$ means the physical vector $\underset{\rightarrow}{u}$ is an element of physical space $\mathbb{P}$.

The length of a physical vector $\underset{\rightarrow}{u} \in \mathbb{P}$ is called the *magnitude* or *Euclidean norm* of $\underset{\rightarrow}{u}$ and is denoted $||\underset{\rightarrow}{u}||_2$. The Euclidian norm of a physical vector is a scalar, that is $||\underset{\rightarrow}{u}||_2 \in \mathbb{R}$, where $\mathbb{R}$ denotes the set of real numbers. A physical vector with zero magnitude is called the *zero vector*, denoted $\underset{\rightarrow}{0}$, and has no particular direction. A physical *unit vector* has unit length, that is $||\underrightarrow{u}||_2 = 1$, because its magnitude is unity. The *direction* of a physical (nonzero) vector such as $\underrightarrow{u} \in \mathbb{P}$ is

$$\frac{\underrightarrow{u}}{||\underrightarrow{u}||_2},$$

that is, the physical vector divided by its magnitude. Two physical vectors are equal if they have the same magnitude and direction [4, pp. 3].

Notice that the definition of a physical vector does not involve any sort of "reference frame" or "coordinate system", just as the notion of a scalar, such as mass or density, is independent of a reference frame or coordinate system. Indeed, eventually reference frames and various types of coordinate systems (such as Cartesian, polar, and spherical coordinate systems) will be defined and used, but a physical vector is an entity that does not depend on reference frames nor coordinate systems. This last point is worth reiterating: physical vectors are independent of reference frames and coordinate systems.

Before continuing, perhaps a comment on notation is in order. In this text a physical vector is denoted by an "underarrow", such as $\underrightarrow{u}$. This notation is also used in [1, 2]. Other texts use an "overarrow", $\vec{u}$, an "overbar", $\bar{u}$, an "underbar", $\underline{u}$, boldface font, $\mathbf{u}$, or boldface-italic font, $\boldsymbol{u}$, to denote a physical vector. A preference for using an underarrow is due to the fact that superscripts are heavily used (e.g., $\underrightarrow{u}^{zw}$), and having both an overarrow and superscripts atop a letter gets rather crowded.

Readers may be curious as to the elements of physical space are called physical vectors (or Gibbsian vectors), rather than just plane old "vectors". This is done for clarity, because there are many types of vectors. In general, a vector is an element of a linear vector space. For instance, the set of all 4×1 matrices, such as

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \tag{2.1}$$

are elements of a linear vector space. As such, to avoid confusion, entities similar to those in Equation (2.1) will be called *column matrices*, rather than vectors [1, pp. 523].

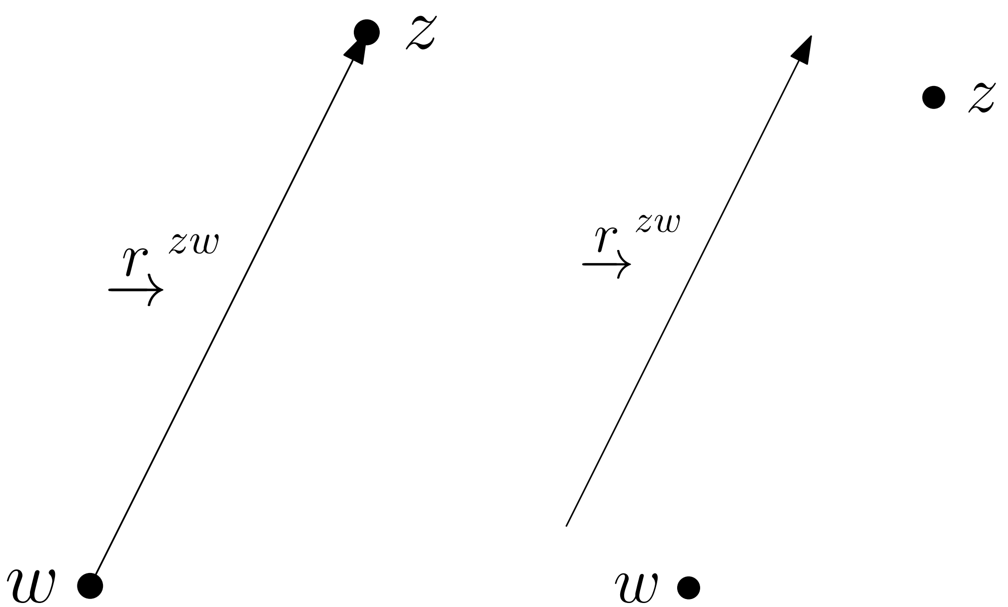
2.1.1 Points and the Notion of Relative Position

Definition 2.2. A point has zero size and zero mass.

The notion of "position" is a relative. For example, it's nonsense to attempt to discuss the position of "point z" in physical space. On the other hand, it is natural to discuss the position of "point z relative to point w" in physical space. Shown in Figure 2.2(a) are two points, points w and z. The position of point z relative to point w is described by $\underrightarrow{r}^{zw}$, where $\underrightarrow{r}^{zw} = -\underrightarrow{r}^{wz}$.

Although points are located relative to each other in physical space, physical vectors do not have a physical location in physical space. A physical vector is described only by magnitude and a direction. There's no mention of "the position" of a physical vector in the definition of a physical vector nor in the definition of physical space. As shown in Figure 2.2, the points w and z have physical locations relative to each other in physical space, but the physical vector $\underrightarrow{r}^{zw}$ does not have a physical location.

In the remainder of this chapter, physical vectors in general will be discussed. A physical vector that describes the position of one point relative to another point is just one kind of physical vector. The properties of physical vectors hold for any and all kinds of physical vectors, not just "position" physical vectors.



(a) The physical vector $\underset{\rightarrow}{r}^{zw}$ describing the position of point z relative to point w.

(b) Points, such as points w and z, in physical space have physical locations. However, $\underset{\rightarrow}{r}^{zw}$ does not have a physical location.

Figure 2.2: Points and physical vectors.

2.1.2 Addition of Physical Vectors

Associated with physical vectors are two operations, *summation* and *the product of a physical vector with a scalar*. In this section the properties of summation (also called addition) will be reviewed [4, pp. 4-5], [5, pp. 636-637], [18, pp. 1-2].

Definition 2.3. Consider $\underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}$. The *sum* of $\underset{\rightarrow}{u} \in \mathbb{P}$ and $\underset{\rightarrow}{v} \in \mathbb{P}$ is

$$\underset{\rightarrow}{u} + \underset{\rightarrow}{v} \in \mathbb{P}.$$

Vector addition satisfies the following properties.

1. (Commutativity of vector addition.)

$$\underset{\rightarrow}{u} + \underset{\rightarrow}{v} = \underset{\rightarrow}{v} + \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}.$$

2. (Associativity of vector addition.)

$$(\underset{\rightarrow}{u} + \underset{\rightarrow}{v}) + \underset{\rightarrow}{w} = \underset{\rightarrow}{v} + (\underset{\rightarrow}{u} + \underset{\rightarrow}{w}), \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}.$$

3. (Additive identity.) $\exists \underset{\rightarrow}{0} \in \mathbb{P}$ such that

$$\underset{\rightarrow}{u} + \underset{\rightarrow}{0} = \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

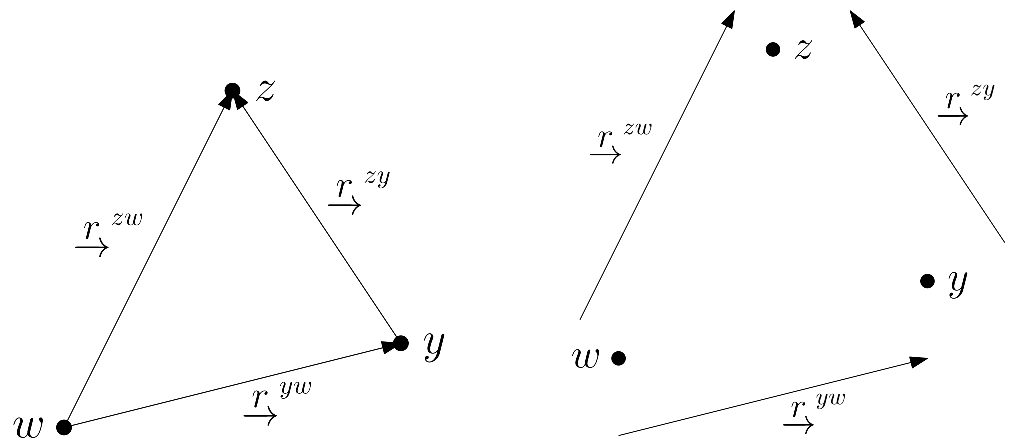
4. (Identity for addition.)

$$\underset{\rightarrow}{u} + (-\underset{\rightarrow}{u}) = (-\underset{\rightarrow}{u}) + \underset{\rightarrow}{u} = \underset{\rightarrow}{0}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

Graphically, physical vector addition is shown in Figure 2.3, where physical vectors are added in a "tip-to-tail" fashion. In particular,

$$\underset{\rightarrow}{r}^{zw} = \underset{\rightarrow}{r}^{zy} + \underset{\rightarrow}{r}^{yw}.$$

Notice how the superscripts "line up" when going from right to left.



(a) The addition of physical vectors $\underset{\rightarrow}{r}^{yw}$ and $\underset{\rightarrow}{r}^{zw}$ yielding $\underset{\rightarrow}{r}^{zw}$.

(b) Although it is often convenient to draw and add physical vectors in a "tip-to-tail" fashion, in general $\underset{\rightarrow}{r}^{yw}$, $\underset{\rightarrow}{r}^{zw}$, and $\underset{\rightarrow}{r}^{zw}$ do not have a physical location.

Figure 2.3: Addition of physical vectors.

2.1.3 Multiplication of a Physical Vector by a Scalar

Now the product of a physical vector and a scalar will be considered [4, pp. 6-7], [5, pp. 636-637], [18, pp. 1-2].

Definition 2.4. Let $\underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}$, $a, b \in \mathbb{R}$, and $|a|$ denote the absolute value of the scalar a. The *product* of $\underset{\rightarrow}{u}$ with a is

$$a\,\underset{\rightarrow}{u} \in \mathbb{P}.$$

The *product* of $\underset{\rightarrow}{u}$ with a satisfies the following properties.

1. (Distributivity of vector sums.)

$$a(\underset{\rightarrow}{u} + \underset{\rightarrow}{v}) = a\,\underset{\rightarrow}{u} + a\,\underset{\rightarrow}{v}, \quad \forall a \in \mathbb{R}, \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}.$$

2. (Distributively of scalar sums.)

$$(a+b)\,\underset{\rightarrow}{u} = a\,\underset{\rightarrow}{u} + b\,\underset{\rightarrow}{u}, \quad \forall a, b \in \mathbb{R}, \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

3. (Associativity of scalar multiplication.)

$$a(b\,\underset{\rightarrow}{u}) = (ab)\,\underset{\rightarrow}{u}, \quad \forall a, b \in \mathbb{R}, \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

4. (Scalar identity.)

$$(1)\,\underset{\rightarrow}{u} = \underset{\rightarrow}{u}(1) = \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

5. $(0)\underrightarrow{u} = \underrightarrow{u}(0) = \underrightarrow{0}, \forall \underrightarrow{u} \in \mathbb{P}$.
6. $||a\,\underrightarrow{u}||_2 = |a|\,||\underrightarrow{u}||_2, \quad \forall a \in \mathbb{R}, \forall \underrightarrow{u} \in \mathbb{P}$.
7. $a\,\underrightarrow{u} = \underrightarrow{0}$ if either $\underrightarrow{u} = \underrightarrow{0}$ or $a = 0$.

2.1.4 Physical Basis Vectors, Reference Frames, Vectrices, and Components

Often it's convenient to express a physical vector in terms of three *dextral, orthonormal physical basis vectors*, which means the three physical basis vectors are arranged in a right-handed fashion, are orthogonal to each other, and are normalized (i.e., they each are of unit length) [2], [1, pp. 6-7]. Because the physical basis vectors are orthogonal to each other, they're also linearly independent, and thus form a basis for $\mathbb{P}$. The physical basis vectors are themselves physical vectors and together they compose a *reference frame*, denoted $\mathcal{F}$. The three physical basis vectors composing a reference frame are often collectively referred to as an *orthonormal triad* [1, pp. 6-7].

Definition 2.5. A *reference frame*, denoted $\mathcal{F}$, is composed of three dextral, orthonormal physical basis vectors that form a basis for $\mathbb{P}$.

In Figure 2.4(a) is a set of three dextral, orthonormal physical basis vectors, $\underrightarrow{a}^1, \underrightarrow{a}^2, \underrightarrow{a}^3 \in \mathbb{P}$, composing reference frame "a" which will be denoted $\mathcal{F}_a$. Given that the physical basis vectors form a dextral orthonormal basis they satisfy

$$\begin{array}{ccc}
\underrightarrow{a}^1 \cdot \underrightarrow{a}^1 = 1\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^3 = 0\,, \\
\underrightarrow{a}^2 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^2 = 1\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^3 = 0\,, \\
\underrightarrow{a}^3 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^3 = 1\,,
\end{array}$$

and

$$\begin{array}{ccc}
\underrightarrow{a}^1 \times \underrightarrow{a}^1 = \underrightarrow{0}\,, & \underrightarrow{a}^1 \times \underrightarrow{a}^2 = \underrightarrow{a}^3\,, & \underrightarrow{a}^1 \times \underrightarrow{a}^3 = -\underrightarrow{a}^2\,, \\
\underrightarrow{a}^2 \times \underrightarrow{a}^1 = -\underrightarrow{a}^3\,, & \underrightarrow{a}^2 \times \underrightarrow{a}^2 = \underrightarrow{0}\,, & \underrightarrow{a}^2 \times \underrightarrow{a}^3 = \underrightarrow{a}^1\,, \\
\underrightarrow{a}^3 \times \underrightarrow{a}^1 = \underrightarrow{a}^2\,, & \underrightarrow{a}^3 \times \underrightarrow{a}^2 = -\underrightarrow{a}^1\,, & \underrightarrow{a}^3 \times \underrightarrow{a}^3 = \underrightarrow{0}\,,
\end{array}$$

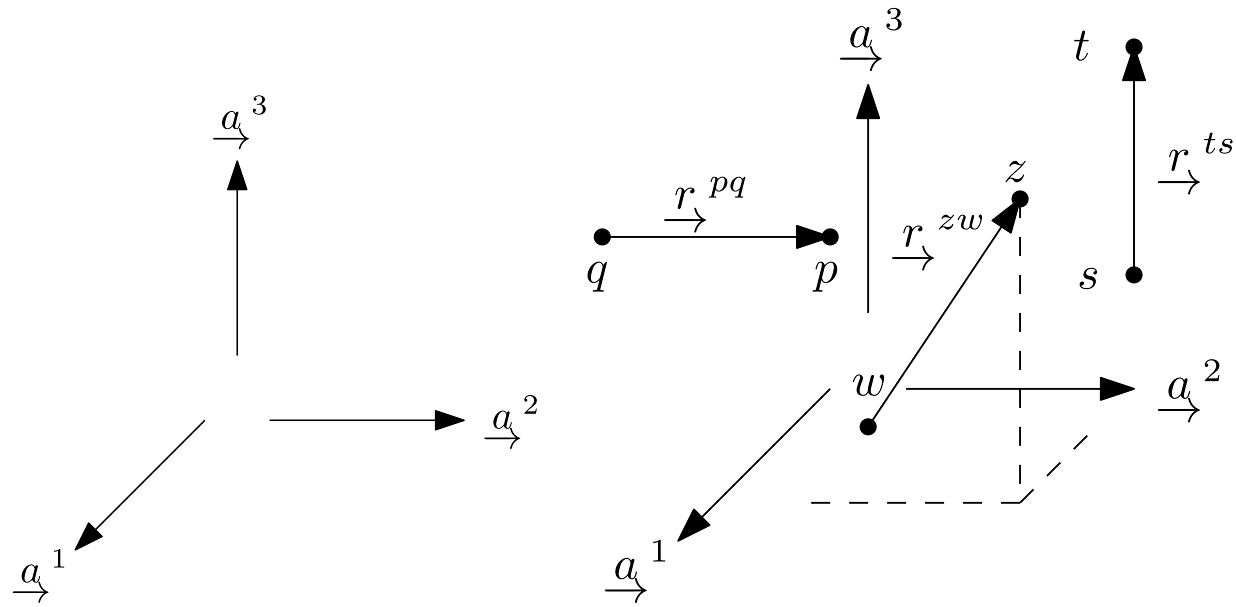
where $(\triangle)\cdot(\square)$ is the dot product and $(\triangle)\times(\square)$ is the cross product (both of which will be officially introduced in Sections 2.1.5 and 2.1.6, respectively).

A physical vector, such as $\underrightarrow{u} \in \mathbb{P}$, can be written as a linear combination of the physical basis vectors $\underrightarrow{a}^1, \underrightarrow{a}^2, \underrightarrow{a}^3 \in \mathbb{P}$. For instance, consider Figure 2.4(b) and the three physical vectors $\underrightarrow{r}^{pq}, \underrightarrow{r}^{ts}, \underrightarrow{r}^{zw} \in \mathbb{P}$. In terms of the physical basis vectors $\underrightarrow{a}^1, \underrightarrow{a}^2, \underrightarrow{a}^3 \in \mathbb{P}$, the three physical vectors $\underrightarrow{u}$, $\underrightarrow{v}$, and $\underrightarrow{w}$ are

$$\begin{aligned}
\underrightarrow{r}^{pq} &= \underrightarrow{a}^1(0) + \underrightarrow{a}^2(r_{a2}^{pq}) + \underrightarrow{a}^3(0), \\
\underrightarrow{r}^{ts} &= \underrightarrow{a}^1(0) + \underrightarrow{a}^2(0) + \underrightarrow{a}^3(r_{a3}^{ts}), \\
\underrightarrow{r}^{zw} &= \underrightarrow{a}^1(r_{a1}^{zw}) + \underrightarrow{a}^2(r_{a2}^{zw}) + \underrightarrow{a}^3(r_{a3}^{zw}),
\end{aligned}$$

where r_{a2}^{pq}, r_{a3}^{ts}, and r_{a1}^{zw}, r_{a3}^{zw}, r_{a3}^{zw} are the *components* of $\underrightarrow{r}^{pq}$, $\underrightarrow{r}^{ts}$, and $\underrightarrow{r}^{zw}$ *resolved in*, or *expressed in*, the reference frame $\mathcal{F}_a$ defined by the physical basis vectors $\underrightarrow{a}^1, \underrightarrow{a}^2, \underrightarrow{a}^3 \in \mathbb{P}$ [2], [1, pp. 6-8].

Although physical vectors are entities that are independent of a particular reference frame, the components of the physical vector do in fact depend on the particular reference frame used. Shortly many reference frames will be used all at once, and it's crucial to understand that, in general, the components of physical vector resolved in one frame, say $\mathcal{F}_a$, are not equal to the components of the same physical vector resolved in another frame, say $\mathcal{F}_b$.



(a) Physical basis vectors composing a reference frame.

(b) Various physical vectors.

Figure 2.4: Dextral orthonormal reference frames.

Referring to Figure 2.4(b) once more, notice that a physical vector does not have to have its "tail" at the origin of the reference frame.

2.1.4.1 Vectrices

In this text the physical basis vectors, such as $\underrightarrow{a}^1, \underrightarrow{a}^2, \underrightarrow{a}^3 \in \mathbb{P}$, and the components of physical vectors, such as $\underrightarrow{u}, \underrightarrow{v}, \underrightarrow{w} \in \mathbb{P}$, will be deliberately and explicitly written in terms of two separate but related matrices as

$$\begin{aligned}
\underrightarrow{u} &= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \begin{bmatrix} 0 \\ u_{a2} \\ 0 \end{bmatrix}, \\
\underrightarrow{v} &= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ v_{a3} \end{bmatrix}, \\
\underrightarrow{w} &= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \begin{bmatrix} w_{a1} \\ w_{a2} \\ w_{a3} \end{bmatrix}.
\end{aligned}$$

In particular, define

$$\underrightarrow{\mathcal{F}}_a = \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix} \tag{2.2}$$

which is called a *vectrix*, a term coined by Dr. Peter C. Hughes [1, pp. 522]. The word vectrix highlights the fact that a vectrix is a (column) matrix of vectors. The plural of the word vectrix is "vectrices", much like the plural of matrix is "matrices". Using a vectrix, $\underrightarrow{\mathcal{F}}_a$, physical vectors such as $\underrightarrow{u}, \underrightarrow{v}, \underrightarrow{w} \in \mathbb{P}$ can be written as

$$\underset{\rightarrow}{u} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 0 \\ u_{a2} \\ 0 \end{bmatrix}}_{\mathbf{u}_a} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a = \mathbf{u}_a^{\mathsf{T}} \underset{\rightarrow}{\mathcal{F}}_a,$$

$$\underset{\rightarrow}{v} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 0 \\ 0 \\ v_{a3} \end{bmatrix}}_{\mathbf{v}_a} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a = \mathbf{v}_a^{\mathsf{T}} \underset{\rightarrow}{\mathcal{F}}_a,$$

$$\underset{\rightarrow}{w} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} w_{a1} \\ w_{a2} \\ w_{a3} \end{bmatrix}}_{\mathbf{w}_a} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{w}_a = \mathbf{w}_a^{\mathsf{T}} \underset{\rightarrow}{\mathcal{F}}_a,$$

where $\mathbf{u}_a, \mathbf{v}_a, \mathbf{w}_a \in \mathbb{R}^3$ are the components of the physical vectors $\underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}$ resolved in $\mathcal{F}_a$. Observe that $\mathbf{u}_a$, $\mathbf{v}_a$, and $\mathbf{w}_a$ are column matrices full of components (which are scalars). From the rules of matrix operation it follows that

$$\begin{aligned} \mathbf{u}_a^{\mathsf{T}} &= \begin{bmatrix} 0 & u_{a2} & 0 \end{bmatrix}, \\ \mathbf{v}_a^{\mathsf{T}} &= \begin{bmatrix} 0 & 0 & v_{a3} \end{bmatrix}, \\ \mathbf{w}_a^{\mathsf{T}} &= \begin{bmatrix} w_{a1} & w_{a2} & w_{a3} \end{bmatrix}. \end{aligned}$$

The vectrix $\underset{\rightarrow}{\mathcal{F}}_a$ is a column matrix of vectors. Accordingly, $\underset{\rightarrow}{\mathcal{F}}_a$ also adheres to the rules of matrix operation, which is why

$$\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}.$$

2.1.4.2 Some Examples

To help solidify the use of vectrices, column matrices, and their relation to physical vectors, consider the following examples.

Example 2.1. Consider Figure 2.5. Express the physical vectors $\underset{\rightarrow}{r}^{zw} = \underset{\rightarrow}{a}^1(0) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(3)$ and $\underset{\rightarrow}{r}^{yz} = \underset{\rightarrow}{a}^1(1) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(0)$ in terms of the vectrix $\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}$ and the column matrices $\mathbf{r}_a^{zw}$ and $\mathbf{r}_a^{yz}$.

Solution.

$$\underset{\rightarrow}{r}^{zw} = \underset{\rightarrow}{a}^1(0) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(3) = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix} \overbrace{\begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}}^{\mathbf{r}_a^{zw}} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{zw},$$

$$\underset{\rightarrow}{r}^{yz} = \underset{\rightarrow}{a}^1(1) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(0) = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}_{\mathbf{r}_a^{yz}} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{yz},$$

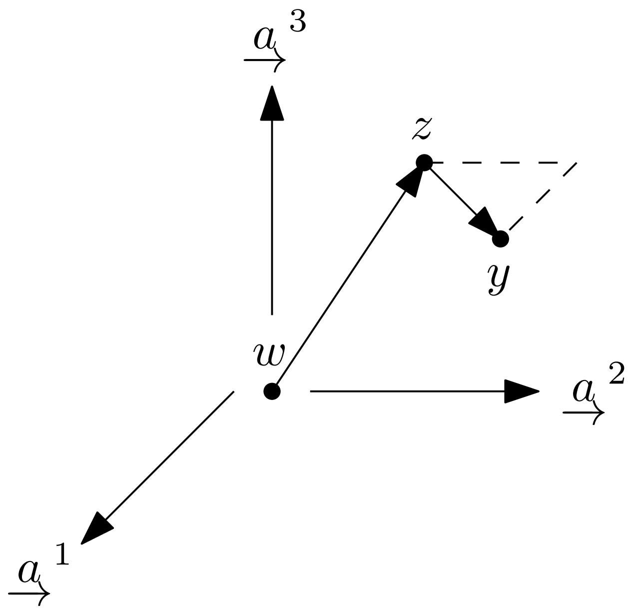


Figure 2.5: Physical vectors $\underset{\rightarrow}{r}^{zw}$ and $\underset{\rightarrow}{r}^{yz}$ to be added.

$\square$

Example 2.2. Consider Figure 2.5 once more. Add the physical vectors $\underset{\rightarrow}{r}^{zw}$ and $\underset{\rightarrow}{r}^{yz}$ to yield $\underset{\rightarrow}{r}^{yw}$, and express $\underset{\rightarrow}{r}^{yw}$ in terms of the vectrix $\underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}$ and the column matrix $\mathbf{r}_a^{yw}$.

Solution. Begin by adding the two physical vectors,

$$\begin{aligned}
\underset{\rightarrow}{r}^{yw} &= \underset{\rightarrow}{r}^{yz} + \underset{\rightarrow}{r}^{zw} \\
&= \left(\underset{\rightarrow}{a}^1(1) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(0)\right) + \left(\underset{\rightarrow}{a}^1(0) + \underset{\rightarrow}{a}^2(2) + \underset{\rightarrow}{a}^3(3)\right) \\
&= \underset{\rightarrow}{a}^1(1) + \underset{\rightarrow}{a}^2(2+2) + \underset{\rightarrow}{a}^3(3) \\
&= \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}}_{\mathbf{r}_a^{yw}} \\
&= \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T} \mathbf{r}_a^{yw}.
\end{aligned}$$

Alternatively,

$$
\begin{aligned}
\underset{\rightarrow}{r}^{yw} &= \underset{\rightarrow}{r}^{yz} + \underset{\rightarrow}{r}^{zw} \\
&= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{yz} + \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{zw} \\
&= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \left(\mathbf{r}_a^{yz} + \mathbf{r}_a^{zw} \right) \\
&= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \right) \\
&= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}}_{\mathbf{r}_a^{yw}} \\
&= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{yw}.
\end{aligned}
$$

□

Before continuing, it's worth emphasizing once again that physical vectors are entities that are independent of a particular reference frame $\mathcal{F}_a$. In due course physical vectors, various different reference frames, and the relationship between the physical vectors resolved in the various reference frames will be considered.

2.1.5 Dot Product

Definition 2.6. Given two physical vectors $\underset{\rightarrow}{u} \in \mathbb{P}$ and $\underset{\rightarrow}{v} \in \mathbb{P}$, the *dot product* (also called the *scalar product*, or more generally referred to as the *inner product*) between the two physical vectors is defined as

$$\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v} = ||\underset{\rightarrow}{u}||_2 \, ||\underset{\rightarrow}{v}||_2 \cos\theta,$$

where $\theta \in [0, 180°]$ is the angle between $\underset{\rightarrow}{u}$ and $\underset{\rightarrow}{v}$.

Let $\underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}$ and $a \in \mathbb{R}$. The dot product satisfies the following properties [4, pp. 7], [5, pp. 638-639].

1. (Commutativity.)
$$\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v} = \underset{\rightarrow}{v} \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}.$$

2. (Distributivity.)
$$(\underset{\rightarrow}{u} + \underset{\rightarrow}{v}) \cdot \underset{\rightarrow}{w} = \underset{\rightarrow}{u} \cdot \underset{\rightarrow}{w} + \underset{\rightarrow}{v} \cdot \underset{\rightarrow}{w}, \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}.$$

3. (Scalar multiplication.)
$$\underset{\rightarrow}{u} \cdot (a\,\underset{\rightarrow}{v}) = a\,\underset{\rightarrow}{v} \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}, \quad \forall a \in \mathbb{R}.$$

4. $\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{u} = 0$ only if $\underset{\rightarrow}{u} = \underset{\rightarrow}{0}$.

Two different physical vectors $\underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}$ are *orthogonal* if

$$\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v} = 0.$$

A set of physical vectors $\left\{\underrightarrow{a}^1, \ldots, \underrightarrow{a}^n\right\}$ are *mutually orthogonal* if

$$\underrightarrow{a}^i \cdot \underrightarrow{a}^j = 0, \quad i \neq j, \quad i, j = 1 \ldots n.$$

The magnitude of the physical vector $\underrightarrow{u} \in \mathbb{P}$ is

$$||\underrightarrow{u}||_2 = \sqrt{\underrightarrow{u} \cdot \underrightarrow{u}}.$$

2.1.5.1 The Dot Product and Its Relation to Vectrices and Components

Consider two physical vectors $\underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}$ expressed in terms of the vectrix $\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \left[\underrightarrow{a}^1 \;\; \underrightarrow{a}^2 \;\; \underrightarrow{a}^3\right]$ and the components of the physical vectors, that is

$$\underrightarrow{u} = \left[\begin{array}{ccc} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{array}\right] \overbrace{\left[\begin{array}{c} u_{a1} \\ u_{a2} \\ u_{a3} \end{array}\right]}^{\mathbf{u}_a} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a = \mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a, \tag{2.3}$$

$$\underrightarrow{v} = \left[\begin{array}{ccc} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{array}\right] \underbrace{\left[\begin{array}{c} v_{a1} \\ v_{a2} \\ v_{a3} \end{array}\right]}_{\mathbf{v}_a} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a = \mathbf{v}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a. \tag{2.4}$$

Recall that the physical basis vectors $\underrightarrow{a}^1$, $\underrightarrow{a}^2$, and $\underrightarrow{a}^3$ are themselves physical vectors. As such,

$$\begin{array}{lll} \underrightarrow{a}^1 \cdot \underrightarrow{a}^1 = 1\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^3 = 0\,, \\ \underrightarrow{a}^2 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^2 = 1\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^3 = 0\,, \\ \underrightarrow{a}^3 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^3 = 1\,. \end{array}$$

Now consider the dot product of $\underrightarrow{u}$ and $\underrightarrow{v}$ [2], [1, pp. 523-526],

$$\begin{aligned}
\underrightarrow{u} \cdot \underrightarrow{v} &= \left(\mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a\right) \cdot \left(\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a\right) \\
&= \mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \left[\begin{array}{c} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{array}\right] \cdot \left[\begin{array}{ccc} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{array}\right] \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \left[\begin{array}{ccc} \underrightarrow{a}^1 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^1 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^1 \cdot \underrightarrow{a}^3 \\ \underrightarrow{a}^2 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^2 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^2 \cdot \underrightarrow{a}^3 \\ \underrightarrow{a}^3 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^3 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^3 \cdot \underrightarrow{a}^3 \end{array}\right] \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \underbrace{\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]}_{\mathbf{1}} \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \mathbf{v}_a \\
&= u_{a1} v_{a1} + u_{a2} v_{a2} + u_{a3} v_{a3},
\end{aligned}$$

where **1** is the *identity matrix*. It's important to realize that both matrix operations and vector operations have been used. First, a matrix operation was used to expand $\underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}$, then the dot product was used to evaluate $\underrightarrow{a}^1 \cdot \underrightarrow{a}^1 = 1$, $\underrightarrow{a}^1 \cdot \underrightarrow{a}^2 = 0$, $\underrightarrow{a}^1 \cdot \underrightarrow{a}^3 = 0$, etc., thus arriving at [2], [1, pp. 523-526]

$$\underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \mathbf{1}. \tag{2.5}$$

Consider the expression for the physical vector $\underrightarrow{u}$ given in Equation (2.3), but expanded. Taking the dot product of Equation (2.3) with $\underrightarrow{a}^1$, $\underrightarrow{a}^2$, and $\underrightarrow{a}^3$ yields

$$\begin{aligned}
\underrightarrow{a}^1 \cdot \underrightarrow{u} &= \underrightarrow{a}^1 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) = \underrightarrow{a}^1 \cdot \underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^1 \cdot \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^1 \cdot \underrightarrow{a}^3(u_{a3}) = u_{a1},\\
\underrightarrow{a}^2 \cdot \underrightarrow{u} &= \underrightarrow{a}^2 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) = \underrightarrow{a}^2 \cdot \underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2 \cdot \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^2 \cdot \underrightarrow{a}^3(u_{a3}) = u_{a2},\\
\underrightarrow{a}^3 \cdot \underrightarrow{u} &= \underrightarrow{a}^3 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) = \underrightarrow{a}^3 \cdot \underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^3 \cdot \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3 \cdot \underrightarrow{a}^3(u_{a3}) = u_{a3}.
\end{aligned}$$

It follows that taking the dot product from the left of Equation (2.3) with $\underrightarrow{\mathcal{F}}_a$ given in Equation (2.2) results in [2], [1, pp. 523-526]

$$\underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{u} = \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix} \cdot \underrightarrow{u} = \begin{bmatrix} \underrightarrow{a}^1 \cdot \underrightarrow{u} \\ \underrightarrow{a}^2 \cdot \underrightarrow{u} \\ \underrightarrow{a}^3 \cdot \underrightarrow{u} \end{bmatrix} = \begin{bmatrix} \underrightarrow{a}^1 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) \\ \underrightarrow{a}^2 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) \\ \underrightarrow{a}^3 \cdot \left(\underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})\right) \end{bmatrix} = \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix} = \mathbf{u}_a.$$

As such, taking the dot product of the vectrix $\underrightarrow{\mathcal{F}}_a$ with a physical vector from the left yields the components of the physical vector (in the form of a column matrix) resolved in the reference frame $\mathcal{F}_a$.

2.1.6 Cross Product

Definition 2.7. Given two physical vectors $\underrightarrow{u} \in \mathbb{P}$ and $\underrightarrow{v} \in \mathbb{P}$, the *cross product* (also called the *vector product*) between the two physical vectors is defined as

$$\underrightarrow{w} = \underrightarrow{u} \times \underrightarrow{v} = ||\underrightarrow{u}||_2\, ||\underrightarrow{v}||_2 \sin\theta\, \underrightarrow{n},$$

where $\theta \in [0, 180^\circ]$ is the angle between $\underrightarrow{u}$ and $\underrightarrow{v}$, and $\underrightarrow{n} \in \mathbb{P}$ is a unit vector in the direction orthogonal to both $\underrightarrow{u}$ and $\underrightarrow{v}$ in a dextral sense (i.e., according to the right-hand-rule). The magnitude of the cross product of $\underrightarrow{u}$ and $\underrightarrow{v}$ is

$$||\underrightarrow{w}||_2 = ||\underrightarrow{u}||_2\, ||\underrightarrow{v}||_2 \sin\theta.$$

Let $\underrightarrow{u}, \underrightarrow{v}, \underrightarrow{w} \in \mathbb{P}$ and $a \in \mathbb{R}$. The cross product satisfies the following properties [4, pp. 8], [5, pp. 639-640].

1. $$\underrightarrow{u} \times \underrightarrow{u} = \underrightarrow{0} \quad \forall \underrightarrow{u} \in \mathbb{P}.$$

2. (Antisymmetry.)
$$\underrightarrow{u} \times \underrightarrow{v} = -\underrightarrow{v} \times \underrightarrow{u}, \quad \forall \underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}.$$

3. (Distributivity.)
$$(\underrightarrow{u} + \underrightarrow{v}) \times \underrightarrow{w} = \underrightarrow{u} \times \underrightarrow{w} + \underrightarrow{v} \times \underrightarrow{w}, \quad \forall \underrightarrow{u}, \underrightarrow{v}, \underrightarrow{w} \in \mathbb{P}.$$

4. (Scalar multiplication.)

$$a(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}) = \underset{\rightarrow}{u} \times (a\underset{\rightarrow}{v}) = (a\underset{\rightarrow}{u}) \times \underset{\rightarrow}{v}, \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}, \quad \forall a \in \mathbb{R}.$$

5. (Scalar triple product.)

$$\underset{\rightarrow}{u} \cdot \left(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}\right) = \left(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}\right) \cdot \underset{\rightarrow}{w} = \underset{\rightarrow}{v} \cdot \left(\underset{\rightarrow}{w} \times \underset{\rightarrow}{u}\right), \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}.$$

6. (Vector triple product.)

$$\underset{\rightarrow}{u} \times \left(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}\right) = \underset{\rightarrow}{v}(\underset{\rightarrow}{w} \cdot \underset{\rightarrow}{u}) - \underset{\rightarrow}{w}(\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v}), \quad \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}.$$

Notice that the cross product does not satisfy the properties of commutativity and associativity, which means that $\underset{\rightarrow}{u} \times \underset{\rightarrow}{v} \neq \underset{\rightarrow}{v} \times \underset{\rightarrow}{u}$, and $\underset{\rightarrow}{u} \times (\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}) \neq (\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}) \times \underset{\rightarrow}{w}, \forall \underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}$.

2.1.6.1 The Cross Product and Its Relation to Vectrices and Components

Consider two physical vectors $\underset{\rightarrow}{u}, \underset{\rightarrow}{v} \in \mathbb{P}$ expressed in terms of the vectrix $\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}$, as in Equations (2.3) and (2.4). Because the physical basis vectors $\underset{\rightarrow}{a}^1$, $\underset{\rightarrow}{a}^2$, and $\underset{\rightarrow}{a}^3$ are themselves physical vectors it follows that

$$\begin{array}{lll} \underset{\rightarrow}{a}^1 \times \underset{\rightarrow}{a}^1 = \underset{\rightarrow}{0}, & \underset{\rightarrow}{a}^1 \times \underset{\rightarrow}{a}^2 = \underset{\rightarrow}{a}^3, & \underset{\rightarrow}{a}^1 \times \underset{\rightarrow}{a}^3 = -\underset{\rightarrow}{a}^2, \\ \underset{\rightarrow}{a}^2 \times \underset{\rightarrow}{a}^1 = -\underset{\rightarrow}{a}^3, & \underset{\rightarrow}{a}^2 \times \underset{\rightarrow}{a}^2 = \underset{\rightarrow}{0}, & \underset{\rightarrow}{a}^2 \times \underset{\rightarrow}{a}^3 = \underset{\rightarrow}{a}^1, \\ \underset{\rightarrow}{a}^3 \times \underset{\rightarrow}{a}^1 = \underset{\rightarrow}{a}^2, & \underset{\rightarrow}{a}^3 \times \underset{\rightarrow}{a}^2 = -\underset{\rightarrow}{a}^1, & \underset{\rightarrow}{a}^3 \times \underset{\rightarrow}{a}^3 = \underset{\rightarrow}{0}. \end{array}$$

Now consider the cross product of $\underrightarrow{u}$ and $\underrightarrow{v}$ [2], [1, pp. 523-526], that being

$$
\begin{aligned}
\underrightarrow{u} \times \underrightarrow{v} &= \left(\mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a\right) \times \left(\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a\right) \\
&= \mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a \times \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix} \times \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \mathbf{v}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \begin{bmatrix} \underrightarrow{a}^1 \times \underrightarrow{a}^1 & \underrightarrow{a}^1 \times \underrightarrow{a}^2 & \underrightarrow{a}^1 \times \underrightarrow{a}^3 \\ \underrightarrow{a}^2 \times \underrightarrow{a}^1 & \underrightarrow{a}^2 \times \underrightarrow{a}^2 & \underrightarrow{a}^2 \times \underrightarrow{a}^3 \\ \underrightarrow{a}^3 \times \underrightarrow{a}^1 & \underrightarrow{a}^3 \times \underrightarrow{a}^2 & \underrightarrow{a}^3 \times \underrightarrow{a}^3 \end{bmatrix} \mathbf{v}_a \\
&= \begin{bmatrix} u_{a1} & u_{a2} & u_{a3} \end{bmatrix} \begin{bmatrix} \underrightarrow{0} & \underrightarrow{a}^3 & -\underrightarrow{a}^2 \\ -\underrightarrow{a}^3 & \underrightarrow{0} & \underrightarrow{a}^1 \\ \underrightarrow{a}^2 & -\underrightarrow{a}^1 & \underrightarrow{0} \end{bmatrix} \mathbf{v}_a \\
&= \begin{bmatrix} \left(-\underrightarrow{a}^3(u_{a2}) + \underrightarrow{a}^2(u_{a3})\right) & \left(\underrightarrow{a}^3(u_{a1}) - \underrightarrow{a}^1(u_{a3})\right) & \left(-\underrightarrow{a}^2(u_{a1}) + \underrightarrow{a}^1(u_{a2})\right) \end{bmatrix} \mathbf{v}_a \\
&= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & -u_{a3} & u_{a2} \\ u_{a3} & 0 & -u_{a1} \\ -u_{a2} & u_{a1} & 0 \end{bmatrix}}_{\mathbf{u}_a^{\times}} \mathbf{v}_a \\
&= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{v}_a, \qquad (2.6)
\end{aligned}
$$

where

$$
\mathbf{u}_a^{\times} = \begin{bmatrix} 0 & -u_{a3} & u_{a2} \\ u_{a3} & 0 & -u_{a1} \\ -u_{a2} & u_{a1} & 0 \end{bmatrix}
$$

is the *cross operator* [2], [1, pp. 523-526]. To derive the expression given in Equation (2.6), namely $\underrightarrow{u} \times \underrightarrow{v} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{v}_a$, both matrix operations and vector operations have been used. First, a matrix operation was used to expand $\underrightarrow{\mathcal{F}}_a \times \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}$, then the cross products of $\underrightarrow{a}^1 \times \underrightarrow{a}^1 = \underrightarrow{0}$, $\underrightarrow{a}^1 \times \underrightarrow{a}^2 = \underrightarrow{a}^3$, $\underrightarrow{a}^1 \times \underrightarrow{a}^3 = -\underrightarrow{a}^2$, etc. were evaluated.

2.1.6.2 The Cross and Uncross Operators

In forthcoming sections and chapters, the cross operator defined in Equation (2.7) will be an invaluable, indispensable, absolutely critical tool. *Always use it*! The cross operator reduces the task of evaluating the cross product of two physical vectors to simple matrix multiplication. Other texts also use the cross operator, but with different notation. For example, $[\mathbf{u}_a \times]$, $[\mathbf{u}_a]$, $[\tilde{\mathbf{u}}_a]$, $\{\mathbf{u}_a\} \otimes$, $-S(\mathbf{u}_a)$, $\hat{\mathbf{u}}_a$, $\mathbf{u}_a^{\wedge}$, and $\tilde{\mathbf{u}}_a$ can be found in other texts [5, 7, 19–24].

Let's now look at the specifics of the cross operator.

Definition 2.8. The *cross operator* is defined as

$$
\mathbf{u}_a^{\times} = \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix}^{\times} = \begin{bmatrix} 0 & -u_{a3} & u_{a2} \\ u_{a3} & 0 & -u_{a1} \\ -u_{a2} & u_{a1} & 0 \end{bmatrix}, \quad \forall \mathbf{u}_a = \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix} \in \mathbb{R}^3. \qquad (2.7)
$$

The matrix $\mathbf{u}_a^\times \in \mathbb{R}^{3\times3}$ is a *skewsymmetric* (also called *antisymmetric*) matrix, which means that

$$\mathbf{u}_a^{\times^\mathsf{T}} = -\mathbf{u}_a^\times, \quad \forall \mathbf{u}_a \in \mathbb{R}^3.$$

The vector space of skewsymmetric matrices in $\mathbb{R}^{3\times3}$ is [21, 22]

$$\mathfrak{so}(3) = \left\{ \mathbf{S} \in \mathbb{R}^{3\times3} \mid \mathbf{S}^\mathsf{T} = -\mathbf{S} \right\}.$$

The cross operator is a mapping from $\mathbb{R}^3$ to $\mathfrak{so}(3)$. Mathematically $(\cdot)^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$ is written. Notice that $\det \mathbf{u}_a^\times = 0$ meaning that $\mathbf{u}_a^\times$ does not have full rank, and is thus singular. As such, the inverse of $\mathbf{u}_a^\times$ cannot be computed. Also, notice that

$$\mathbf{u}_a^\times \mathbf{u}_a = \begin{bmatrix} 0 & -u_{a3} & u_{a2} \\ u_{a3} & 0 & -u_{a1} \\ -u_{a2} & u_{a1} & 0 \end{bmatrix} \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix} = \begin{bmatrix} -u_{a3}u_{a2} + u_{a2}u_{a3} \\ u_{a3}u_{a1} - u_{a1}u_{a3} \\ -u_{a2}u_{a1} + u_{a1}u_{a2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0},$$

as it should owing to the fact that $\underset{\rightarrow}{u} \times \underset{\rightarrow}{u} = \underset{\rightarrow}{0}$. Additionally,

$$\mathbf{u}_a^\times \mathbf{v}_a = -\mathbf{v}_a^\times \mathbf{u}_a, \quad \forall \mathbf{u}_a, \mathbf{v}_a \in \mathbb{R}^3, \tag{2.8}$$

which follows from the fact that $\underset{\rightarrow}{u} \times \underset{\rightarrow}{v} = -\underset{\rightarrow}{v} \times \underset{\rightarrow}{u}$.

Related to the cross operator is the *uncross operator* [21, 22].

Definition 2.9. The *uncross operator* is defined as

$$\mathbf{U}_a^\vee = \begin{bmatrix} 0 & -u_{a3} & u_{a2} \\ u_{a3} & 0 & -u_{a1} \\ -u_{a2} & u_{a1} & 0 \end{bmatrix}^\vee = \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix} = \mathbf{u}_a, \quad \forall \mathbf{U}_a \in \mathfrak{so}(3). \tag{2.9}$$

Given Definitions 2.8 and 2.9 it follows that

$$\mathbf{u}_a^{\times^\vee} = \mathbf{u}_a, \quad \forall \mathbf{u}_a \in \mathbb{R}^3.$$

The uncross operator is a mapping from $\mathfrak{so}(3)$ to $\mathbb{R}^3$. Mathematically $(\cdot)^\vee : \mathfrak{so}(3) \to \mathbb{R}^3$ is written.

Readers not familiar with skewsymmetric matrices might find it helpful to know that any 3×3 matrix, such as $\mathbf{Q} \in \mathbb{R}^{3\times3}$ (or any $n \times n$ matrix for that matter) can be decomposed into symmetric and skewsymmetric parts using the symmetric and antisymmetric projection operators:

$$\mathbf{Q} = \tfrac{1}{2}\mathbf{Q} + \tfrac{1}{2}\mathbf{Q} = \tfrac{1}{2}\mathbf{Q} + \tfrac{1}{2}\mathbf{Q} + \tfrac{1}{2}\mathbf{Q}^\mathsf{T} - \tfrac{1}{2}\mathbf{Q}^\mathsf{T} = \underbrace{\tfrac{1}{2}\left(\mathbf{Q} + \mathbf{Q}^\mathsf{T}\right)}_{\boldsymbol{\mathcal{P}}_s(\mathbf{Q})} + \underbrace{\tfrac{1}{2}\left(\mathbf{Q} - \mathbf{Q}^\mathsf{T}\right)}_{\boldsymbol{\mathcal{P}}_a(\mathbf{Q})}$$

where the symmetric and antisymmetric projection operators are $\boldsymbol{\mathcal{P}}_s(\mathbf{Q}) = \frac{1}{2}(\mathbf{Q} + \mathbf{Q}^\mathsf{T})$ and $\boldsymbol{\mathcal{P}}_a(\mathbf{Q}) = \frac{1}{2}(\mathbf{Q} - \mathbf{Q}^\mathsf{T})$, respectively.

2.1.6.3 Useful Identities

There are many very useful identities that accompany the cross and uncross operators. Here are a few of them.

$$
\begin{aligned}
\mathbf{u}^\times \mathbf{v} &= -\mathbf{v}^\times \mathbf{u}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3 \\
-\mathbf{u}^\times \mathbf{v}^\times &= \mathbf{u}^\mathsf{T}\mathbf{v}\mathbf{1} - \mathbf{v}\mathbf{u}^\mathsf{T}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3 \\
\left(\mathbf{u}^\times \mathbf{v}\right)^\times &= \mathbf{u}^\times \mathbf{v}^\times - \mathbf{v}^\times \mathbf{u}^\times = -\mathbf{u}\mathbf{v}^\mathsf{T} + \mathbf{v}\mathbf{u}^\mathsf{T}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3 \\
(\mathbf{u} + \mathbf{v})^\times &= \mathbf{u}^\times + \mathbf{v}^\times, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3 \\
(a\mathbf{u})^\times &= a\mathbf{u}^\times, \quad \forall a \in \mathbb{R}, \forall \mathbf{u} \in \mathbb{R}^3 \\
\tfrac{1}{2}\mathrm{tr}\left(\mathbf{v}^\times \mathbf{Q}\right) &= -\mathbf{v}^\mathsf{T}\boldsymbol{\mathcal{P}}_a(\mathbf{Q})^\vee, \quad \forall \mathbf{v} \in \mathbb{R}^3, \forall \mathbf{Q} \in \mathbb{R}^{3\times 3}.
\end{aligned}
$$

The subscript "a" have been neglected because these identities hold for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ resolved in any frame.

2.1.6.4 Some Examples

Example 2.3. Consider the physical vectors $\underset{\rightarrow}{u} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{u}_a \in \mathbb{P}$ and $\underset{\rightarrow}{v} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{v}_a \in \mathbb{P}$ where $\mathbf{u}_a = [1 \;\; 2 \;\; 1]^\mathsf{T} \in \mathbb{R}^3$ and $\mathbf{v}_a = [2 \;\; 1 \;\; 2]^\mathsf{T} \in \mathbb{R}^3$. Compute $\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v}$, $||\underset{\rightarrow}{u}||_2$, and $\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}$.

Solution.

$$
\begin{aligned}
\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{v} &= \mathbf{u}_a^\mathsf{T}\mathbf{v}_a = \mathbf{v}_a^\mathsf{T}\mathbf{u}_a = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = 6, \\
||\underset{\rightarrow}{u}||_2 &= \sqrt{\underset{\rightarrow}{u} \cdot \underset{\rightarrow}{u}} = \sqrt{\mathbf{u}_a^\mathsf{T}\mathbf{u}_a} = \sqrt{6}, \\
\underset{\rightarrow}{u} \times \underset{\rightarrow}{v} &= \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{u}_a^\times\mathbf{v}_a = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T} \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T} \begin{bmatrix} -1+4 \\ 2-2 \\ -4+1 \end{bmatrix} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T} \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}.
\end{aligned}
$$

□

Example 2.4. Consider the physical vectors $\underset{\rightarrow}{u} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{u}_a \in \mathbb{P}$, $\underset{\rightarrow}{v} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{v}_a \in \mathbb{P}$, and $\underset{\rightarrow}{w} = \underset{\rightarrow}{\mathcal{F}}_a^\mathsf{T}\mathbf{w}_a \in \mathbb{P}$ and the scalar triple product $\underset{\rightarrow}{u} \cdot \left(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}\right)$. Show that

$$
\underset{\rightarrow}{u} \cdot \left(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}\right) = \underset{\rightarrow}{w} \cdot \left(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}\right) = \left(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}\right) \cdot \underset{\rightarrow}{w} = \underset{\rightarrow}{v} \cdot \left(\underset{\rightarrow}{w} \times \underset{\rightarrow}{u}\right). \tag{2.10}
$$

Solution. First note that the transpose of a scalar is just the original scalar, that is $a^\mathsf{T} = a$, $\forall a \in \mathbb{R}$. Now, to show the equality of part of Equation (2.10) consider the following:

$$
\begin{aligned}
\underset{\rightarrow}{u} \cdot \left(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}\right) &= \mathbf{u}_a^{\mathsf{T}} \underbrace{\underset{\rightarrow}{\mathcal{F}}_a \cdot \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}}}_{\mathbf{1}} \mathbf{v}_a^{\times} \mathbf{w}_a \\
&= \mathbf{u}_a^{\mathsf{T}} \mathbf{v}_a^{\times} \mathbf{w}_a \quad \text{(This is a scalar; it can be transposed.)} \\
&= (\mathbf{v}_a^{\times} \mathbf{w}_a)^{\mathsf{T}} \mathbf{u}_a \\
&= \mathbf{w}_a^{\mathsf{T}} \mathbf{v}_a^{\times\mathsf{T}} \mathbf{u}_a \\
&= -\mathbf{w}_a^{\mathsf{T}} \mathbf{v}_a^{\times} \mathbf{u}_a \quad \text{(Use Equation (2.8).)} \\
&= \mathbf{w}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{v}_a \quad \text{(Put } \mathbf{1} = \underset{\rightarrow}{\mathcal{F}}_a \cdot \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \text{ in between the "}\mathbf{w}_a^{\mathsf{T}}\text{" and the "}\mathbf{u}_a^{\times}\text{".)} \\
&= \underset{\rightarrow}{w} \cdot \left(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}\right) \quad \text{(Use the communicatively property of the dot product.)} \\
&= \left(\underset{\rightarrow}{u} \times \underset{\rightarrow}{v}\right) \cdot \underset{\rightarrow}{w} \\
&= \left(\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{v}_a\right) \cdot \left(\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{w}_a\right) \\
&= \left((\mathbf{u}_a^{\times} \mathbf{v}_a)^{\mathsf{T}} \underset{\rightarrow}{\mathcal{F}}_a\right) \cdot \left(\underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \mathbf{w}_a\right) \\
&= -\mathbf{v}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{w}_a \quad \text{(Use Equation (2.8) once more.)} \\
&= \mathbf{v}_a^{\mathsf{T}} \mathbf{w}_a^{\times} \mathbf{u}_a \\
&= \underset{\rightarrow}{v} \cdot \left(\underset{\rightarrow}{w} \times \underset{\rightarrow}{u}\right).
\end{aligned}
$$

□

It can also be shown that $\det [\mathbf{u}_a \;\; \mathbf{v}_a \;\; \mathbf{w}_a] = \mathbf{u}_a^{\mathsf{T}} \mathbf{v}_a^{\times} \mathbf{w}_a$ where $[\mathbf{u}_a \;\; \mathbf{v}_a \;\; \mathbf{w}_a] \in \mathbb{R}^{3\times 3}$.

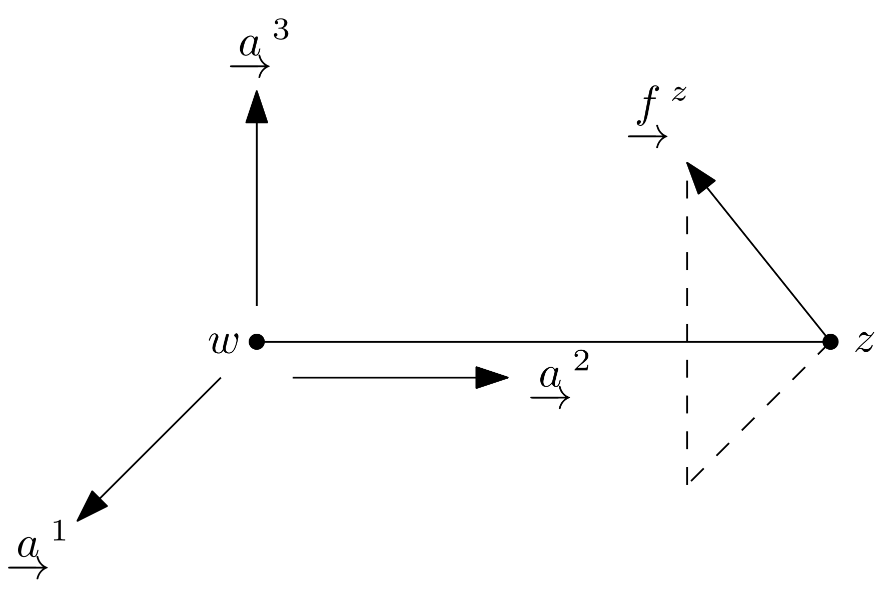


Figure 2.6: A statics problem.

Example 2.5. Consider Figure 2.6 where a $2\ (m)$ beam is subjected to the force $\underset{\rightarrow}{f}^z = \underset{\rightarrow}{a}^1(1) + \underset{\rightarrow}{a}^2(0) + \underset{\rightarrow}{a}^3(3)\ (N)$. What's the moment on point z relative to point w due to $\underset{\rightarrow}{f}^z$?

Solution. The moment on point z relative to point w due to $\underrightarrow{f}^{z}$ is $\underrightarrow{m}^{zw} = \underrightarrow{r}^{zw} \times \underrightarrow{f}^{z}$ where

$$\underrightarrow{r}^{zw} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}}_{\mathbf{r}_a^{zw}} (m), \quad \underrightarrow{f}^{z} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}}_{\mathbf{f}_a^{z}} (N).$$

It follows that

$$\begin{aligned} \underrightarrow{m}^{zw} &= \underrightarrow{r}^{zw} \times \underrightarrow{f}^{z} \\ &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{r}_a^{zw^\times} \mathbf{f}_a^{z} \\ &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \\ &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} 6 \\ 0 \\ -2 \end{bmatrix}}_{\mathbf{m}_a^{zw}} (N \cdot m) \\ &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{m}_a^{zw} \end{aligned}$$

□

2.2 Tensors

Physical vectors are in fact a special kind of *tensor*. Like physical vectors, tensors are independent of a particular reference frame and coordinate system. In physical space, which is three-dimensional, the number of components a tensor may have is 3^N, where N is the *order* or *rank* of the tensor. A tensor of *order zero* (i.e., $N = 0$) is specified by *one* component; these are simply scalars. Physical quantities that have magnitude only, such as mass, are represented by scalars. Tensors of *order one*, or *first-order* tensors, have three components; these are physical vectors, which have magnitude and direction. Tensors of *order two*, or *second-order* tensors [25, pp. 4-5], are now discussed.

Definition 2.10. A *second-order tensor* $\underrightarrow{T} : \mathbb{P} \rightarrow \mathbb{P}$ is a linear operator such that [4, pp. 10], [26, pp. 648-649]

$$\underrightarrow{T} \cdot \underrightarrow{u} \in \mathbb{P}, \quad \forall \underrightarrow{u} \in \mathbb{P}$$

A second-order tensor is a linear map or linear transformation from one physical vector to another physical vector [4, pp. 10] [26, pp. 648-649]. Second-order tensors are also called *dyadics* [2], [1, pp. 526-527], [25, pp. 4-5], [27, pp. 307-308]. To be concise, when referring to a second-order tensor, “tensor” or “dyadic” will be written for simplicity.

Owing to the fact that any tensor is a linear operator, a tensor such as $\underrightarrow{T} : \mathbb{P} \rightarrow \mathbb{P}$ satisfies the following properties [4, pp. 10].

1. (Distributivity.)

$$\underrightarrow{T} \cdot \left(\underrightarrow{u} + \underrightarrow{v} \right) = \underrightarrow{T} \cdot \underrightarrow{v} + \underrightarrow{T} \cdot \underrightarrow{u}, \quad \forall \underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}.$$

2. (Scalar multiplication.)

$$\underset{\rightarrow}{T} \cdot \left(a\,\underset{\rightarrow}{u}\right) = a\,\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}, \quad \forall a \in \mathbb{R}.$$

3. (Zero tensor.) $\exists \underset{\rightarrow}{O} : \mathbb{P} \to \mathbb{P}$ such that

$$\underset{\rightarrow}{O} \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{0}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

4. (Identity tensor.) $\exists \underset{\rightarrow}{I} : \mathbb{P} \to \mathbb{P}$ such that

$$\underset{\rightarrow}{I} \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

Now consider three tensors, $\underset{\rightarrow}{T} : \mathbb{P} \to \mathbb{P}$, $\underset{\rightarrow}{S} : \mathbb{P} \to \mathbb{P}$, $\underset{\rightarrow}{R} : \mathbb{P} \to \mathbb{P}$. The product of two tensors, such as $\underset{\rightarrow}{T}$ and $\underset{\rightarrow}{S}$, is $\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}$. Together the tensors satisfy the following properties [26, pp. 648-650].

1. $$\left(\underset{\rightarrow}{T} + \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{S} + \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{S} \cdot \underset{\rightarrow}{u} + \underset{\rightarrow}{T} \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

2. $$\left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{T} \cdot \left(\underset{\rightarrow}{S} \cdot \underset{\rightarrow}{u}\right), \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

3. $$\left(\left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{R}\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{T} \cdot \left(\underset{\rightarrow}{S} \cdot \underset{\rightarrow}{R}\right)\right) \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

4. $$\left(\underset{\rightarrow}{T} \cdot \left(\underset{\rightarrow}{R} + \underset{\rightarrow}{S}\right)\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{R} + \underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

5. $$\left(\left(\underset{\rightarrow}{R} + \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{R} \cdot \underset{\rightarrow}{T} + \underset{\rightarrow}{S} \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

6. $$\left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{I}\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{I} \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{T} \cdot \underset{\rightarrow}{u}, \quad \forall \underset{\rightarrow}{u} \in \mathbb{P}.$$

In genearl, $\left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{u} \neq \left(\underset{\rightarrow}{S} \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u}$. However, if $\left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{S}\right) \cdot \underset{\rightarrow}{u} = \left(\underset{\rightarrow}{S} \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u}$ the two tensors are said to *commute*.

Definition 2.11. A tensor, such as $\underset{\rightarrow}{T} : \mathbb{P} \to \mathbb{P}$, is said to be *invertible* when there is another tensor called the *inverse* of $\underset{\rightarrow}{T}$, denoted $\underset{\rightarrow}{T}^{-1} : \mathbb{P} \to \mathbb{P}$, such that [4, pp. 11], [26, pp. 650-651]

$$\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{T}^{-1} = \underset{\rightarrow}{T}^{-1} \cdot \underset{\rightarrow}{T} = \underset{\rightarrow}{I}.$$

Consider $\underset{\rightarrow}{T} : \mathbb{P} \to \mathbb{P}$ and $\underset{\rightarrow}{u} \in \mathbb{P}$ such that $\underset{\rightarrow}{v} = \underset{\rightarrow}{T} \cdot \underset{\rightarrow}{u} \in \mathbb{P}$. Using the definition of the inverse of $\underset{\rightarrow}{T}$ it follows that [26, pp. 650-651]

$$\underset{\rightarrow}{T}^{-1} \cdot \underset{\rightarrow}{v} = \underset{\rightarrow}{T}^{-1} \cdot \left(\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{u}\right) = \left(\underset{\rightarrow}{T}^{-1} \cdot \underset{\rightarrow}{T}\right) \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{I} \cdot \underset{\rightarrow}{u} = \underset{\rightarrow}{u}.$$

2.2.1 Direct Product Between Physical vectors

Definition 2.12. Consider two physical vectors, $\underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}$. The *direct product*, *tensor product*, or *dyadic product* of $\underrightarrow{u}$ and $\underrightarrow{v}$ is [4, pp. 11]

$$\underrightarrow{T} = \underrightarrow{u}\,\underrightarrow{v},$$

where $\underrightarrow{T} : \mathbb{P} \to \mathbb{P}$ is the tensor that results from the direct product of $\underrightarrow{u}$ and $\underrightarrow{v}$. The direct product forms a *dyad*, which means "a pair" [2], [1, pp. 526-527], [25, pp. 4-5], [27, pp. 307-308].

2.2.1.1 The Direct Product, Tensors, and Their Relation to Vectrices and Components

Consider two physical vectors, $\underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}$, and the direct product of $\underrightarrow{u}$ and $\underrightarrow{v}$, that is $\underrightarrow{T} = \underrightarrow{u}\,\underrightarrow{v}$. Also consider a reference frame $\mathcal{F}_a$ and associated vectrix $\underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} = \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix}$ so that $\underrightarrow{u} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a$, $\underrightarrow{v} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{v}_a$. The direct product of the two physical vectors can be concisely written as

$$\underrightarrow{T} = \underrightarrow{u}\,\underrightarrow{v} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a\mathbf{v}_a^{\mathsf{T}}\underrightarrow{\boldsymbol{\mathcal{F}}}_a = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{T}_a\underrightarrow{\boldsymbol{\mathcal{F}}}_a,$$

where $\mathbf{T}_a = \mathbf{u}_a\mathbf{v}_a^{\mathsf{T}} \in \mathbb{R}^{3\times 3}$ is a matrix of scalars representing the tensor $\underrightarrow{T}$ expressed in reference frame $\mathcal{F}_a$.

Now consider any tensor $\underrightarrow{T} : \mathbb{P} \to \mathbb{P}$, a reference frame $\mathcal{F}_a$, and associated vectrix $\underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} = \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix}$. In terms of the reference frame $\mathcal{F}_a$ and vectrix $\underrightarrow{\boldsymbol{\mathcal{F}}}_a$ the tensor $\underrightarrow{T}$ can be written as

$$\underrightarrow{T} = \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \underbrace{\begin{bmatrix} t_{a11} & t_{a12} & t_{a13} \\ t_{a21} & t_{a22} & t_{a23} \\ t_{a31} & t_{a32} & t_{a33} \end{bmatrix}}_{\mathbf{T}_a} \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{T}_a\underrightarrow{\boldsymbol{\mathcal{F}}}_a, \tag{2.11}$$

where $\mathbf{T}_a \in \mathbb{R}^{3\times 3}$ is a matrix of scalars representing the tensor $\underrightarrow{T}$ expressed in reference frame $\mathcal{F}_a$. Equation (2.11) can be expanded, leading to

$$\begin{aligned}
\underrightarrow{T} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{T}_a\underrightarrow{\boldsymbol{\mathcal{F}}}_a \\
&= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \begin{bmatrix} t_{a11} & t_{a12} & t_{a13} \\ t_{a21} & t_{a22} & t_{a23} \\ t_{a31} & t_{a32} & t_{a33} \end{bmatrix} \begin{bmatrix} \underrightarrow{a}^1 \\ \underrightarrow{a}^2 \\ \underrightarrow{a}^3 \end{bmatrix} \\
&= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \begin{bmatrix} \underrightarrow{a}^1(t_{a11}) + \underrightarrow{a}^2(t_{a12}) + \underrightarrow{a}^3(t_{a13}) \\ \underrightarrow{a}^1(t_{a21}) + \underrightarrow{a}^2(t_{a22}) + \underrightarrow{a}^3(t_{a23}) \\ \underrightarrow{a}^1(t_{a31}) + \underrightarrow{a}^2(t_{a32}) + \underrightarrow{a}^3(t_{a33}) \end{bmatrix} \\
&= \underrightarrow{a}^1\underrightarrow{a}^1(t_{a11}) + \underrightarrow{a}^1\underrightarrow{a}^2(t_{a12}) + \underrightarrow{a}^1\underrightarrow{a}^3(t_{a13}) \\
&\quad + \underrightarrow{a}^2\underrightarrow{a}^1(t_{a21}) + \underrightarrow{a}^2\underrightarrow{a}^2(t_{a22}) + \underrightarrow{a}^2\underrightarrow{a}^3(t_{a23}) \\
&\quad + \underrightarrow{a}^3\underrightarrow{a}^1(t_{a31}) + \underrightarrow{a}^3\underrightarrow{a}^2(t_{a32}) + \underrightarrow{a}^3\underrightarrow{a}^3(t_{a33}).
\end{aligned} \tag{2.12}$$

Equation (2.12) can be written as

$$\underrightarrow{T} = \sum_{i=1}^{3}\sum_{j=1}^{3} \underrightarrow{a}^i\underrightarrow{a}^j(t_{aij}).$$

Notice that the tensor $\underrightarrow{T}$ has been written as linear combination of dyads, which is why tensors are also called *dyadics*.

From Equation (2.12) it should be clear that $\underset{\rightarrow}{T}$ is not a matrix of scalars. On the other hand, $\mathbf{T}_a \in \mathbb{R}^{3\times 3}$ in Equation (2.11) *is* a matrix of scalars. In particular, $\mathbf{T}_a$ a 3×3 matrix of scalars representing $\underset{\rightarrow}{T}$ resolved in frame $\mathcal{F}_a$ with physical basis vectors $\underset{\rightarrow}{a}^1$, $\underset{\rightarrow}{a}^2$, and $\underset{\rightarrow}{a}^3$.

2.2.2 Dot and Cross Product Operations Involving Dyads

Let $\underset{\rightarrow}{T} = \underset{\rightarrow}{u}\,\underset{\rightarrow}{v}$ and $\underset{\rightarrow}{w}$ where $\underset{\rightarrow}{u}, \underset{\rightarrow}{v}, \underset{\rightarrow}{w} \in \mathbb{P}$. Then [2], [1, pp. 526-527], [25, pp. 4-5], [27, pp. 307-308]

$$\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{w} = (\underset{\rightarrow}{u}\,\underset{\rightarrow}{v}) \cdot \underset{\rightarrow}{w} = \underset{\rightarrow}{u}(\underset{\rightarrow}{v} \cdot \underset{\rightarrow}{w}),$$

and

$$\underset{\rightarrow}{w} \cdot \underset{\rightarrow}{T} = \underset{\rightarrow}{w} \cdot (\underset{\rightarrow}{u}\,\underset{\rightarrow}{v}) = (\underset{\rightarrow}{w} \cdot \underset{\rightarrow}{u})\underset{\rightarrow}{v}.$$

Notice that both $\underset{\rightarrow}{T} \cdot \underset{\rightarrow}{w}$ and $\underset{\rightarrow}{w} \cdot \underset{\rightarrow}{T}$ yield physical vectors. Similarly, [2], [1, pp. 526-527], [25, pp. 4-5], [27, pp. 307-308]

$$\underset{\rightarrow}{T} \times \underset{\rightarrow}{w} = (\underset{\rightarrow}{u}\,\underset{\rightarrow}{v}) \times \underset{\rightarrow}{w} = \underset{\rightarrow}{u}(\underset{\rightarrow}{v} \times \underset{\rightarrow}{w}),$$

and

$$\underset{\rightarrow}{w} \times \underset{\rightarrow}{T} = \underset{\rightarrow}{w} \times (\underset{\rightarrow}{u}\,\underset{\rightarrow}{v}) = (\underset{\rightarrow}{w} \times \underset{\rightarrow}{u})\underset{\rightarrow}{v}.$$

Here, the results of $\underset{\rightarrow}{T} \times \underset{\rightarrow}{w}$ and $\underset{\rightarrow}{w} \times \underset{\rightarrow}{T}$ yield tensors (well, dyads really).

2.2.3 An Example

Example 2.6. Consider two physical vectors, $\underset{\rightarrow}{u} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a \in \mathbb{P}$ and $\underset{\rightarrow}{v} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{v}_a \in \mathbb{P}$, and their cross product, $\underset{\rightarrow}{w} = \underset{\rightarrow}{u} \times \underset{\rightarrow}{v}$. Find the tensor, $\underset{\rightarrow}{T} : \mathbb{P} \rightarrow \mathbb{P}$, representing this cross-product operation [4, pp. 13].

Solution. Consider the cross production of $\underset{\rightarrow}{u}$ and $\underset{\rightarrow}{v}$, that being

$$\begin{aligned} \underset{\rightarrow}{w} &= \underset{\rightarrow}{u} \times \underset{\rightarrow}{v} \\ &= \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a^{\times}\mathbf{v}_a. \end{aligned}$$

In between $\mathbf{u}_a^{\times}$ and $\mathbf{v}_a$ place $\mathbf{1} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a \cdot \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}$ to get

$$\begin{aligned} \underset{\rightarrow}{w} &= \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a^{\times}\mathbf{1}\mathbf{v}_a \\ &= \underbrace{\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a^{\times}\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a}_{\underset{\rightarrow}{T}} \cdot \underbrace{\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{v}_a}_{\underset{\rightarrow}{v}} \\ &= \underset{\rightarrow}{T} \cdot \underset{\rightarrow}{v}. \end{aligned}$$

As such, $\underset{\rightarrow}{T} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{u}_a^{\times}\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a$ is the tensor representation of the cross product.

□

2.3 Many Reference Frames

2.3.1 Labelling Convention When Using Many Reference Frames

Consider two reference frames, reference frames "a" and "b", denoted $\mathcal{F}_a$ and $\mathcal{F}_b$, respectively. The physical basis vectors forming the orthonormal triads associated with $\mathcal{F}_a$ and $\mathcal{F}_b$ are $\underset{\rightarrow}{a}^1$, $\underset{\rightarrow}{a}^2$, $\underset{\rightarrow}{a}^3$ and $\underset{\rightarrow}{b}^1$, $\underset{\rightarrow}{b}^2$, $\underset{\rightarrow}{b}^3$,

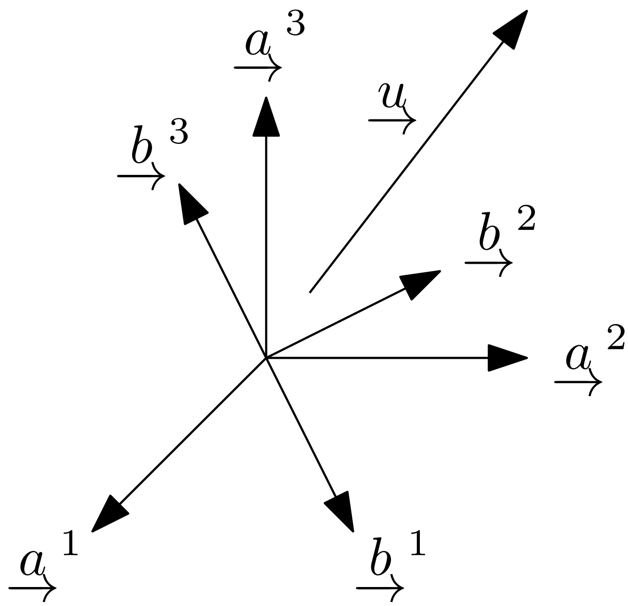


Figure 2.7: Two reference fames, $\mathcal{F}_a$ and $\mathcal{F}_b$, and a generic physical vector $\underrightarrow{u}$. Although the tails of the physical basis vectors are drawn such that their tails are collocated at a point, in general, the physical basis vectors composing $\mathcal{F}_a$ and $\mathcal{F}_b$ do not have a physical location.

respectively, where $\underrightarrow{a}^i, \underrightarrow{b}^i \in \mathbb{P}$, $i = 1, 2, 3$. The two reference frames are not oriented so that $\underrightarrow{a}^1$ is parallel to $\underrightarrow{b}^1$, $\underrightarrow{a}^2$ is parallel to $\underrightarrow{b}^2$, and $\underrightarrow{a}^3$ is parallel to $\underrightarrow{b}^3$. Rather, the reference frames are oriented differently, as shown in Figure 2.7.

Readers may be tempted to ask why the physical basis vectors of $\mathcal{F}_a$ are $\underrightarrow{a}^1$, $\underrightarrow{a}^2$, $\underrightarrow{a}^3$, and not $\underrightarrow{i}$, $\underrightarrow{j}$, and $\underrightarrow{k}$? To answer that question, let's ask another question: if the physical basis vectors of $\mathcal{F}_a$ are $\underrightarrow{i}$, $\underrightarrow{j}$, and $\underrightarrow{k}$, what are the physical basis vectors of $\mathcal{F}_b$ going to be? Suppose $\underrightarrow{i}'$, $\underrightarrow{j}'$, and $\underrightarrow{k}'$ are used. Now what if there was a frame $\mathcal{F}_c$? Are the physical basis vectors of $\mathcal{F}_c$ to be labeled $\underrightarrow{i}''$, $\underrightarrow{j}''$, and $\underrightarrow{k}''$? Clearly, if multiple reference frames are needed, it's best to use a labeling convention that is clear, concise, and consistent. For this reason the "$\underrightarrow{i}$, $\underrightarrow{j}$, and $\underrightarrow{k}$" labeling convention is not used. The physical basis vectors of $\mathcal{F}_\chi$, where $\chi = a, b, c, d, e$, etc., will be $\underrightarrow{\chi}^i \in \mathbb{P}$, $i = 1, 2, 3$. The same labeling convention can be found in [1, 2, 18].

2.3.2 Resolving a Physical Vector in Different Reference Frames

Consider Figure 2.7 once more, and the physical vector $\underrightarrow{u} \in \mathbb{P}$. Now that there are two reference frames, $\underrightarrow{u}$ can be resolved in either $\mathcal{F}_a$ or $\mathcal{F}_b$ [1, pp. 527] [2], that is

$$\underrightarrow{u} = \underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3}) = \underbrace{\begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix}}_{\underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}} \underbrace{\begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix}}_{\mathbf{u}_a} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{u}_a, \tag{2.13}$$

or

$$\underset{\rightarrow}{u} = \underset{\rightarrow}{b}^1(u_{b1}) + \underset{\rightarrow}{b}^2(u_{b2}) + \underset{\rightarrow}{b}^3(u_{b3}) = \underbrace{\begin{bmatrix} \underset{\rightarrow}{b}^1 & \underset{\rightarrow}{b}^2 & \underset{\rightarrow}{b}^3 \end{bmatrix}}_{\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}}} \underbrace{\begin{bmatrix} u_{b1} \\ u_{b2} \\ u_{b3} \end{bmatrix}}_{\mathbf{u}_b} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{u}_b, \tag{2.14}$$

where $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a$ and $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b$ are the vectrices associated with $\mathcal{F}_a$ and $\mathcal{F}_b$, and $\mathbf{u}_a$ and $\mathbf{u}_b$ are the *components* of the physical vector $\underset{\rightarrow}{u}$ *resolved in*, or *expressed in*, reference frames $\mathcal{F}_a$ and $\mathcal{F}_b$. Now, although the components of $\underset{\rightarrow}{u}$ are not equal in general, Equations (2.13) and (2.14) are both expressions for $\underset{\rightarrow}{u}$, and are therefore equivalent:

$$\underset{\rightarrow}{u} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{u}_a = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{u}_b. \tag{2.15}$$

One must be careful with Equation (2.15); Equation (2.15) does *not* mean that $\mathbf{u}_a \stackrel{?}{=} \mathbf{u}_b$. In fact, unless both $\mathcal{F}_a$ and $\mathcal{F}_b$ are aligned (so that they are effectively the same reference frame), in general $\mathbf{u}_a \neq \mathbf{u}_b$. What Equation (2.15) says is that a physical vector, such as $\underset{\rightarrow}{u}$, can be expressed in either $\mathcal{F}_a$ or $\mathcal{F}_b$, or any other reference frame one sees fit.

It is important to stress, once again, that physical vectors (as well as tensors) are entities that are entirely independent of a particular reference frame. Moreover, given the above discussion, it is important to understand that any physical vector can be expressed in any reference frame (as can any tensor).

2.3.2.1 An Example

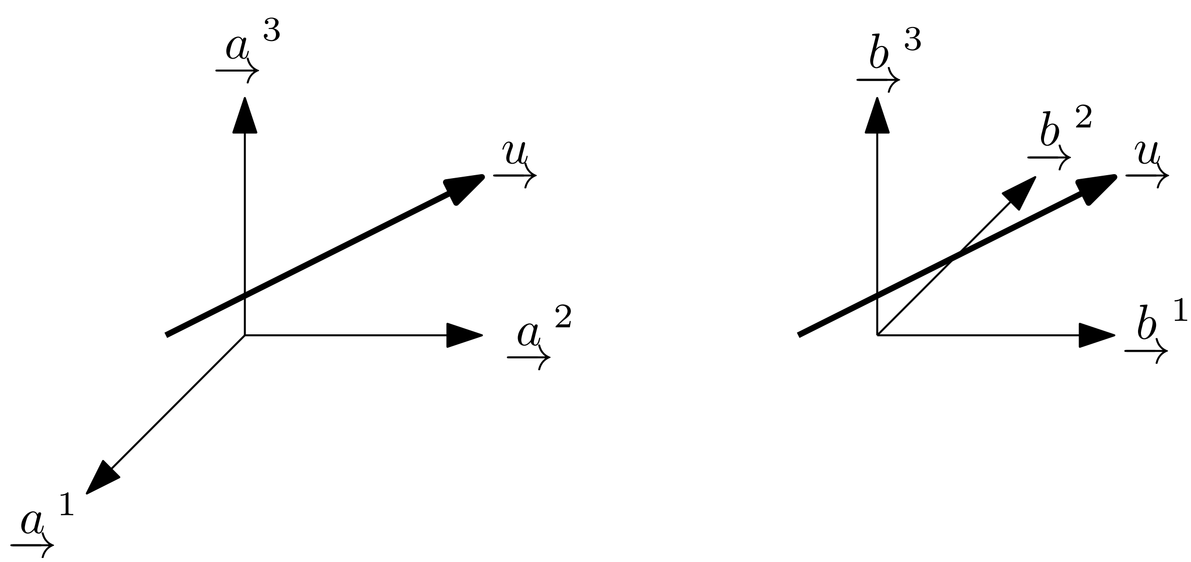


Figure 2.8: Two reference fames, $\mathcal{F}_a$ and $\mathcal{F}_b$, facing different directions.

Example 2.7. Consider Figure 2.8 where $\underset{\rightarrow}{u} \in \mathbb{P}$ is the same physical vector, and $\mathcal{F}_a$ and $\mathcal{F}_b$ are different reference frames. Express $\underset{\rightarrow}{u}$ in terms of both $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}$ and $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{b}^1 & \underset{\rightarrow}{b}^2 & \underset{\rightarrow}{b}^3 \end{bmatrix}$.

Solution.

$$
\begin{aligned}
\underrightarrow{u} &= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \overbrace{\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}}^{\mathbf{u}_a} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{u}_a, \\
\underrightarrow{u} &= \begin{bmatrix} \underrightarrow{b}^1 & \underrightarrow{b}^2 & \underrightarrow{b}^3 \end{bmatrix} \underbrace{\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{u}_b} = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{u}_b.
\end{aligned}
$$

The components of the physical vector $\underrightarrow{u}$ expressed in $\mathcal{F}_a$ are $\mathbf{u}_a$, while $\mathbf{u}_b$ are the components of the same vector expressed in $\mathcal{F}_b$.

□

2.4 Direction Cosine Matrices

2.4.1 The Direction Cosine Matrix $\mathbf{C}_{ba}$

Given a physical vector $\underrightarrow{u} \in \mathbb{P}$, the reference frames $\mathcal{F}_a$ and $\mathcal{F}_b$, and the components of $\underrightarrow{u}$ expressed in $\mathcal{F}_a$ and $\mathcal{F}_b$, $\mathbf{u}_a$ and $\mathbf{u}_b$, what's the relationship between $\mathbf{u}_a$ and $\mathbf{u}_b$? To answer this question recall Equation (2.15) [2] [1, pp. 527] [18, pp. 11-14],

$$
\begin{aligned}
\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{u}_b &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{u}_a, \\
\begin{bmatrix} \underrightarrow{b}^1 & \underrightarrow{b}^2 & \underrightarrow{b}^3 \end{bmatrix} \mathbf{u}_b &= \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix} \mathbf{u}_a.
\end{aligned}
$$

Taking the dot product from the left with $\underrightarrow{\boldsymbol{\mathcal{F}}}_b$ gives

$$
\begin{bmatrix} \underrightarrow{b}^1 \\ \underrightarrow{b}^2 \\ \underrightarrow{b}^3 \end{bmatrix} \cdot \begin{bmatrix} \underrightarrow{b}^1 & \underrightarrow{b}^2 & \underrightarrow{b}^3 \end{bmatrix} \mathbf{u}_b = \underbrace{\begin{bmatrix} \underrightarrow{b}^1 \\ \underrightarrow{b}^2 \\ \underrightarrow{b}^3 \end{bmatrix} \cdot \begin{bmatrix} \underrightarrow{a}^1 & \underrightarrow{a}^2 & \underrightarrow{a}^3 \end{bmatrix}}_{\underrightarrow{\boldsymbol{\mathcal{F}}}_b \cdot \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}} \mathbf{u}_a.
$$

As in Section 2.1.5 it follows that

$$
\begin{array}{lll}
\underrightarrow{b}^1 \cdot \underrightarrow{b}^1 = 1\,, & \underrightarrow{b}^1 \cdot \underrightarrow{b}^2 = 0\,, & \underrightarrow{b}^1 \cdot \underrightarrow{b}^3 = 0\,, \\
\underrightarrow{b}^2 \cdot \underrightarrow{b}^1 = 0\,, & \underrightarrow{b}^2 \cdot \underrightarrow{b}^2 = 1\,, & \underrightarrow{b}^2 \cdot \underrightarrow{b}^3 = 0\,, \\
\underrightarrow{b}^3 \cdot \underrightarrow{b}^1 = 0\,, & \underrightarrow{b}^3 \cdot \underrightarrow{b}^2 = 0\,, & \underrightarrow{b}^3 \cdot \underrightarrow{b}^3 = 1\,.
\end{array}
$$

Therefore,

$$
\begin{aligned}
\begin{bmatrix} \underrightarrow{b}^{1} \cdot \underrightarrow{b}^{1} & \underrightarrow{b}^{1} \cdot \underrightarrow{b}^{2} & \underrightarrow{b}^{1} \cdot \underrightarrow{b}^{3} \\ \underrightarrow{b}^{2} \cdot \underrightarrow{b}^{1} & \underrightarrow{b}^{2} \cdot \underrightarrow{b}^{2} & \underrightarrow{b}^{2} \cdot \underrightarrow{b}^{3} \\ \underrightarrow{b}^{3} \cdot \underrightarrow{b}^{1} & \underrightarrow{b}^{3} \cdot \underrightarrow{b}^{2} & \underrightarrow{b}^{3} \cdot \underrightarrow{b}^{3} \end{bmatrix} \mathbf{u}_b &= \begin{bmatrix} \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{3} \end{bmatrix} \mathbf{u}_a, \\
\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{1}} \mathbf{u}_b &= \underbrace{\begin{bmatrix} \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{1} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{2} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{1} & \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{2} & \underrightarrow{b}^{3} \cdot \underrightarrow{a}^{3} \end{bmatrix}}_{\mathbf{C}_{ba}} \mathbf{u}_a, \\
\mathbf{u}_b &= \mathbf{C}_{ba} \mathbf{u}_a, \qquad (2.16)
\end{aligned}
$$

where $\mathbf{C}_{ba}$ is the *direction cosine matrix* that describes the orientation of $\mathcal{F}_b$ relative to $\mathcal{F}_a$. Therefore,

$$\mathbf{C}_{ba} = \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}. \qquad (2.17)$$

Notice in Equation (2.16) that the subscripts all "line up". Specifically, when $\mathbf{u}_b = \mathbf{C}_{ba}\mathbf{u}_a$ is evaluated from right-to-left the subscript a on $\mathbf{u}_a$ lines up with the subscript a on $\mathbf{C}_{ba}$, then the subscript b on $\mathbf{C}_{ba}$ lines up with the subscript b on $\mathbf{u}_b$. Here is an example of how the notation is clear, concise, and consistent.

For a moment, return to Equation (2.16) and multiply both sides by $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}$ to yield

$$\underrightarrow{u} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba} \mathbf{u}_a.$$

Comparing the above expression to Equation (2.15),

$$\underrightarrow{u} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b,$$

gives an alternative way of relating $\mathbf{C}_{ba}$ to $\underrightarrow{\mathcal{F}}_a$ and $\underrightarrow{\mathcal{F}}_b$, that being

$$\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba}, \qquad (2.18)$$

which can also be written

$$\underrightarrow{\mathcal{F}}_a = \mathbf{C}_{ba}^{\mathsf{T}} \underrightarrow{\mathcal{F}}_b.$$

2.4.2 The Direction Cosine Matrix $\mathbf{C}_{ab}$

Similar to the derivation of $\mathbf{C}_{ba}$, an expression for $\mathbf{C}_{ab}$ can be derived. To do so, recall Equation (2.15),

$$
\begin{aligned}
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b, \\
\begin{bmatrix} \underrightarrow{a}^{1} & \underrightarrow{a}^{2} & \underrightarrow{a}^{3} \end{bmatrix} \mathbf{u}_a &= \begin{bmatrix} \underrightarrow{b}^{1} & \underrightarrow{b}^{2} & \underrightarrow{b}^{3} \end{bmatrix} \mathbf{u}_b.
\end{aligned}
$$

Taking the dot product from the left with $\underrightarrow{\mathcal{F}}_a$ gives

$$\begin{bmatrix} \underrightarrow{a}^{1} \\ \underrightarrow{a}^{2} \\ \underrightarrow{a}^{3} \end{bmatrix} \cdot \begin{bmatrix} \underrightarrow{a}^{1} & \underrightarrow{a}^{2} & \underrightarrow{a}^{3} \end{bmatrix} \mathbf{u}_a = \underbrace{\begin{bmatrix} \underrightarrow{a}^{1} \\ \underrightarrow{a}^{2} \\ \underrightarrow{a}^{3} \end{bmatrix} \cdot \begin{bmatrix} \underrightarrow{b}^{1} & \underrightarrow{b}^{2} & \underrightarrow{b}^{3} \end{bmatrix}}_{\underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}} \mathbf{u}_b.$$

Again, from Section 2.1.5,

$$\begin{array}{lll} \underrightarrow{a}^1 \cdot \underrightarrow{a}^1 = 1\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^1 \cdot \underrightarrow{a}^3 = 0\,, \\ \underrightarrow{a}^2 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^2 = 1\,, & \underrightarrow{a}^2 \cdot \underrightarrow{a}^3 = 0\,, \\ \underrightarrow{a}^3 \cdot \underrightarrow{a}^1 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^2 = 0\,, & \underrightarrow{a}^3 \cdot \underrightarrow{a}^3 = 1\,. \end{array}$$

It follows that

$$\begin{aligned} \begin{bmatrix} \underrightarrow{a}^1 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^1 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^1 \cdot \underrightarrow{a}^3 \\ \underrightarrow{a}^2 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^2 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^2 \cdot \underrightarrow{a}^3 \\ \underrightarrow{a}^3 \cdot \underrightarrow{a}^1 & \underrightarrow{a}^3 \cdot \underrightarrow{a}^2 & \underrightarrow{a}^3 \cdot \underrightarrow{a}^3 \end{bmatrix} \mathbf{u}_a &= \begin{bmatrix} \underrightarrow{a}^1 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^1 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^1 \cdot \underrightarrow{b}^3 \\ \underrightarrow{a}^2 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^2 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^2 \cdot \underrightarrow{b}^3 \\ \underrightarrow{a}^3 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^3 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^3 \cdot \underrightarrow{b}^3 \end{bmatrix} \mathbf{u}_b, \\ \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{1}} \mathbf{u}_a &= \underbrace{\begin{bmatrix} \underrightarrow{a}^1 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^1 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^1 \cdot \underrightarrow{b}^3 \\ \underrightarrow{a}^2 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^2 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^2 \cdot \underrightarrow{b}^3 \\ \underrightarrow{a}^3 \cdot \underrightarrow{b}^1 & \underrightarrow{a}^3 \cdot \underrightarrow{b}^2 & \underrightarrow{a}^3 \cdot \underrightarrow{b}^3 \end{bmatrix}}_{\mathbf{C}_{ab}} \mathbf{u}_b, \\ \mathbf{u}_a &= \mathbf{C}_{ab}\mathbf{u}_b, \end{aligned} \tag{2.19}$$

leading us to

$$\mathbf{C}_{ab} = \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}. \tag{2.20}$$

Multiplying both sides of Equation (2.19) by $\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}$ to gives

$$\underrightarrow{u} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{u}_a = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{C}_{ab}\mathbf{u}_b.$$

Comparing the above expression to Equation (2.15),

$$\underrightarrow{u} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{u}_a = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{u}_b,$$

an alternative way of relating $\mathbf{C}_{ba}$ to $\underrightarrow{\mathcal{F}}_a$ and $\underrightarrow{\mathcal{F}}_b$ is given by

$$\underrightarrow{\mathcal{F}}_b^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{C}_{ab}, \tag{2.21}$$

which can also be written

$$\underrightarrow{\mathcal{F}}_b = \mathbf{C}_{ab}^{\mathsf{T}}\underrightarrow{\mathcal{F}}_a.$$

2.4.3 The Relationship Between $\mathbf{C}_{ba}$ and $\mathbf{C}_{ab}$

Now, how are $\mathbf{C}_{ba}$ and $\mathbf{C}_{ab}$ related? Starting with Equation (2.18) and taking the dot product with $\underrightarrow{\mathcal{F}}_a$ gives

$$\begin{aligned} \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} &= \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{C}_{ba}, \\ \mathbf{1} &= \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{C}_{ba}. \end{aligned} \tag{2.22}$$

Substitution of Equation (2.20) into Equation (2.22) yields

$$\mathbf{C}_{ab}\mathbf{C}_{ba} = \mathbf{1}. \tag{2.23}$$

Alternatively, if both sides of Equation (2.17) are transposed,

$$\mathbf{C}_{ba}^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}},$$

which can be substituded into Equation (2.22) to get

$$\mathbf{C}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} = \mathbf{1}. \tag{2.24}$$

Together Equations (2.23) and (2.24) give

$$\mathbf{C}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} = \mathbf{C}_{ab}\mathbf{C}_{ba} = \mathbf{1},$$

leading to

$$\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab}. \tag{2.25}$$

Through similar arguments, it can be shown that

$$\mathbf{C}_{ba}\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ba}\mathbf{C}_{ab} = \mathbf{1}.$$

Although the relation $\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab}$ in Equation (2.25) is an interesting and useful one, there is yet another useful relation between $\mathbf{C}_{ab}$ and the inverse of $\mathbf{C}_{ba}$. First, assume the inverse of $\mathbf{C}_{ba}$ exits, which will formally be discussed in Section 2.4.4. Next, consider Equation (2.23), and postmultiply Equation (2.23) by $\mathbf{C}_{ba}^{-1}$ to yield

$$\begin{aligned} \mathbf{C}_{ab}\overbrace{\mathbf{C}_{ba}\mathbf{C}_{ba}^{-1}}^{\mathbf{1}} &= \mathbf{C}_{ba}^{-1}, \\ \mathbf{C}_{ab} &= \mathbf{C}_{ba}^{-1}. \end{aligned}$$

Combining this result with Equation (2.25) gives

$$\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab} = \mathbf{C}_{ba}^{-1}. \tag{2.26}$$

Thus, the relationship between $\mathbf{C}_{ba}$, $\mathbf{C}_{ba}^{\mathsf{T}}$, $\mathbf{C}_{ba}^{-1}$, and $\mathbf{C}_{ab}$ can be summarized as

$$\mathbf{C}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} = \mathbf{C}_{ba}^{-1}\mathbf{C}_{ba} = \mathbf{C}_{ab}\mathbf{C}_{ba} = \mathbf{1}. \tag{2.27}$$

2.4.4 Orthonormality of Direction Cosine Matrices

Definition 2.13. Consider $\mathbf{Q} \in \mathbb{R}^{n\times n}$. Denote the rows and columns of $\mathbf{Q}$ as, respectively, $\mathbf{q}_r^i$ and $\mathbf{q}_c^i$, $i = 1 \dots n$, so that

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_r^1 \\ \vdots \\ \mathbf{q}_r^n \end{bmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{q}_c^1 & \dots & \mathbf{q}_c^n \end{bmatrix}.$$

The matrix $\mathbf{Q}$ is said to be an *orthonormal* matrix if [2] [1, pp. 8-9] [18, pp. 11-14]

$$\mathbf{q}_r^i \mathbf{q}_r^{j\,\mathsf{T}} = \delta_{ij}, \quad i, j = 1, \dots, n,$$

and

$$\mathbf{q}_c^{i\,\mathsf{T}} \mathbf{q}_c^j = \delta_{ij}, \quad i, j = 1, \dots, n,$$

meaning that the matrix $\mathbf{Q}$ is not only an orthogonal matrix, but also that the rows and columns of the matrix $\mathbf{Q}$ are normalized (i.e., they have unit length). Furthermore, orthonormal matrices satisfy

$$\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{1}, \quad \mathbf{Q}\mathbf{Q}^{\mathsf{T}} = \mathbf{1}, \quad \mathbf{Q}^{\mathsf{T}} = \mathbf{Q}^{-1}, \quad \det(\mathbf{Q}^{\mathsf{T}}\mathbf{Q}) = \det(\mathbf{Q}\mathbf{Q}^{\mathsf{T}}) = \det{}^2\mathbf{Q} = 1, \quad \det\mathbf{Q} = \pm 1.$$

Based on the discussion in Section 2.4.3, and Equation (2.26) in particular, it should be clear that $\mathbf{C}_{ba}$, $\mathbf{C}_{ab}$, and all direction cosine matrices for that matter, are orthonormal. For instance, in Equation (2.27) is shown that $\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ba}^{-1}$.

In order to drive home the fact direction cosine matrices are orthonormal, first recall that $\underrightarrow{a}^{i}$, $\underrightarrow{b}^{i} \in \mathbb{P}$, $i = 1, 2, 3$ are physical vectors, and although they are used to composed $\mathcal{F}_a$ and $\mathcal{F}_b$, they too can be resolved in any frame. In particular,

$$\underrightarrow{a}^{i} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{1}_i = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{a}_b^i, \quad i = 1, 2, 3,$$

where $\mathbf{1}_1 = [1 \;\; 0 \;\; 0]^{\mathsf{T}}$, $\mathbf{1}_2 = [0 \;\; 1 \;\; 0]^{\mathsf{T}}$, and $\mathbf{1}_3 = [0 \;\; 0 \;\; 1]^{\mathsf{T}}$, and $\mathbf{a}_b^i$ are the components of the physical vector $\underrightarrow{a}^{i}$ resolved in the reference frame $\mathcal{F}_b$. Placing the three physical vectors $\underrightarrow{a}^{1}$, $\underrightarrow{a}^{2}$, and $\underrightarrow{a}^{3}$ in a vectrix (transposed) results in

$$\underbrace{\left[\begin{array}{ccc} \underrightarrow{a}^{1} & \underrightarrow{a}^{2} & \underrightarrow{a}^{3} \end{array} \right]}_{\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}} = \left[\begin{array}{ccc} \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{a}_b^1 & \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{a}_b^2 & \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{a}_b^3 \end{array} \right] = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left[\begin{array}{ccc} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{array} \right].$$

Taking the dot product from the left with $\underrightarrow{\mathcal{F}}_b$ gives

$$\underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left[\begin{array}{ccc} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{array} \right].$$

Comparing this result to Equation (2.17),

$$\mathbf{C}_{ba} = \left[\begin{array}{ccc} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{array} \right]. \tag{2.28}$$

Next, consider $\mathbf{C}_{ba}^{\mathsf{T}} \mathbf{C}_{ba} = \mathbf{1}$ written using Equation (2.28), that being

$$\mathbf{C}_{ab}^{\mathsf{T}} \mathbf{C}_{ba} = \left[\begin{array}{c} \mathbf{a}_b^{1\,\mathsf{T}} \\ \mathbf{a}_b^{2\,\mathsf{T}} \\ \mathbf{a}_b^{3\,\mathsf{T}} \end{array} \right] \left[\begin{array}{ccc} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{array} \right] = \left[\begin{array}{ccc} \mathbf{a}_b^{1\,\mathsf{T}} \mathbf{a}_b^1 & \mathbf{a}_b^{1\,\mathsf{T}} \mathbf{a}_b^2 & \mathbf{a}_b^{1\,\mathsf{T}} \mathbf{a}_b^3 \\ \mathbf{a}_b^{2\,\mathsf{T}} \mathbf{a}_b^1 & \mathbf{a}_b^{2\,\mathsf{T}} \mathbf{a}_b^2 & \mathbf{a}_b^{2\,\mathsf{T}} \mathbf{a}_b^3 \\ \mathbf{a}_b^{3\,\mathsf{T}} \mathbf{a}_b^1 & \mathbf{a}_b^{3\,\mathsf{T}} \mathbf{a}_b^2 & \mathbf{a}_b^{3\,\mathsf{T}} \mathbf{a}_b^3 \end{array} \right].$$

Using the fact that $\mathbf{a}_b^{i\,\mathsf{T}} \mathbf{a}_b^j = \mathbf{a}_b^{i\,\mathsf{T}} \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{a}_b^j = \underrightarrow{a}^{i} \cdot \underrightarrow{a}^{j}$, $i, j = 1, 2, 3$, it follows that

$$\mathbf{C}_{ab}^{\mathsf{T}} \mathbf{C}_{ba} = \left[\begin{array}{ccc} \underrightarrow{a}^{1} \cdot \underrightarrow{a}^{1} & \underrightarrow{a}^{1} \cdot \underrightarrow{a}^{2} & \underrightarrow{a}^{1} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{a}^{2} \cdot \underrightarrow{a}^{1} & \underrightarrow{a}^{2} \cdot \underrightarrow{a}^{2} & \underrightarrow{a}^{2} \cdot \underrightarrow{a}^{3} \\ \underrightarrow{a}^{3} \cdot \underrightarrow{a}^{1} & \underrightarrow{a}^{3} \cdot \underrightarrow{a}^{2} & \underrightarrow{a}^{3} \cdot \underrightarrow{a}^{3} \end{array} \right] = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right] = \mathbf{1},$$

clearly indicating $\mathbf{C}_{ab}^{\mathsf{T}} \mathbf{C}_{ba} = \mathbf{1}$ [18, pp. 11-14].

Now return to Equation (2.24). Taking the determinant of both sides and using the fact that $\forall \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{3 \times 3}$, $\det(\mathbf{PQ}) = \det \mathbf{P} \det \mathbf{Q}$, and $\det \mathbf{P}^{\mathsf{T}} = \det \mathbf{P}$, it follows that [18, pp. 11-14]

$$\det(\mathbf{C}_{ba}^{\mathsf{T}} \mathbf{C}_{ba}) = \det \mathbf{C}_{ba}^{\mathsf{T}} \det \mathbf{C}_{ba} = \det{}^2 \mathbf{C}_{ba} = 1.$$

Note that the right-hand side is equal to 1 because $\det \mathbf{1} = 1$. Now, taking the square root of both sides gives

$$\det \mathbf{C}_{ba} = \pm 1.$$

In fact, and as shown next, $\det \mathbf{C}_{ba} = +1$.

To show that $\det \mathbf{C}_{ba} = +1$, recall Equation (2.28),

$$\mathbf{C}_{ba} = \begin{bmatrix} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{bmatrix}. \tag{2.29}$$

It can be verified that

$$\det \mathbf{C}_{ba} = \mathbf{a}_b^{1\mathsf{T}} \mathbf{a}_b^{2\times} \mathbf{a}_b^3.$$

From Example (2.4), it is known that $\underrightarrow{u} \cdot \left(\underrightarrow{v} \times \underrightarrow{w} \right) = \mathbf{u}_b^{\mathsf{T}} \mathbf{v}_b^{\times} \mathbf{w}_b$ for any three physical vectors $\underrightarrow{u}$, $\underrightarrow{v}$, and $\underrightarrow{w}$. Using this result gives [18, pp. 11-14]

$$\det \mathbf{C}_{ba} = \mathbf{a}_b^{1\mathsf{T}} \mathbf{a}_b^{2\times} \mathbf{a}_b^3 = \underrightarrow{a}^1 \cdot \left(\underrightarrow{a}^2 \times \underrightarrow{a}^3 \right) = \underrightarrow{a}^1 \cdot \underrightarrow{a}^1 = +1, \tag{2.30}$$

where $\underrightarrow{a}^1 = \underrightarrow{a}^2 \times \underrightarrow{a}^3$ has been used to simplify. As such,

$$\det \mathbf{C}_{ba} = +1.$$

Recall that $\det(\mathbf{Q}) \neq 0$ means the matrix $\mathbf{Q}$ is invertible. Thus, direction cosine matrices are invertible because $\det \mathbf{C}_{ba} = +1$.

2.4.4.1 Definition of a Direction Cosine Matrix and Some Discussion

Based on Sections 2.4.1, 2.4.2, 2.4.3, and 2.4.4 the definition of the direction cosine matrix can now be formally given.

Definition 2.14. Consider two reference frames, $\mathcal{F}_a$ and $\mathcal{F}_b$, and associated vectrices, $\underrightarrow{\mathcal{F}}_a$ and $\underrightarrow{\mathcal{F}}_b$. The direction cosine matrix describing the orientation of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ is [2], [1, pp. 8,22]

$$\mathbf{C}_{ba} = \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}},$$

also written as

$$\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba},$$

or

$$\underrightarrow{\mathcal{F}}_b = \mathbf{C}_{ba} \underrightarrow{\mathcal{F}}_a,$$

where

$$\mathbf{C}_{ba}^{\mathsf{T}} \mathbf{C}_{ba} = \mathbf{C}_{ab} \mathbf{C}_{ab}^{\mathsf{T}} = \mathbf{1}, \quad \mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab} = \mathbf{C}_{ba}^{-1}, \quad \det \mathbf{C}_{ba} = +1.$$

A direction cosine matrix characterizes the attitude or orientation of one frame relative to another. As noted in Definition 2.14, the direction cosine matrix $\mathbf{C}_{ba}$ describes the orientation of $\mathcal{F}_b$ relative to $\mathcal{F}_a$.

Although the direction cosine matrix has 9 elements, for numerical integration purposes, during simulation, for example, it is possible to construct the full direction cosine matrix from only 6 elements. To see this, recall Equations (2.29) and (2.30),

$$\mathbf{C}_{ba} = \begin{bmatrix} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{bmatrix}, \quad \det \mathbf{C}_{ba} = \mathbf{a}_b^{1\mathsf{T}} \mathbf{a}_b^{2\times} \mathbf{a}_b^3.$$

Using the first and second column of $\mathbf{C}_{ba}$ the third can be computed via

$$\mathbf{a}_b^3 = \mathbf{a}_b^{1\times} \mathbf{a}_b^2,$$

which follows from the fact that $\underrightarrow{a}^3 = \underrightarrow{a}^1 \times \underrightarrow{a}^2$. Therefore, during numerical integration, or when using the direction cosine matrix for estimation or control on-board a vehicle, only 6 elements of the direction cosine matrix need to computed, and the remaining 3 elements can be computed as needed.

Note that $\mathbf{C}_{ba}$ does *not* "rotate" the physical vector $\underrightarrow{u}$. Rather, $\mathbf{C}_{ba}$ changes the point of view the physical vector $\underrightarrow{u}$ is being observed from. To see this, consider Equation (2.16) once again,

$$\mathbf{u}_b = \mathbf{C}_{ba}\mathbf{u}_a.$$

From right-to-left, start off with $\mathbf{u}_a$, the components of $\underrightarrow{u}$ resolved in $\mathcal{F}_a$. The direction cosine matrix $\mathbf{C}_{ba}$ is then multiplied by $\mathbf{u}_a$ yielding $\mathbf{u}_b$, the components of $\underrightarrow{u}$ resolved in $\mathcal{F}_b$.

Although $\mathbf{C}_{ba}$ does not rotate a physical vector such as $\underrightarrow{u}$, one interpretation of $\mathbf{C}_{ba}$ is that $\mathbf{C}_{ba}$ rotates the physical basis vectors of $\mathcal{F}_a$ thus yielding the physical basis vectors $\mathcal{F}_b$.

2.4.4.2 An Example

Example 2.8. Is the matrix

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

a valid direction cosine matrix?

Solution. First, is $\mathbf{Q}$ is orthonormal? Let

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}^1 & \mathbf{q}^2 & \mathbf{q}^3 \end{bmatrix},$$

an observer that

$$\begin{array}{lll} \mathbf{q}^{1\,\mathsf{T}}\mathbf{q}^1 = 1\,, & \mathbf{q}^{1\,\mathsf{T}}\mathbf{q}^2 = 0\,, & \mathbf{q}^{1\,\mathsf{T}}\mathbf{q}^3 = 0\,, \\ \mathbf{q}^{2\,\mathsf{T}}\mathbf{q}^1 = 0\,, & \mathbf{q}^{2\,\mathsf{T}}\mathbf{q}^2 = 1\,, & \mathbf{q}^{2\,\mathsf{T}}\mathbf{q}^3 = 0\,, \\ \mathbf{q}^{3\,\mathsf{T}}\mathbf{q}^1 = 0\,, & \mathbf{q}^{3\,\mathsf{T}}\mathbf{q}^2 = 0\,, & \mathbf{q}^{3\,\mathsf{T}}\mathbf{q}^3 = 1\,. \end{array}$$

The columns of the matrix $\mathbf{Q}$ have unit length and are orthogonal to each other; as such, yes, $\mathbf{Q}$ is orthonormal. Second, what's the determinant of $\mathbf{Q}$ equal to?

$$\begin{aligned} \det \mathbf{Q} &= \det \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{matrix} \\ &= \left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right)(1) + 0 + 0 - 0 - 0 - (1)\left(-\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right) \\ &= 1. \end{aligned}$$

Thus, $\mathbf{Q}$ is orthonormal and has determinant equal to $+1$. As such, $\mathbf{Q}$ is indeed a valid direction cosine matrix.

□

2.4.5 Compounding Direction Cosine Matrices

In the previous sections only two frames, $\mathcal{F}_a$ and $\mathcal{F}_b$, have been dealt with. Now consider many reference frames, say $\mathcal{F}_a$, $\mathcal{F}_b$, $\mathcal{F}_c$, ..., $\mathcal{F}_r$, $\mathcal{F}_s$, $\mathcal{F}_t$ [2], [1, pp. 19], [18, pp. 24,27]. As discussed previously, a physical vector, such as $\underrightarrow{u} \in \mathbb{P}$, can be resolved in any reference frame, that is

$$\underrightarrow{u} = \underrightarrow{\mathcal{F}}_t^{\mathsf{T}} \mathbf{u}_t = \underrightarrow{\mathcal{F}}_s^{\mathsf{T}} \mathbf{u}_s = \underrightarrow{\mathcal{F}}_r^{\mathsf{T}} \mathbf{u}_r = \cdots = \underrightarrow{\mathcal{F}}_c^{\mathsf{T}} \mathbf{u}_c = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a.$$

The relationship between the components of $\underrightarrow{u}$ in each frame is

$$\begin{aligned}
\mathbf{u}_b &= \mathbf{C}_{ba}\mathbf{u}_a, &(2.31)\\
\mathbf{u}_c &= \mathbf{C}_{cb}\mathbf{u}_b \ (= \mathbf{C}_{cb}\mathbf{C}_{ba}\mathbf{u}_a), &(2.32)\\
\mathbf{u}_d &= \mathbf{C}_{dc}\mathbf{u}_c \ (= \mathbf{C}_{dc}\mathbf{C}_{cb}\mathbf{C}_{ba}\mathbf{u}_a), &(2.33)\\
&\vdots \\
\mathbf{u}_r &= \mathbf{C}_{rq}\mathbf{u}_q, \\
\mathbf{u}_s &= \mathbf{C}_{sr}\mathbf{u}_r, \\
\mathbf{u}_t &= \mathbf{C}_{ts}\mathbf{u}_s.
\end{aligned}$$

What's the relationship between $\mathbf{C}_{ta}$ and $\mathbf{C}_{ba}$, $\mathbf{C}_{cb}$, $\mathbf{C}_{dc}$, ..., $\mathbf{C}_{rq}$, $\mathbf{C}_{sr}$, $\mathbf{C}_{ts}$? If Equation (2.31) is substituted into Equation (2.32), then that result and substitute it into Equation (2.33), and so on, it follows that

$$\mathbf{u}_t = \underbrace{\mathbf{C}_{ts}\mathbf{C}_{sr}\mathbf{C}_{rq}\cdots\mathbf{C}_{dc}\mathbf{C}_{cb}\mathbf{C}_{ba}}_{\mathbf{C}_{ta}}\mathbf{u}_a.$$

As such,

$$\mathbf{C}_{ta} = \mathbf{C}_{ts}\mathbf{C}_{sr}\mathbf{C}_{rq}\cdots\mathbf{C}_{dc}\mathbf{C}_{cb}\mathbf{C}_{ba}.$$

2.4.6 The Special Orthogonal Group of Rigid-Body Rotations

Direction cosine matrices are 3×3 matrices that are orthonormal and have determinate equal to positive 1. These properties define what is called the special orthogonal group of rigid-body rotations, denoted $SO(3)$.

Definition 2.15. The special orthogonal group of rigid-body rotations is defined by [21, pp. 250] [22, pp. 24]

$$SO(3) = \left\{ \mathbf{C} \in \mathbb{R}^{3\times 3} \ \mid \ \mathbf{C}^{\mathsf{T}}\mathbf{C} = \mathbf{1}, \ \det \mathbf{C} = +1 \right\}.$$

Note that two direction cosine matrices can be added together but the result is *not* another direction cosine matrix, that is

$$\mathbf{C}_{rq} + \mathbf{C}_{qp} \notin SO(3).$$

Given this result, $SO(3)$ is not a linear vector space. However, multiplying two direction cosine matrices does yield another direction cosine matrix, that is

$$\mathbf{C}_{rq}\mathbf{C}_{qp} \in SO(3).$$

In fact, $SO(3)$ is a *group* under the operation of matrix multiplication [22, pp. 24]. A set $\boldsymbol{G}$ with a binary operation $\circ$ defined on the elements of $\boldsymbol{G}$ is a group if the following properties hold.

1. (Closure.)
$$\boldsymbol{g}^1 \circ \boldsymbol{g}^2 \in \boldsymbol{G}, \quad \forall \boldsymbol{g}^1, \boldsymbol{g}^2 \in \boldsymbol{G}$$

2. (Identity.) There exists an identity element, $\mathbf{1}$, such that
$$\boldsymbol{g} \circ \mathbf{1} = \mathbf{1} \circ \boldsymbol{g}, \quad \forall \boldsymbol{g} \in \boldsymbol{G}.$$

3. (Inverse.) $\exists \boldsymbol{g}^{-1} \in \boldsymbol{G}$ such that

$$\boldsymbol{g} \circ \boldsymbol{g}^{-1} = \boldsymbol{g}^{-1} \circ \boldsymbol{g} = \mathbf{1}, \quad \forall \boldsymbol{g} \in \boldsymbol{G}.$$

4. (Associativity.)

$$\left(\boldsymbol{g}^1 \circ \boldsymbol{g}^2\right) \circ \boldsymbol{g}^3 = \boldsymbol{g}^1 \circ \left(\boldsymbol{g}^2 \circ \boldsymbol{g}^3\right), \quad \forall \boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3 \in \boldsymbol{G}.$$

In the case of $SO(3)$, where the binary operation is matrix multiplication, the following holds.

1. (Closure.)

$$\mathbf{C}_{rq}\mathbf{C}_{qp} \in SO(3), \quad \mathbf{C}_{rq}, \mathbf{C}_{qp} \in SO(3),$$

where

$$\begin{aligned} (\mathbf{C}_{rq}\mathbf{C}_{qp})^{\mathsf{T}}\mathbf{C}_{rq}\mathbf{C}_{qp} = \mathbf{C}_{qp}^{\mathsf{T}} \underbrace{\mathbf{C}_{rq}^{\mathsf{T}}\mathbf{C}_{rq}}_{\mathbf{1}} \mathbf{C}_{qp} = \mathbf{C}_{qp}^{\mathsf{T}}\mathbf{C}_{qp} &= \mathbf{1}, \\ \det(\mathbf{C}_{rq}\mathbf{C}_{qp}) = \det \mathbf{C}_{rq} \det \mathbf{C}_{qp} &= +1. \end{aligned}$$

2. (Identity.) The identity element is the 3×3 identity matrix $\mathbf{1}$.
3. (Inverse.) From Equation (2.26) the inverse of $\mathbf{C}_{qp} \in SO(3)$ is $\mathbf{C}_{qp}^{\mathsf{T}} = \mathbf{C}_{pq} \in SO(3)$.
4. (Associativity.) Matrix multiplication is associative, thus

$$\left(\mathbf{C}_{sr}\mathbf{C}_{rq}\right)\mathbf{C}_{qp} = \mathbf{C}_{sr}\left(\mathbf{C}_{rq}\mathbf{C}_{qp}\right).$$

2.4.7 The Eigenvalues of Direction Cosine Matrices

To glean further insight into the properties of the direction cosine matrix, as in [2], the eigenvalues of the direction cosine matrix will be computed. First recall that the complex conjugate of a complex number $\lambda = \sigma + \jmath\omega \in \mathbb{C}$ is $\lambda^{\mathsf{H}} = \sigma - \jmath\omega \in \mathbb{C}$, and the complex-conjugate transpose, or the Hermitian transpose, of a complex matrix $\mathbf{\Lambda} = \mathbf{\Sigma} + \jmath\mathbf{\Omega} \in \mathbb{C}^{n\times n}$ is $\mathbf{\Lambda}^{\mathsf{H}} = \mathbf{\Sigma}^{\mathsf{T}} - \jmath\mathbf{\Omega}^{\mathsf{T}} \in \mathbb{C}^{n\times n}$. Consider

$$\mathbf{C}_{qp}\mathbf{e}^i = \lambda_i\mathbf{e}^i, \quad i = 1, 2, 3,$$

where $\lambda_i \in \mathbb{C}$ is an eigenvalue of $\mathbf{C}_{qp}$ and $\mathbf{e}^i \in \mathbb{C}^3$ is an eigenvector. Rearranging,

$$\begin{aligned} \mathbf{e}^{i^{\mathsf{H}}}\mathbf{C}_{qp}^{\mathsf{T}} &= \lambda_i^{\mathsf{H}}\mathbf{e}^{i^{\mathsf{H}}}, \\ \mathbf{e}^{i^{\mathsf{H}}}\mathbf{C}_{qp}^{\mathsf{T}}\mathbf{C}_{qp} &= \lambda_i^{\mathsf{H}}\mathbf{e}^{i^{\mathsf{H}}}\mathbf{C}_{qp}, \\ \mathbf{e}^{i^{\mathsf{H}}}\underbrace{\mathbf{C}_{qp}^{\mathsf{T}}\mathbf{C}_{qp}}_{\mathbf{1}} &= \lambda_i^{\mathsf{H}}\mathbf{e}^{i^{\mathsf{H}}}\mathbf{C}_{qp}, \\ \mathbf{e}^{i^{\mathsf{H}}}\mathbf{e}^i &= \lambda_i^{\mathsf{H}}\mathbf{e}^{i^{\mathsf{H}}}\underbrace{\mathbf{C}_{qp}\mathbf{e}^i}_{\lambda_i\mathbf{e}^i}, \\ \mathbf{e}^{i^{\mathsf{H}}}\mathbf{e}^i &= \lambda_i^{\mathsf{H}}\lambda_i\mathbf{e}^{i^{\mathsf{H}}}\mathbf{e}^i, \\ \left(\lambda_i^{\mathsf{H}}\lambda_i - 1\right)\mathbf{e}^{i^{\mathsf{H}}}\mathbf{e}^i &= 0. \end{aligned} \tag{2.34}$$

where $\mathbf{e}^{i^{\mathsf{H}}}\mathbf{e}^i$ is a real number. For $\mathbf{e}^i \neq 0$ Equation (2.34) gives $\lambda_i^{\mathsf{H}}\lambda_i = 1$ meaning that the magnitude of each eigenvalue is equal to one. The eigenvalues are, in general, complex. Let $\lambda_i = x_i + y_i j = r_i e^{j\phi_i}$ where

$r_i e^{j\phi_i}$ is the polar form of $\lambda_i = x_i + y_i j$. In particular, $r_i = \sqrt{x_i^2 + y_i^2}$ and $\tan\phi_i = y_i/x_i$. It follows that $\lambda_i^{\mathsf{H}}\lambda_i = 1$ can be written

$$\begin{aligned} \lambda_i^{\mathsf{H}}\lambda_i &= 1, \\ r_i^2 e^{j\phi_i} e^{-j\phi_i} &= 1, \\ r_i^2 e^{j\phi_i - j\phi_i} &= 1, \\ r_i^2 &= 1. \end{aligned}$$

This results in the conclusion that the eigenvalues are $\lambda_1 = \pm 1, \lambda_2 = e^{j\phi}, \lambda_3 = e^{-j\phi}$. Note that the two complex eigenvalues are complex-conjugate pairs. In order to assess if $\lambda_1 = +1$ or $\lambda_1 = -1$ recall that

$$\det \mathbf{C}_{qp} = \prod_{i=1}^{3} \lambda_i = \lambda_1\lambda_2\lambda_3 = (\pm 1)(e^{j\phi})(e^{-j\phi}) = \pm 1.$$

By definition

$$\det \mathbf{C}_{qp} = +1,$$

and it follows that

$$\begin{aligned} \det \mathbf{C}_{qp} &= \lambda_1\lambda_2\lambda_3, \\ +1 &= \lambda_1 e^{j\phi} e^{-j\phi}, \end{aligned}$$

and hence $\lambda_1 = +1$. Therefore the eigenvalues of $\mathbf{C}_{qp}$ are $\lambda_1 = +1, \lambda_2 = e^{j\phi}, \lambda_3 = e^{-j\phi}$.

2.4.8 Direction Cosine Matrix Parameterizations

A direction cosine matrix, such as $\mathbf{C}_{qp}$, describes the attitude or orientation of $\mathcal{F}_q$ relative to $\mathcal{F}_p$ *globally* and *uniquely*. Despite the global and unique nature of direction cosine matrices, rotation matrix parameterization, which are all deficient in some way (i.e., they either are not a global attitude representation or are not a unique attitude representation), are quite popular for computations and use in estimation and control. Attitude parameterizations will now be discussed.

2.4.8.1 Axis/Angle Parameters and Euler's Theorem

The first direction cosine matrix parameterization discussed are axis/angle parameters [1, pp. 10-14]. To begin, recall that

$$\mathbf{C}_{qp}\mathbf{e}^1 = \lambda_1 \mathbf{e}^1,$$

where $\lambda_1 = +1$. Denoting the normalized eigenvector $\mathbf{e}^1$ as $\mathbf{a}$ (i.e., $\mathbf{e}^1 = \mathbf{a}$), notice that

$$\mathbf{C}_{qp}\mathbf{a} = \mathbf{a}. \tag{2.35}$$

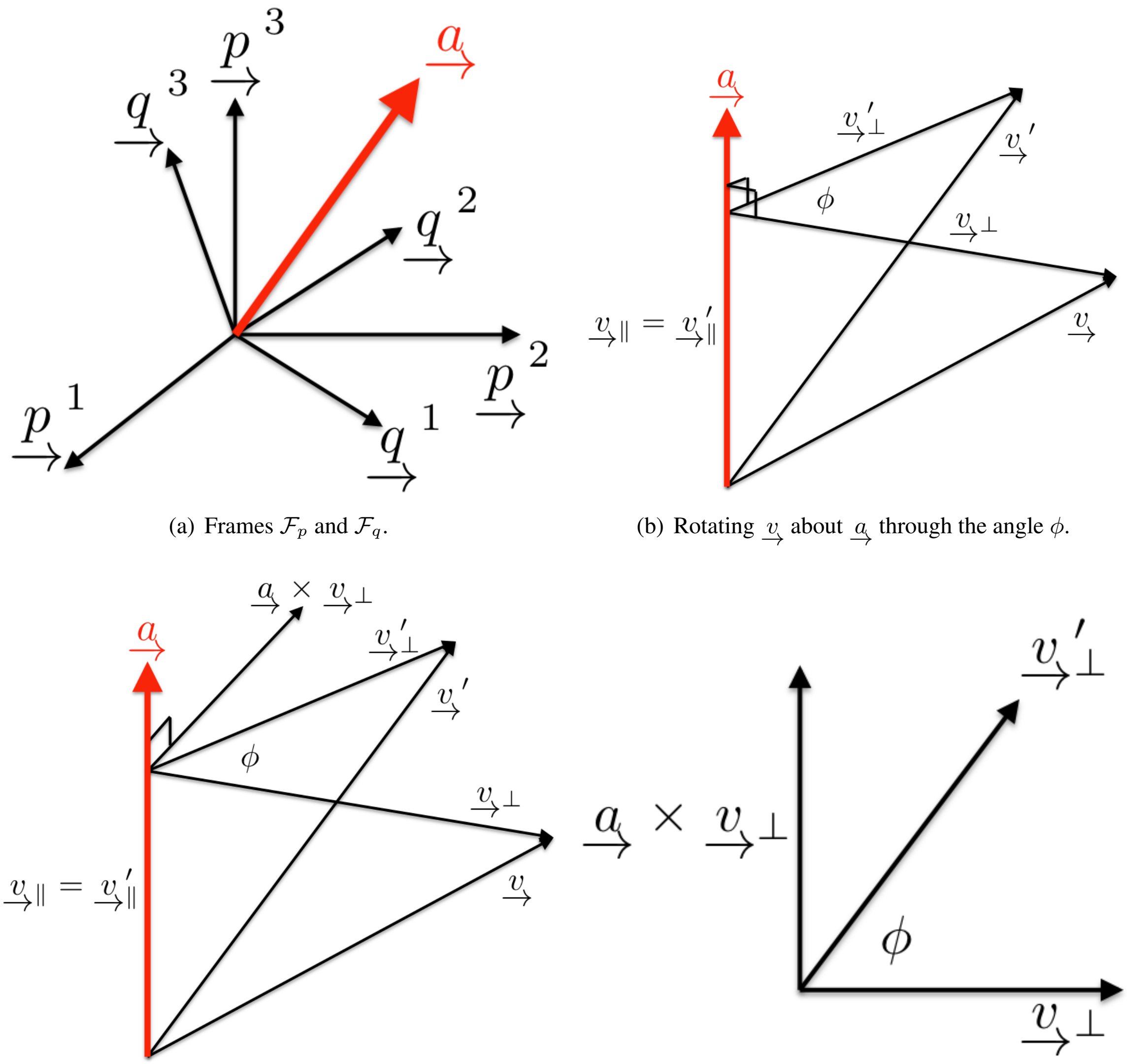
Recall that for any physical vector, say $\underrightarrow{v} \in \mathbb{P}$, that $\mathbf{v}_q = \mathbf{C}_{qp}\mathbf{v}_p$. Equation (2.35) says that the unit-length physical vector $\underrightarrow{a} \in \mathbb{P}$ resolved in either frame $\mathcal{F}_p$ or $\mathcal{F}_q$ is the same, that is

$$\mathbf{a}_p = \mathbf{a}_q = \mathbf{a}.$$

Because $\underrightarrow{a} \in \mathbb{P}$ is of unit length, $\mathbf{a} \in \mathbb{S}^2$ where $\mathbb{S}^2 = \left\{\mathbf{a} \in \mathbb{R}^3 \mid \sqrt{\mathbf{a}^{\mathsf{T}}\mathbf{a}} = 1\right\}$ is the unit two sphere.

Graphically the situation at hand is shown in Figure 2.9(a). Indeed, the unit-length physical vector $\underrightarrow{a}$

resolved in either frame $\mathcal{F}_p$ or $\mathcal{F}_q$ is the same. Another way of interpreting this result is that the frame $\mathcal{F}_p$ rotated about $\underrightarrow{a}$ by an angle ϕ yielding $\mathcal{F}_q$. As such, the physical vector $\underrightarrow{a}$ is called the axis of rotation. It is of unit length (i.e., $\underrightarrow{a} \cdot \underrightarrow{a} = 1$) because the original eigenvector was normalized. The scalar $\phi \in \mathbb{R}$ is called the angle of rotation.



(a) Frames $\mathcal{F}_p$ and $\mathcal{F}_q$.

(b) Rotating $\underrightarrow{v}$ about $\underrightarrow{a}$ through the angle ϕ.

(c) The cross product of $\underrightarrow{a}$ and $\underrightarrow{v}_\perp$.

(d) Looking down $\underrightarrow{a}$.

Figure 2.9: Rotation of frame $\mathcal{F}_p$ about $\underrightarrow{a}$ through an angle ϕ yielding $\mathcal{F}_q$.

Of interest presently is the relationship between $\mathbf{C}_{qp}$, $\mathbf{a}$, and ϕ. Consider an arbitrary physical vector $\underrightarrow{v}$ that represents any one of $\underrightarrow{p}^1$, $\underrightarrow{p}^2$, or $\underrightarrow{p}^3$, the physical basis vectors that compose $\mathcal{F}_p$, as shown in Figure 2.9(b). Consider the rotation of $\underrightarrow{v}$ about $\underrightarrow{a}$ through the angle ϕ. Call $\underrightarrow{v}'$ the rotated physical vector that represents $\underrightarrow{q}^1$, $\underrightarrow{q}^2$, or $\underrightarrow{q}^3$ corresponding to the rotated $\underrightarrow{p}^1$, $\underrightarrow{p}^2$, or $\underrightarrow{p}^3$, as shown in Figure 2.9(b).

The physical vector $\underrightarrow{v}$ can be decomposed into parallel and perpendicular parts using $\underrightarrow{a}$:

$$\begin{aligned}
\underrightarrow{v} &= \underrightarrow{v}_{\parallel} + \underrightarrow{v}_{\perp}, && (2.36)\\
\underrightarrow{v}_{\parallel} &= (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a}, && (2.37)\\
\underrightarrow{v}_{\perp} &= \underrightarrow{v} - (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a}. && (2.38)
\end{aligned}$$

The parallel part of $\underrightarrow{v}$ and $\underrightarrow{v}'$ are equal, thus

$$\underrightarrow{v}' = \underrightarrow{v}_{\parallel} + \underrightarrow{v}'_{\perp}.$$

As shown in Figure 2.9(d), by looking down $\underrightarrow{a}$ "from above" it follows that

$$\underrightarrow{v}'_{\perp} = \cos\phi\, \underrightarrow{v}_{\perp} + \sin\phi\, \underrightarrow{a} \times \underrightarrow{v}_{\perp}. \tag{2.39}$$

where $\underrightarrow{a} \times \underrightarrow{v}_{\perp}$ is visually shown in Figure 2.9(c). Therefore, using Equation (2.36), (2.37), (2.38), and (2.39), $\underrightarrow{v}'$ can be written

$$\begin{aligned}
\underrightarrow{v}' &= \underrightarrow{v}_{\parallel} + \underrightarrow{v}'_{\perp}\\
&= (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a} + \cos\phi\, \underrightarrow{v}_{\perp} + \sin\phi\, \underrightarrow{a} \times \underrightarrow{v}_{\perp}\\
&= (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a} + \cos\phi(\underrightarrow{v} - (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a}) + \sin\phi\, \underrightarrow{a} \times (\underrightarrow{v} - (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a})\\
&= (\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a} + \cos\phi\, \underrightarrow{v} - \cos\phi(\underrightarrow{a} \cdot \underrightarrow{v})\underrightarrow{a} + \sin\phi\, \underrightarrow{a} \times \underrightarrow{v} - \sin\phi(\underrightarrow{a} \cdot \underrightarrow{v})\underbrace{\underrightarrow{a} \times \underrightarrow{a}}_{\underrightarrow{0}}\\
&= \cos\phi\, \underrightarrow{v} + (1 - \cos\phi)\underrightarrow{a}(\underrightarrow{a} \cdot \underrightarrow{v}) + \sin\phi\, \underrightarrow{a} \times \underrightarrow{v}.
\end{aligned}$$

Thus,

$$\underrightarrow{v}' = \cos\phi\, \underrightarrow{v} + (1 - \cos\phi)\underrightarrow{a}(\underrightarrow{a} \cdot \underrightarrow{v}) + \sin\phi\, \underrightarrow{a} \times \underrightarrow{v}. \tag{2.40}$$

Recall that $\underrightarrow{v}$ represents one of $\underrightarrow{p}^1$, $\underrightarrow{p}^2$, or $\underrightarrow{p}^3$, and resolved in $\mathcal{F}_p$, $\underrightarrow{v}$ is

$$\underrightarrow{v} = \underrightarrow{\boldsymbol{\mathcal{F}}}_p^{\mathsf{T}} \mathbf{1}_i, \quad i = 1, 2, 3,$$

where $\mathbf{1}_1 = [1 \;\; 0 \;\; 0]^{\mathsf{T}}$, $\mathbf{1}_2 = [0 \;\; 1 \;\; 0]^{\mathsf{T}}$, and $\mathbf{1}_3 = [0 \;\; 0 \;\; 1]^{\mathsf{T}}$. Similarly, $\underrightarrow{v}'$ represents one of $\underrightarrow{q}^1$, $\underrightarrow{q}^2$, or $\underrightarrow{q}^3$ corresponding to the rotated $\underrightarrow{p}^1$, $\underrightarrow{p}^2$, or $\underrightarrow{p}^3$, and resolved in $\mathcal{F}_q$, $\underrightarrow{v}'$ is

$$\underrightarrow{v}' = \underrightarrow{\boldsymbol{\mathcal{F}}}_q^{\mathsf{T}} \mathbf{1}_i, \quad i = 1, 2, 3.$$

For a moment, say $\underrightarrow{v} = \underrightarrow{p}^1$; thus $\mathbf{v}_p = \mathbf{1}_1$. Because $\underrightarrow{v}'$ represents the rotated physical vector $\underrightarrow{q}^1$, $\underrightarrow{q}^2$, or $\underrightarrow{q}^3$ corresponding to the rotated $\underrightarrow{p}^1$, $\underrightarrow{p}^2$, or $\underrightarrow{p}^3$, in the case where $\underrightarrow{v} = \underrightarrow{p}^1$, $\underrightarrow{v}' = \underrightarrow{q}^1$ and therefore $\mathbf{v}'_q = \mathbf{1}_1$. Notice that $\mathbf{v}_p = \mathbf{v}'_q$ in this example; this holds true when $\underrightarrow{v} = \underrightarrow{p}^2$ and $\underrightarrow{v}' = \underrightarrow{q}^2$, and when $\underrightarrow{v} = \underrightarrow{p}^3$ and $\underrightarrow{v}' = \underrightarrow{q}^3$ as well. Therefore, it follows that

$$\mathbf{v}_p = \mathbf{v}'_q = \mathbf{v}.$$

Returning to Equation (2.40) and expressing the left-hand-side in $\mathcal{F}_q$ and the right-hand-side in $\mathcal{F}_q$ gives

$$\underrightarrow{\mathcal{F}}_q^{\mathsf{T}} \mathbf{v}'_q = \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \left(\cos\phi \mathbf{v}_p + (1 - \cos\phi) \mathbf{a}_p \mathbf{a}_p^{\mathsf{T}} \mathbf{v}_p + \sin\phi \mathbf{a}_p^{\times} \mathbf{v}_p \right). \tag{2.41}$$

Using the fact that $\mathbf{v} = \mathbf{v}_p = \mathbf{v}'_q$ and $\mathbf{a} = \mathbf{a}_p = \mathbf{a}_q$ Equation (2.41) can be written

$$\begin{aligned}
\underrightarrow{\mathcal{F}}_q^{\mathsf{T}} \mathbf{v} &= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \left(\cos\phi \mathbf{1} + (1 - \cos\phi) \mathbf{a}\mathbf{a}^{\mathsf{T}} + \sin\phi \mathbf{a}^{\times} \right) \mathbf{v}, \\
\overbrace{\underrightarrow{\mathcal{F}}_p \cdot \underrightarrow{\mathcal{F}}_q^{\mathsf{T}}}^{\mathbf{C}_{pq}} \mathbf{v} &= \overbrace{\underrightarrow{\mathcal{F}}_p \cdot \underrightarrow{\mathcal{F}}_p^{\mathsf{T}}}^{\mathbf{1}} \left(\cos\phi \mathbf{1} + (1 - \cos\phi) \mathbf{a}\mathbf{a}^{\mathsf{T}} + \sin\phi \mathbf{a}^{\times} \right) \mathbf{v}, \\
\mathbf{C}_{pq} \mathbf{v} &= \left(\cos\phi \mathbf{1} + (1 - \cos\phi) \mathbf{a}\mathbf{a}^{\mathsf{T}} + \sin\phi \mathbf{a}^{\times} \right) \mathbf{v},
\end{aligned}$$

where the dot product from the left with $\underrightarrow{\mathcal{F}}_p$ has been taken. Therefore,

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

However, of interest is $\mathbf{C}_{qp}$. Recalling that $\mathbf{C}_{qp} = \mathbf{C}_{pq}^{\mathsf{T}}$ it follows that

$$\mathbf{C}_{qp} = \cos\phi \mathbf{1} + (1 - \cos\phi) \mathbf{a}\mathbf{a}^{\mathsf{T}} - \sin\phi \mathbf{a}^{\times}, \tag{2.42}$$

the direction cosine matrix written in terms of $\mathbf{a}$ and ϕ.

The above derivation leads to what is known as Euler's Theorem.

Theorem 2.1 (Euler's Theorem)**.** The most general motion of a rigid body with one point fixed is a rotation about an axis through that point.

The derivation presented in this section is based on [1, 2, 18].

Is $(\mathbf{a}, \phi)$ a Global and Unique Attitude Representation?

There are no singularities associated with $(\mathbf{a}, \phi)$ in Equation (2.42). Specifically, there are no terms that go to infinity, and no terms that go to zero that are in the denominator of a fraction. Together $(\mathbf{a}, \phi)$ form a global attitude representation, meaning any $\mathbf{C} \in SO(3)$ can be represented by an $\mathbf{a}$ and an ϕ.

From Equation (2.42) it follows that for a given axis $\mathbf{a}$ and a given angle ϕ there is a unique $\mathbf{C}_{qp}$ [1, pp. 14]. However, given $\mathbf{C}_{qp}$ there is no unique $(\mathbf{a}, \phi)$. The axis $\mathbf{a}$ and $\phi \pm n2\pi$, $n \in \mathbb{Z}$, represents the exact same direction cosine matrix. Additionally, $(\mathbf{a}, \phi)$ and $(-\mathbf{a}, 2\pi - \phi)$ also represent the exact same direction cosine matrix.

It is tempting to say "Well, restrict ϕ such that $-\pi \leq 0 < \pi$, or $0 \leq \phi < 2\pi$". Doing so leads to a discontinuity, which from a simulation, estimation, and control point of view is not acceptable.

2.4.8.2 The Quaternion Parameterization

A quaternion are a very popular attitude representation. The quaternion is defined as

$$\mathbf{q} = \begin{bmatrix} \boldsymbol{\epsilon} \\ \eta \end{bmatrix} = \begin{bmatrix} \mathbf{a} \sin\left(\frac{\phi}{2}\right) \\ \cos\left(\frac{\phi}{2}\right) \end{bmatrix}, \tag{2.43}$$

where $\mathbf{a}$ is the axis of rotation and ϕ is the angle of rotation defined in Section 2.4.8.1. Quaternions, also called Euler parameters, are elements of the three sphere, $\mathbb{S}^3 = \left\{ \mathbf{q} \in \mathbb{R}^4 \mid \sqrt{\mathbf{q}^\mathsf{T}\mathbf{q}} = 1 \right\}$ [28]. It is straightforward to verify that the quaternion defined in Equation (2.43) is indeed an element of $\mathbb{S}^3$,

$$\sqrt{\mathbf{q}^\mathsf{T}\mathbf{q}} = \sqrt{\boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon} + \eta^2} = \sqrt{\mathbf{a}^\mathsf{T}\mathbf{a}\sin^2\left(\frac{\phi}{2}\right) + \cos^2\left(\frac{\phi}{2}\right)} = 1,$$

where $\mathbf{a}^\mathsf{T}\mathbf{a} = 1$ and $\sin^2\left(\frac{\phi}{2}\right) + \cos^2\left(\frac{\phi}{2}\right) = 1$.

The direction cosine matrix $\mathbf{C}_{qp}$ can be parameterized (i.e., written in terms of) the quaternion. To do so, first recall that

$$\begin{aligned}
\cos\phi &= \cos^2\left(\frac{\phi}{2}\right) - \sin^2\left(\frac{\phi}{2}\right) = 1 - 2\sin^2\left(\frac{\phi}{2}\right), \\
\sin\phi &= 2\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right).
\end{aligned}$$

Next, from Equation (2.42), consider $\mathbf{C}_{qp}$ parameterized using axis/angle variables rearranged as

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

where $\boldsymbol{\epsilon} = \mathbf{a}\sin\left(\frac{\phi}{2}\right)$ and $\eta = \cos\left(\frac{\phi}{2}\right)$. Thus,

$$\mathbf{C}_{qp} = \left(1 - 2\boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon}\right)\mathbf{1} + 2\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\mathsf{T} - 2\eta\boldsymbol{\epsilon}^\times. \tag{2.44}$$

Notice that there are no trigonometric functions in Equation (2.44). This is an advantage in terms of simulation, estimation, and control.

Is $\mathbf{q} = [\boldsymbol{\epsilon}^\mathsf{T} \;\; \eta]^\mathsf{T}$ a Global and Unique Attitude Representation?

In Equation (2.44) there are no singularities associated with $\boldsymbol{\epsilon}$ and η. Therefore, a quaternion can represent attitude globally [28].

Given an $\boldsymbol{\epsilon}$ and η there is a unique $\mathbf{C}_{qp}$. However, given $\mathbf{C}_{qp}$ there are two antipodal quaternions that both represent $\mathbf{C}_{qp}$, that is, $\mathbf{q} = [\boldsymbol{\epsilon}^\mathsf{T} \;\; \eta]^\mathsf{T}$ and $-\mathbf{q} = [-\boldsymbol{\epsilon}^\mathsf{T} \;\; -\eta]^\mathsf{T}$ represent the same $\mathbf{C}_{qp}$ [28]. As such, a quaternion is not a unique representation of attitude.

2.4.8.3 Rodrigues Parameters

Another attitude parameterization is the three-parameter set known as Rodrigues parameters that are defined as [29, pp. 48-50]

$$\mathbf{p} = \frac{\boldsymbol{\epsilon}}{\eta}. \tag{2.45}$$

It can be shown that, in terms of Rodrigues parameters, the direction cosine matrix $\mathbf{C}_{qp}$ can be written as [1, pp. 30-31] [29, pp. 48-50]

$$\mathbf{C}_{qp} = \mathbf{1} + \frac{2}{1 + \mathbf{p}^\mathsf{T}\mathbf{p}}\left(\mathbf{p}^\times\mathbf{p}^\times - \mathbf{p}^\times\right).$$

Are Rodrigues Parameters a Global and Unique Attitude Representation?

Equation (2.45) can alternatively be written as [1, pp. 30-31]

$$\mathbf{p} = \mathbf{a} \tan\left(\frac{\phi}{2}\right), \tag{2.46}$$

where $\mathbf{a} \in \mathbb{S}^2$ and ϕ are axis/angle parameters. Notice that as as $\phi \to \pm\pi$ that $\|\mathbf{p}\|_2 \to \infty$. As such, rotations of $\pm\pi$ about any axis $\mathbf{a}$ cannot be represented by $\mathbf{p}$. If the original definition of the Rodrigues parameters given in Equation (2.45) is used, the same conclusion can be draw: as $\eta \to 0$, $\|\mathbf{p}\|_2 \to \infty$. Therefore, Rodrigues parameters are not a global attitude representation.

Recall that a quaternion parameterization of attitude is not unique because $\mathbf{q} = [\boldsymbol{\epsilon}^\mathsf{T} \quad \eta]^\mathsf{T}$ and $-\mathbf{q} = [-\boldsymbol{\epsilon}^\mathsf{T} \quad -\eta]^\mathsf{T}$ represent the same attitude. Using Equation (2.45) it can be see that Rodrigues parameters represent attitude uniquely owing to the fact that

$$\frac{\boldsymbol{\epsilon}}{\eta} = \frac{-\boldsymbol{\epsilon}}{-\eta} = \mathbf{p}.$$

2.4.8.4 Modified Rodrigues Parameters

The most modern attitude representation is the so-called modified Rodrigues parameters (MRPs) defined as [29, pp. 50-52]

$$\mathbf{s} = \frac{\boldsymbol{\epsilon}}{1+\eta}. \tag{2.47}$$

MRPs were discovered in 1962 by T. F. Wiener. MRPs were rediscovered in 1982 by Milenkovic (who referred to MRPs as the conformal rotation vector) [30], and then again in 1987 by Marandi and Modi [29, pp. 50-52]. In terms of MRPs the direction cosine matrix $\mathbf{C}_{qp}$ can be written as [29, pp. 51]

$$\mathbf{C}_{qp} = \mathbf{1} + \frac{4}{(1+\mathbf{s}^\mathsf{T}\mathbf{s})^2}\left(2\mathbf{s}^\times\mathbf{s}^\times - (1-\mathbf{s}^\mathsf{T}\mathbf{s})\mathbf{s}^\times\right).$$

Are Modified Rodrigues Parameters a Global and Unique Attitude Representation?

The expression for the MRP $\mathbf{s}$ given in Equation (2.47) is equivalent to

$$\mathbf{s} = \mathbf{a} \tan\left(\frac{\phi}{4}\right), \tag{2.48}$$

where $\mathbf{a} \in \mathbb{S}^2$ and ϕ are axis/angle parameters. As $\phi \to \pm 2\pi$ the Euclidan norm of $\mathbf{s}$ goes to infinity, that is $\|\mathbf{s}\|_2 \to \infty$. However, any attitude can be represented, and therefore MRPs are a global attitude representation.

It can be shown that MRPs are not a unique attitude representation.

2.4.8.5 Principal Direction Cosine Matrices

Consider two reference frames, $\mathcal{F}_p$ and $\mathcal{F}_q$. Often we're interested in the direction cosine matrix $\mathbf{C}_{qp}$ associated with a pure rotation about one of the principal axis, that is, one of $\underrightarrow{p}^{i} \in \mathbb{P}$, $i = 1, 2, 3$.

Consider two reference frames, $\mathcal{F}_p$ and $\mathcal{F}_q$, and a pure rotation about the 3 axis. What's the direction cosine matrix, $\mathbf{C}_{qp}$, associated with this rotation? Well, recall from Definition 2.6 in Section 2.1.5 that the dot product between two physical vectors $\underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}$ is

$$\underrightarrow{u} \cdot \underrightarrow{v} = ||\underrightarrow{u}||_2 \, ||\underrightarrow{v}||_2 \cos\theta$$

where $\theta \in [0, 180°]$ is the angle between $\underrightarrow{u}$ and $\underrightarrow{v}$. Also, recall from Equation (2.17) in Section 2.4.1 (or Definition 2.14 in Section 2.4.4.1) that

$$\mathbf{C}_{qp} = \underrightarrow{\mathcal{F}}_q \cdot \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} = \begin{bmatrix} \underrightarrow{q}^1 \cdot \underrightarrow{p}^1 & \underrightarrow{q}^1 \cdot \underrightarrow{p}^2 & \underrightarrow{q}^1 \cdot \underrightarrow{p}^3 \\ \underrightarrow{q}^2 \cdot \underrightarrow{p}^1 & \underrightarrow{q}^2 \cdot \underrightarrow{p}^2 & \underrightarrow{q}^2 \cdot \underrightarrow{p}^3 \\ \underrightarrow{q}^3 \cdot \underrightarrow{p}^1 & \underrightarrow{q}^3 \cdot \underrightarrow{p}^2 & \underrightarrow{q}^3 \cdot \underrightarrow{p}^3 \end{bmatrix},$$

where $||\underrightarrow{p}^i||_2 = ||\underrightarrow{q}^i||_2 = 1$, $i = 1, 2, 3$. Note that

$$\begin{array}{lll} \underrightarrow{q}^1 \cdot \underrightarrow{p}^1 = \cos\theta\,, & \underrightarrow{q}^1 \cdot \underrightarrow{p}^2 = \cos\left(\frac{\pi}{2} - \theta\right) = \sin\theta\,, & \underrightarrow{q}^1 \cdot \underrightarrow{p}^3 = 0\,, \\ \underrightarrow{q}^2 \cdot \underrightarrow{p}^1 = \cos\left(\frac{\pi}{2} + \theta\right) = -\sin\theta\,, & \underrightarrow{q}^2 \cdot \underrightarrow{p}^2 = \cos\theta\,, & \underrightarrow{q}^2 \cdot \underrightarrow{p}^3 = 0\,, \\ \underrightarrow{q}^3 \cdot \underrightarrow{p}^1 = 0\,, & \underrightarrow{q}^3 \cdot \underrightarrow{p}^2 = 0\,, & \underrightarrow{q}^3 \cdot \underrightarrow{p}^3 = 1\,, \end{array}$$

and hence

$$\mathbf{C}_{qp} = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $c_\theta = \cos\theta$ and $s_\theta = \sin\theta$. This rotation about the 3 axis is denoted

$$\mathbf{C}_3(\theta) = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In a similar fashion, $\mathbf{C}_1(\phi)$ and $\mathbf{C}_2(\psi)$, the direction cosine matrices associated with rotations about the 1 and 2 axis, can be found. To this end,

$$\mathbf{C}_1(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & s_\phi \\ 0 & -s_\phi & c_\phi \end{bmatrix}, \quad \mathbf{C}_2(\psi) = \begin{bmatrix} c_\psi & 0 & -s_\psi \\ 0 & 1 & 0 \\ s_\psi & 0 & c_\psi \end{bmatrix}.$$

The matrices $\mathbf{C}_1(\phi)$, $\mathbf{C}_2(\psi)$, $\mathbf{C}_3(\theta)$ are the *principal direction cosine matrices* [2], [1, pp. 15], [18, pp. 14-15].

There's an alternative way to derive the principal direction cosine matrices. Recall the axis/angle representation of $\mathbf{C}_{qp}$ given in Equation (2.42),

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

Consider the case where $\mathbf{a} = [0 \;\; 0 \;\; 1]^{\mathsf{T}}$:

$$\mathbf{C}_{qp} = \begin{bmatrix} \cos\phi & 0 & 0 \\ 0 & \cos\phi & 0 \\ 0 & 0 & \cos\phi \end{bmatrix} + (1-\cos\phi) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\sin\phi & 0 \\ \sin\phi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This is a principal 3 rotation. In a similar way, by setting $\mathbf{a} = [1 \;\; 0 \;\; 0]^{\mathsf{T}}$ and $\mathbf{a} = [0 \;\; 1 \;\; 0]^{\mathsf{T}}$ principal 1 and 2 rotations can be derived.

2.4.8.6 Parameterizing the Direction Cosine Matrix Using an Euler-Angle Sequence

Attitude can be represented as the product of three principal direction cosine matrices. For instance, consider the following $3-2-1$ *Euler-angle sequence*,

$$\mathbf{C}_{da} = \mathbf{C}_{dc}\mathbf{C}_{cb}\mathbf{C}_{ba} = \mathbf{C}_1(\gamma)\mathbf{C}_2(\beta)\mathbf{C}_3(\alpha)$$

where $\boldsymbol{\theta} = [\alpha \;\; \beta \;\; \gamma]^{\mathsf{T}}$ are called the *Euler angles*. Note, in order to pronounce the name "Euler" correctly, just say "The Edmonton Oilers", and replace "Oilers" with "Euler" leading to "The Edmonton Eulers". The Edmonton Eulers score a lot of sharp angle goals!

Euler-angle sequences can be used to parameterize attitude. For instance, $3-2-1$ and $3-1-3$ Euler-angle sequences are popular attitude representations. In total there are 12 different Euler-angle sequences that can be used to represent attitude. Unfortunately, all Euler-angle sequences suffer from a condition where at a particular attitude two Euler angles in the sequence *coalesce* into one degree of freedom. Consider a $3-1-3$ Euler-angle sequence, $\mathbf{C}_{da} = \mathbf{C}_3(\psi)\mathbf{C}_1(\kappa)\mathbf{C}_3(\upsilon)$ where $\kappa = \pm n2\pi$, $n \in \mathbb{Z}$. In this case, $\mathbf{C}_1(\pm n2\pi) = \mathbf{1}$ and

$$\mathbf{C}_{da} = \mathbf{C}_3(\psi)\mathbf{C}_1(n2\pi)\mathbf{C}_3(\upsilon) = \mathbf{C}_3(\psi)\mathbf{C}_3(\upsilon) = \mathbf{C}_3(\psi+\upsilon).$$

Here ψ and υ are associated with the same degree of freedom.

Before continuing it is worth pausing to ask whether or not, in general, direction cosine matrix multiplication is commutative,

$$\mathbf{C}_2(\beta)\mathbf{C}_3(\alpha) \stackrel{?}{=} \mathbf{C}_3(\alpha)\mathbf{C}_2(\beta)?$$

The answer is definitely "no" [2], [18, pp. 24]. To see why, let $\mathbf{C}_{ba} = \mathbf{C}_3(\alpha)$ and $\mathbf{C}_{cb} = \mathbf{C}_2(\beta)$, and then note that

$$\mathbf{C}_{cb}\mathbf{C}_{ba} \neq \mathbf{C}_{ba}\mathbf{C}_{cb}.$$

Notice that the subscripts do not line up.

Is An Euler-angle sequence a Global and Unique Attitude Representation?

Any Euler-angle sequence reduces to the product of sine and cosine functions. Given any Euler angles, one can compute the corresponding direction cosine matrix, and as such Euler angles are able to represent attitude globally. Given any direction cosine matrix there exists a set of Euler angles that can generate said direction cosine matrix. However, as discussed next, the Euler-angle sequence is not unique.

Recall from Section 2.4.8.6 that all Euler-angle sequences suffer from a condition where at a particular attitude two Euler angles in the sequence coalesce into one degree of freedom. In the case of a $3-1-3$ Euler-angle sequence, $\mathbf{C}_{da} = \mathbf{C}_3(\psi)\mathbf{C}_1(\kappa)\mathbf{C}_3(\upsilon)$, this condition occurs when $\kappa = \pm n2\pi$, $n \in \mathbb{Z}$. There are an infinite number of combination ψ and υ that represent the same direction cosine matrix $\mathbf{C}_{da}$ when $\kappa = \pm n2\pi$, $n \in \mathbb{Z}$. Thus, an Euler-angle sequence is not a unique attitude representation.

2.4.8.7 Some Examples

Example 2.9. Compute $\mathbf{C}_{da} = \mathbf{C}_1(\gamma)\mathbf{C}_2(\beta)\mathbf{C}_1(\alpha)$ explicitly.

Solution. Via matrix multiplication it follows that

$$
\begin{aligned}
\mathbf{C}_{da} &= \mathbf{C}_1(\gamma)\mathbf{C}_2(\beta)\mathbf{C}_3(\alpha) \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\gamma & s_\gamma \\ 0 & -s_\gamma & c_\gamma \end{bmatrix} \begin{bmatrix} c_\beta & 0 & -s_\beta \\ 0 & 1 & 0 \\ s_\beta & 0 & c_\beta \end{bmatrix} \begin{bmatrix} c_\alpha & s_\alpha & 0 \\ -s_\alpha & c_\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\gamma & s_\gamma \\ 0 & -s_\gamma & c_\gamma \end{bmatrix} \begin{bmatrix} c_\beta c_\alpha & c_\beta s_\alpha & -s_\beta \\ -s_\alpha & c_\alpha & 0 \\ s_\beta c_\alpha & s_\beta s_\alpha & c_\beta \end{bmatrix} \\
&= \begin{bmatrix} c_\beta c_\alpha & c_\beta s_\alpha & -s_\beta \\ -c_\gamma s_\alpha + s_\gamma s_\beta c_\alpha & c_\gamma c_\alpha + s_\gamma s_\beta s_\alpha & s_\gamma c_\beta \\ s_\gamma s_\alpha + c_\gamma s_\beta c_\alpha & -s_\gamma c_\alpha + c_\gamma s_\beta s_\alpha & c_\gamma c_\beta \end{bmatrix}.
\end{aligned}
$$

□

Example 2.10. Verify that $\mathbf{C}_1(\phi)$ satisfies $\mathbf{C}_1^{\mathsf{T}}(\phi)\mathbf{C}_1(\phi) = \mathbf{1}$ and $\det \mathbf{C}_1(\phi) = +1$.

Solution. Observe that

$$
\begin{aligned}
\mathbf{C}_1^{\mathsf{T}}(\phi)\mathbf{C}_1(\phi) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi & c_\phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & s_\phi \\ 0 & -s_\phi & c_\phi \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi^2 + s_\phi^2 & c_\phi s_\phi - s_\phi c_\phi \\ 0 & s_\phi c_\phi - c_\phi s_\phi & c_\phi^2 + s_\phi^2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},
\end{aligned}
$$

where the identity $c_\phi^2 + s_\phi^2 = 1$ has been used. Also notice that

$$
\det \mathbf{C}_1(\phi) = \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi & c_\phi \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi & c_\phi \end{bmatrix} \begin{matrix} 1 & 0 \\ 0 & c_\phi \\ 0 & s_\phi \end{matrix} = c_\phi^2 + s_\phi^2 = 1.
$$

Thus, $\mathbf{C}_1(\phi)$ satisfies $\mathbf{C}_1^{\mathsf{T}}(\phi)\mathbf{C}_1(\phi) = \mathbf{1}$ and $\det \mathbf{C}_1(\phi) = +1$, as it should. □

Example 2.11. Given

$$
\underrightarrow{u} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a, \quad \mathbf{u}_a = [0 \;\; 2 \;\; 1]^{\mathsf{T}}, \quad \mathbf{C}_{ba} = \mathbf{C}_3(\pi/2),
$$

what is $\mathbf{u}_b$?

Solution. Recall that

$$
\begin{aligned}
\underrightarrow{u} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b, \\
\underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a &= \underrightarrow{\mathcal{F}}_b \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b, \\
\mathbf{C}_{ba} \mathbf{u}_a &= \mathbf{u}_b.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\mathbf{u}_b &= \mathbf{C}_{ba}\mathbf{u}_a \\
&= \begin{bmatrix} c_{\pi/2} & s_{\pi/2} & 0 \\ -s_{\pi/2} & c_{\pi/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.
\end{aligned}
$$

Readers are encouraged to compare this result to Example 2.7. □

2.4.8.8 Deficiencies of Direction Cosine Matrix Parameterizations

Recall that the direction cosine matrix is a global and unique representation of attitude. All parameterizations of the direction cosine matrix are deficient in some way because attitude represented by a parameterization is either not a global representation or not a unique representation. Moreover, no three-parameter attitude parameterization can be both global and finite [1, pp. 29]. Table 2.1 summarizes the properties of the direction cosine matrix parameterizations discussed in this chapter.

Parameterization	Global	Unique	Advantages	Disadvantages
Direction Cosine Matrix	Yes	Yes	Global and unique.	9 parameters and 6 constraints.
Axis/Angle Parameters	Yes	No	Easy to visualize. 4 parameters.	1 constraint; trig. functions.
Quaternions	Yes	No	No trig. functions. 4 parameters.	1 constraint; not unique.
Rodrigues Parameters	No	Yes	3 parameters.	Goes to infinity for large angles.
MRPs	Yes	No	3 parameters.	Not unique.
Euler-angle sequence	Yes	No	Easy to visualize; 3 parameters.	Not unique.

Table 2.1: Attitude parameterizations and their properties.

Recall from Section 2.4.4.1 that the direction cosine matrix has 9 elements, however, for simulation purposes it is possible to construct the full direction cosine matrix from only 6 elements via

$$\mathbf{a}_b^3 = \mathbf{a}_b^{1^\times}\mathbf{a}_b^2,$$

where

$$\mathbf{C}_{ba} = \begin{bmatrix} \mathbf{a}_b^1 & \mathbf{a}_b^2 & \mathbf{a}_b^3 \end{bmatrix}.$$

This result follows from the fact that $\underrightarrow{a}^3 = \underrightarrow{a}^1 \times \underrightarrow{a}^2$. Therefore, one does not have to use all 9 elements of the direction cosine matrix, rather, one can use 6 and compute the remaining 3 as needed.

2.4.9 Relations Between Physical Vectors and Tensors Resolved in Different Reference Frames

2.4.9.1 Relations Involving Physical Vectors

Consider the physical vectors $\underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}$ and the reference frames $\mathcal{F}_a$ and $\mathcal{F}_c$ where

$$\begin{aligned}
\underrightarrow{u} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{C}_{ab} \mathbf{u}_b = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b, \\
\mathbf{u}_b &= \mathbf{C}_{ba} \mathbf{u}_a, \\
\underrightarrow{v} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{C}_{ab} \mathbf{v}_b = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{v}_b, \\
\mathbf{v}_b &= \mathbf{C}_{ba} \mathbf{v}_a,
\end{aligned}$$

and $\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab}$. The relationship between $\underrightarrow{u}$ and $\underrightarrow{v}$, when each are resolved in different reference frames, will now be explored. Equations (2.18) and (2.21), $\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba}$ and $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{C}_{ab}$, will be needed.

First, consider the addition of $\underrightarrow{u}$ and $\underrightarrow{v}$, that being

$$\begin{aligned}
\underrightarrow{u} + \underrightarrow{v} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{v}_b \\
&= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \left(\mathbf{u}_a + \mathbf{C}_{ab} \mathbf{v}_b\right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{C}_{ba} \mathbf{u}_a + \mathbf{v}_b\right).
\end{aligned}$$

Notice that $\mathbf{u}_a$ is multiplied by $\mathbf{C}_{ba}$ yielding a set of components that is then resolved in $\mathcal{F}_b$, which is then added to $\mathbf{v}_b$, another set of components resolved in $\mathcal{F}_b$. The components of physical vectors can only be added if the components of each vector being added are resolved in the same reference frame.

Next, consider the dot product of $\underrightarrow{u}$ and $\underrightarrow{v}$, that is

$$\begin{aligned}
\underrightarrow{u} \cdot \underrightarrow{v} &= \left(\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a\right) \cdot \left(\underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{v}_b\right) \\
&= \mathbf{u}_a^{\mathsf{T}} \underrightarrow{\mathcal{F}}_a \cdot \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{v}_b \\
&= \mathbf{u}_a^{\mathsf{T}} \mathbf{C}_{ab} \mathbf{v}_b \\
&= \left(\mathbf{C}_{ba} \mathbf{u}_a\right)^{\mathsf{T}} \mathbf{v}_b,
\end{aligned}$$

where $\left(\mathbf{C}_{ba} \mathbf{u}_a\right)^{\mathsf{T}} = \mathbf{u}_a^{\mathsf{T}} \mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{u}_a^{\mathsf{T}} \mathbf{C}_{ab}$ owing to the fact that $\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab}$.

A very useful identity related to the cross product of $\underrightarrow{u}$ and $\underrightarrow{v}$ will now be derived. Consider

$$\begin{aligned}
\underrightarrow{u} \times \underrightarrow{v} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{u}_a^{\times} \mathbf{v}_a \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba} \mathbf{u}_a^{\times} \mathbf{C}_{ab} \mathbf{v}_b \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba} \mathbf{u}_a^{\times} \mathbf{C}_{ba}^{\mathsf{T}} \mathbf{v}_b,
\end{aligned} \tag{2.49}$$

where Equation (2.18), $\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba}$, and $\mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ab}$ have been used. In a similar fashion,

$$\begin{aligned}
\underrightarrow{u} \times \underrightarrow{v} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{u}_b^{\times} \mathbf{v}_b \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{C}_{ba} \mathbf{u}_a\right)^{\times} \mathbf{v}_b.
\end{aligned} \tag{2.50}$$

Equating Equations (2.49) and (2.50) leads to the identity [2], [1, pp. 529], [18, pp. 14]

$$\left(\mathbf{C}_{ba} \mathbf{u}_a\right)^{\times} = \mathbf{C}_{ba} \mathbf{u}_a^{\times} \mathbf{C}_{ba}^{\mathsf{T}}. \tag{2.51}$$

This identity is extremely useful.

2.4.9.2 Relations Involving Tensors

Consider a tensor $\underrightarrow{T} : \mathbb{P} \rightarrow \mathbb{P}$ and the reference frames $\mathcal{F}_a$ and $\mathcal{F}_b$ where

$$\begin{aligned}
\underrightarrow{T} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{T}_a \underrightarrow{\boldsymbol{\mathcal{F}}}_a & (2.52)\\
&= \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{T}_b \underrightarrow{\boldsymbol{\mathcal{F}}}_b, & (2.53)
\end{aligned}$$

$\mathbf{T}_a \in \mathbb{R}^{3\times3}$ is a matrix of scalars representing the tensor $\underrightarrow{T}$ expressed in reference frame $\mathcal{F}_a$, and $\mathbf{T}_b \in \mathbb{R}^{3\times3}$ is a matrix of scalars representing the tensor $\underrightarrow{T}$ expressed in reference frame $\mathcal{F}_b$. Using $\underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{C}_{ba}$ from Equations (2.18), Equation (2.52) can be written as

$$\underrightarrow{T} = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{C}_{ba} \mathbf{T}_a \mathbf{C}_{ba}^{\mathsf{T}} \underrightarrow{\boldsymbol{\mathcal{F}}}_b = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{C}_{ba} \mathbf{T}_a \mathbf{C}_{ab} \underrightarrow{\boldsymbol{\mathcal{F}}}_b.$$

Combining this result with Equation (2.53) gives [2], [1, pp. 530]

$$\mathbf{T}_b = \mathbf{C}_{ba} \mathbf{T}_a \mathbf{C}_{ba}^{\mathsf{T}} = \mathbf{C}_{ba} \mathbf{T}_a \mathbf{C}_{ab}.$$

In a similar manner it's straightforward to show that

$$\mathbf{T}_a = \mathbf{C}_{ab} \mathbf{T}_b \mathbf{C}_{ab}^{\mathsf{T}} = \mathbf{C}_{ab} \mathbf{T}_b \mathbf{C}_{ba}.$$

Notice that the notation is clear, concise, and consistent; all the subscripts "line up".

Chapter 3

Kinematics

Kinematics is the study of the geometry of motion, without concern or regard for the cause of motion [7]. Kinematics is the subject of this chapter.

3.1 Newtonian Time

In Newtonian mechanics it is assumed that *time* is invariant with respect to the frame of reference from which is observed. Said another way, observations of time are the same in all reference frames. For instance, given reference frames $\mathcal{F}_a$ and $\mathcal{F}_b$, then the time observed in $\mathcal{F}_a$ equals the time observed in $\mathcal{F}_b$ [4, pp. 29-30]. This reference-frame-invariant time is refereed to as Newtonian Time [2].

Well, at this point readers are probably wondering "when is time not reference-frame-invariant"? The theory of special relativity asserts that time is not reference-frame-invariant, but rather the speed of light is, where $c = 300 \times 10^6$ (m/s) is the speed of light. This assumption leads to what is known as the Lorentz transformation, which forms the basis of the kinematics of special relativity. What's important in the context of Newtonian mechanics is that provided objects are moving relative to each other much slower than the speed of light, clocks affixed to all object observe time to "flow" at the same rate [2], [31, pp. 277-281]. For example, consider a spacecraft orbiting the Earth in a polar orbit (so that the spacecraft passes right over the poles of the Earth) with an altitude of 644 (km) at a velocity of $27\ 080$ (km/h), and two clocks, one on-board the spacecraft, the other at one of the poles. Then, each orbit the clock on-board the spacecraft would slow by $0.000\ 001\ 85$ (s) compared to the clock at the pole [32, pp. 120]. As such, in many practical situations the theory of special relativity is not needed.

3.2 Derivative With Respect to Time of Scalars, Physical Vectors, and Vectrices

3.2.1 Derivative With Respect to Time of a Scalar Function of Time

Definition 3.1. Consider a scalar function of time, such as $u : \mathbb{R}^+ \to \mathbb{R}$. The derivative with respect to time of u is

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = \dot{u}(t) = \lim_{\Delta t \to 0} \frac{u(t + \Delta t) - u(t)}{\Delta t} \tag{3.1}$$

Unless clarity is required, often the argument "t" will be neglected.

Notice that the definition of the scalar derivative does not involve any sort of reference frame. As such, the derivative with respect to time of a scalar, such as u, is independent of a particular reference frame [2] [4, pp. 28].

3.2.2 Derivative With Respect to Time of a Physical Vector Function of Time

Although the derivative with respect to time of a scalar is independent of a particular reference frame, the derivative with respect to time of a physical vector is not. To illustrate this point, consider Figure 3.1 depicting a rotating table with a lady bug, named Ms. Bug, walking away from the center of the rotating table out to the perimeter of the rotating table. Consider two reference frames, $\mathcal{F}_a$ and $\mathcal{F}_b$. Frame $\mathcal{F}_a$ is fixed to the "ground" and does not rotate with the table, while $\mathcal{F}_b$ is fixed to the table and rotates with the table. With respect to an observer situated in $\mathcal{F}_b$, the observer sees Ms. Bug simply strolling down the $\underrightarrow{b}^1$ axis of $\mathcal{F}_b$. However, with respect to an observer situated in $\mathcal{F}_a$, the observer not only sees Ms. Bug move down $\underrightarrow{b}^1$, but also sees Ms. Bug rotate as the table rotates. Clearly, the motion of Ms. Bug with respect to an observer situated in $\mathcal{F}_a$ is different than the motion of Ms. Bug with respect to an observer situated in $\mathcal{F}_b$. As such, the derivative of a physical vector is meaningless without explicitly stating which frame the derivative is with respect to [2] [4, pp. 32].

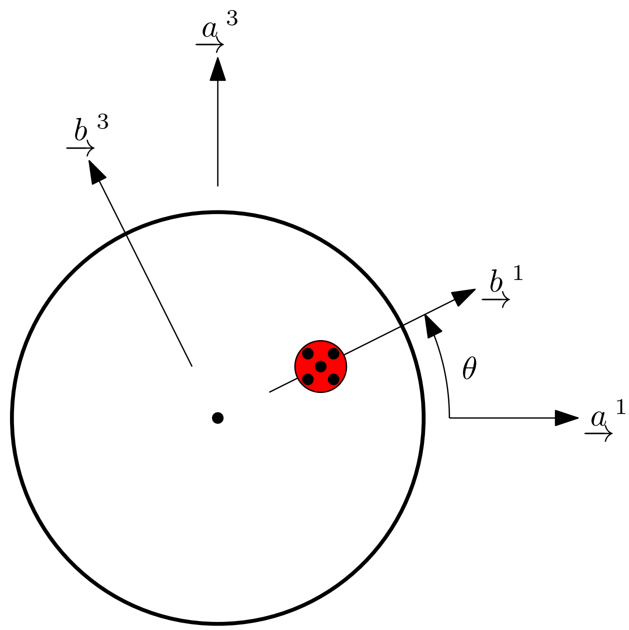


Figure 3.1: Spinning table with Ms. Bug.

Definition 3.2. Consider a physical vector function of time, such as $\underrightarrow{u} : \mathbb{R}^+ \to \mathbb{P}$, and a reference frame, such as $\mathcal{F}_a$. The time derivative *with respect to* $\mathcal{F}_a$ of $\underrightarrow{u}(\cdot)$ is

$$\left.\frac{\mathrm{d}\underrightarrow{u}(t)}{\mathrm{d}t}\right|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a}(t) = \overset{\mathcal{F}_a}{\lim_{\Delta t \to 0}} \frac{\underrightarrow{u}(t+\Delta t) - \underrightarrow{u}(t)}{\Delta t} \tag{3.2}$$

where

$$\overset{\mathcal{F}_a}{\lim_{\Delta t \to 0}}$$

means the limit is taken with respect to $\mathcal{F}_a$ [4, pp. 32-33].

For simplicity, the temporal argument "t" will be dropped. The limit in Equation (3.2) is computed given the time-rate-of-change of $\underrightarrow{u}$ with respect to reference frame $\mathcal{F}_a$. The short-hand notation $\underrightarrow{u}^{\bullet a}$ is used to simplify writing the derivative of a physical vector. A similar notation is used in [3]. To write the time derivative of $\underset{\rightarrow}{u}$ with respect to $\mathcal{F}_b$, $\mathcal{F}_c$, $\mathcal{F}_d$, etc., is

$$\underset{\rightarrow}{u}^{\bullet b}, \quad \underset{\rightarrow}{u}^{\bullet c}, \quad \underset{\rightarrow}{u}^{\bullet d}.$$

3.2.3 Derivative With Respect to Time of a Vectrix

Consider $\mathcal{F}_a$ and the corresponding vectrix $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix}$. With respect to $\mathcal{F}_a$, the physical basis vectors $\underset{\rightarrow}{a}^1$, $\underset{\rightarrow}{a}^2$, and $\underset{\rightarrow}{a}^3$ do not change with time. As such,

$$\underset{\rightarrow}{a}^{1^{\bullet a}} = \underset{\rightarrow}{0}, \quad \underset{\rightarrow}{a}^{2^{\bullet a}} = \underset{\rightarrow}{0}, \quad \underset{\rightarrow}{a}^{3^{\bullet a}} = \underset{\rightarrow}{0}.$$

As such, the time derivative of the vectrix $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a$ yields the null vectrix,

$$\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\bullet a} = \underset{\rightarrow}{\boldsymbol{\mathcal{O}}} = \begin{bmatrix} \underset{\rightarrow}{0} \\ \underset{\rightarrow}{0} \\ \underset{\rightarrow}{0} \end{bmatrix}.$$

Similarly, consider $\mathcal{F}_b$ and the corresponding vectrix $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} = \begin{bmatrix} \underset{\rightarrow}{b}^1 & \underset{\rightarrow}{b}^2 & \underset{\rightarrow}{b}^3 \end{bmatrix}$. With respect to $\mathcal{F}_b$ the physical basis vectors $\underset{\rightarrow}{b}^1$, $\underset{\rightarrow}{b}^2$, and $\underset{\rightarrow}{b}^3$ do not change with time. It follows that

$$\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\bullet b} = \underset{\rightarrow}{\boldsymbol{\mathcal{O}}} = \begin{bmatrix} \underset{\rightarrow}{0} \\ \underset{\rightarrow}{0} \\ \underset{\rightarrow}{0} \end{bmatrix}.$$

Now, what about $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\bullet b}$ and $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\bullet a}$? This question will be answered momentarily, but for now, readers are encouraged to ponder what the result might be. Will the derivatives be the null vectrix once again, or something else?

3.2.4 Time Derivatives of Physical Vectors in Terms of Components

Consider a physical vector that changes with time, $\underset{\rightarrow}{u} : \mathbb{R}^+ \to \mathbb{P}$, and $\mathcal{F}_a$ with corresponding vectrix $\underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a$, that being

$$\underset{\rightarrow}{u} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{u}_a = \begin{bmatrix} \underset{\rightarrow}{a}^1 & \underset{\rightarrow}{a}^2 & \underset{\rightarrow}{a}^3 \end{bmatrix} \begin{bmatrix} u_{a1} \\ u_{a2} \\ u_{a3} \end{bmatrix} = (u_{a1})\underset{\rightarrow}{a}^1 + (u_{a2})\underset{\rightarrow}{a}^2 + (u_{a3})\underset{\rightarrow}{a}^3 = \sum_{i=1}^{3} u_{ai}\underset{\rightarrow}{a}^i,$$

where u_{ai}, $i = 1, 2, 3$ are the components of $\underset{\rightarrow}{u}$ resolved in $\mathcal{F}_a$. First, recall that the physical basis vectors composing $\mathcal{F}_a$ are constant with respect to $\mathcal{F}_a$. Next, using Equation (3.2) the time derivative of $\underset{\rightarrow}{u}$ with respect to $\mathcal{F}_a$ is [4, pp. 32-33]

$$\begin{aligned}
\underset{\rightarrow}{u}^{\bullet a} &= \overset{\mathcal{F}_a}{\lim_{\Delta t \to 0}} \sum_{i=1}^{3} \frac{u_{ai}(t+\Delta t) - u_{ai}(t)}{\Delta t} \underset{\rightarrow}{a}^{i} \\
&= \sum_{i=1}^{3} \overset{\mathcal{F}_a}{\lim_{\Delta t \to 0}} \frac{u_{ai}(t+\Delta t) - u_{ai}(t)}{\Delta t} \underset{\rightarrow}{a}^{i} \\
&= \sum_{i=1}^{3} \left(\lim_{\Delta t \to 0} \frac{u_{ai}(t+\Delta t) - u_{ai}(t)}{\Delta t} \right) \underset{\rightarrow}{a}^{i},
\end{aligned}$$

where the fact that the summation and the limit can be interchanged has been exploited. Using Equation (3.1), notice that

$$\dot{u}_{ai}(t) = \lim_{\Delta t \to 0} \frac{u_{ai}(t+\Delta t) - u_{ai}(t)}{\Delta t}, \quad i = 1, 2, 3, \tag{3.3}$$

and as such

$$\underset{\rightarrow}{u}^{\bullet a} = \sum_{i=1}^{3} \dot{u}_{ai} \underset{\rightarrow}{a}^{i} = (\dot{u}_{a1}) \underset{\rightarrow}{a}^{1} + (\dot{u}_{a2}) \underset{\rightarrow}{a}^{2} + (\dot{u}_{a3}) \underset{\rightarrow}{a}^{3} = \begin{bmatrix} \underset{\rightarrow}{a}^{1} & \underset{\rightarrow}{a}^{2} & \underset{\rightarrow}{a}^{3} \end{bmatrix} \underbrace{\begin{bmatrix} \dot{u}_{a1} \\ \dot{u}_{a2} \\ \dot{u}_{a3} \end{bmatrix}}_{\dot{\mathbf{u}}_a} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \dot{\mathbf{u}}_a.$$

Note that the limiting procedure in Equation (3.3) is the same limiting procedure as in Equation (3.1); it's a limiting procedure on three scalars, which is why the "$\mathcal{F}_a$" on top of the limit symbol has been removed.

Now, as discussed at length in Section 2.3, the physical vector $\underset{\rightarrow}{u} : \mathbb{R}^+ \rightarrow \mathbb{P}$ can be resolved in any other frame. To this end, consider $\underset{\rightarrow}{u}$ resolved in $\mathcal{F}_b$,

$$\underset{\rightarrow}{u} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \mathbf{u}_b = \begin{bmatrix} \underset{\rightarrow}{b}^{1} & \underset{\rightarrow}{b}^{2} & \underset{\rightarrow}{b}^{3} \end{bmatrix} \begin{bmatrix} u_{b1} \\ u_{b2} \\ u_{b3} \end{bmatrix} = (u_{b1}) \underset{\rightarrow}{b}^{1} + (u_{b2}) \underset{\rightarrow}{b}^{2} + (u_{b3}) \underset{\rightarrow}{b}^{3} = \sum_{i=1}^{3} u_{bi} \underset{\rightarrow}{b}^{i},$$

where $\underset{\rightarrow}{b}^{i}$ and u_{bi}, $i = 1, 2, 3$, are the physical basis vectors that compose $\mathcal{F}_b$ and the components of $\underset{\rightarrow}{u}$ resolved in $\mathcal{F}_b$, respectively. Recall that, in general, $u_{bi} \neq u_{ai}$, $i = 1, 2, 3$.

Again, realize that the physical basis vectors composing $\mathcal{F}_b$ are constant with respect to $\mathcal{F}_b$. Using Equation (3.2) the time derivative of $\underset{\rightarrow}{u}$ with respect to $\mathcal{F}_b$ is then

$$\begin{aligned}
\underset{\rightarrow}{u}^{\bullet b} &= \overset{\mathcal{F}_b}{\lim_{\Delta t \to 0}} \sum_{i=1}^{3} \frac{u_{bi}(t+\Delta t) - u_{bi}(t)}{\Delta t} \underset{\rightarrow}{b}^{i} \\
&= \sum_{i=1}^{3} \overset{\mathcal{F}_b}{\lim_{\Delta t \to 0}} \frac{u_{bi}(t+\Delta t) - u_{bi}(t)}{\Delta t} \underset{\rightarrow}{b}^{i} \\
&= \sum_{i=1}^{3} \left(\lim_{\Delta t \to 0} \frac{u_{bi}(t+\Delta t) - u_{bi}(t)}{\Delta t} \right) \underset{\rightarrow}{b}^{i},
\end{aligned}$$

where, again, the fact that the summation and the limit can be interchanged has been exploited. Using

Equation (3.1), notice that

$$\dot{u}_{bi}(t) = \lim_{\Delta t \to 0} \frac{u_{bi}(t + \Delta t) - u_{bi}(t)}{\Delta t}, \quad i = 1, 2, 3, \tag{3.4}$$

and thus

$$\underrightarrow{u}^{\bullet b} = \sum_{i=1}^{3} \dot{u}_{bi} \underrightarrow{b}^{i} = (\dot{u}_{b1}) \underrightarrow{b}^{1} + (\dot{u}_{b2}) \underrightarrow{b}^{2} + (\dot{u}_{b3}) \underrightarrow{b}^{3} = \begin{bmatrix} \underrightarrow{b}^{1} & \underrightarrow{b}^{2} & \underrightarrow{b}^{3} \end{bmatrix} \underbrace{\begin{bmatrix} \dot{u}_{b1} \\ \dot{u}_{b2} \\ \dot{u}_{b3} \end{bmatrix}}_{\dot{\mathbf{u}}_b} = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \dot{\mathbf{u}}_b.$$

Note that the limiting procedure in Equation (3.4) is the same limiting procedure as in Equation (3.1); it's a limiting procedure on three scalars, which is why the "$\mathcal{F}_b$" on top of the limit symbol has been removed.

3.2.5 A Discussion on "Dots"

At this point it is worth pausing to make sure readers clearly understand the difference between all the "dots" used to define different kinds of time differentiation.

Regarding $\underrightarrow{u} : \mathbb{R}^+ \to \mathbb{P}$,

$$\begin{aligned}
\left.\frac{\mathrm{d}\,\underrightarrow{u}}{\mathrm{d}t}\right|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a} &= \text{time derivative of } \underrightarrow{u} \text{ with respect to } \mathcal{F}_a, \\
\dot{\mathbf{u}}_a = \begin{bmatrix} \dot{u}_{a1} \\ \dot{u}_{a2} \\ \dot{u}_{a3} \end{bmatrix} &= \text{time derivative of the components of } \underrightarrow{u} \text{ resolved in } \mathcal{F}_a, \\
\left.\frac{\mathrm{d}\,\underrightarrow{u}}{\mathrm{d}t}\right|_{\mathcal{F}_b} = \underrightarrow{u}^{\bullet b} &= \text{time derivative of } \underrightarrow{u} \text{ with respect to } \mathcal{F}_b, \\
\dot{\mathbf{u}}_b = \begin{bmatrix} \dot{u}_{b1} \\ \dot{u}_{b2} \\ \dot{u}_{b3} \end{bmatrix} &= \text{time derivative of the components of } \underrightarrow{u} \text{ resolved in } \mathcal{F}_b, \\
\left.\frac{\mathrm{d}\,\underrightarrow{u}}{\mathrm{d}t}\right|_{\mathcal{F}_\chi} = \underrightarrow{u}^{\bullet \chi} &= \text{time derivative of } \underrightarrow{u} \text{ with respect to } \mathcal{F}_\chi, \\
\dot{\mathbf{u}}_\chi = \begin{bmatrix} \dot{u}_{\chi 1} \\ \dot{u}_{\chi 2} \\ \dot{u}_{\chi 3} \end{bmatrix} &= \text{time derivative of the components of } \underrightarrow{u} \text{ resolved in } \mathcal{F}_\chi,
\end{aligned}$$

where $\chi = c, d, e,$ etc. It's important to understand that, in general, $\underrightarrow{u}^{\bullet a} \neq \underrightarrow{u}^{\bullet b} \neq \underrightarrow{u}^{\bullet c}$, and so on. Similarly, in general $\dot{\mathbf{u}}_a \neq \dot{\mathbf{u}}_a \neq \dot{\mathbf{u}}_c$, and so on.

Now, why is the same kind of "dot" used for the derivative of the components of the vector? That is, why write

$$\dot{\mathbf{u}}_a \quad \text{and} \quad \dot{\mathbf{u}}_b,$$

and not

$$\mathbf{u}_a^{\bullet a} \quad \text{and} \quad \mathbf{u}_b^{\bullet b}?$$

The reason traces back to Section 3.1. Both $\mathbf{u}_a$ and $\mathbf{u}_b$ are column matrices who's entries are scalars. As discussed in Section 3.1, time evolves uniformly in all reference frames. As such, the time derivative of any scalar in any frame is the same "kind" of derivative. If time were not uniform, and time evolved differently in different reference frames, then some additional notation would be needed [5, pp. 92].

3.2.6 Vector Derivative Properties

Let $\underrightarrow{u}, \underrightarrow{v}, \underrightarrow{\omega} \in \mathbb{P}$ be physical vectors that are functions of time, and $f \in \mathbb{R}$ be a scalar function of time. The following hold [4, pp. 34-35], [5, pp. 58, 88-93].

1. (Addition rule.)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\underrightarrow{u}(t) + \underrightarrow{v}(t)\right)\bigg|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a}(t) + \underrightarrow{v}^{\bullet a}(t), \quad \forall \underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}.$$

2. (Product rule with a scalar.)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f(t)\,\underrightarrow{u}(t)\right)\bigg|_{\mathcal{F}_a} = \dot{f}(t)\,\underrightarrow{u}(t) + f(t)\,\underrightarrow{u}^{\bullet a}(t), \quad \forall f \in \mathbb{R}, \forall \underrightarrow{u} \in \mathbb{P}.$$

3. (Product rule with the dot product.)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\underrightarrow{u}(t) \cdot \underrightarrow{v}(t)\right)\bigg|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a}(t) \cdot \underrightarrow{v}(t) + \underrightarrow{u}(t) \cdot \underrightarrow{v}^{\bullet a}(t), \quad \forall \underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}.$$

4. (Product rule with the cross product.)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\underrightarrow{u}(t) \times \underrightarrow{v}(t)\right)\bigg|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a}(t) \times \underrightarrow{v}(t) + \underrightarrow{u}(t) \times \underrightarrow{v}^{\bullet a}(t), \quad \forall \underrightarrow{u}, \underrightarrow{v} \in \mathbb{P}.$$

5. (Chain rule for a physical vector.)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\underrightarrow{u}(f(t))\right]\bigg|_{\mathcal{F}_a} = \underrightarrow{u}^{\bullet a}(f(t))\dot{f}(t), \quad \forall f \in \mathbb{R}, \forall \underrightarrow{u} \in \mathbb{P}.$$

6. (Chain rule for a scalar function of a physical vector.)

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\underrightarrow{u}(t))\bigg|_{\mathcal{F}_a} = \underrightarrow{\nabla} f(t) \cdot \underrightarrow{u}^{\bullet a}(t), \quad \forall f \in \mathbb{R}, \forall \underrightarrow{u} \in \mathbb{P},$$

where $\underrightarrow{u} = \underrightarrow{a}^1(u_{a1}) + \underrightarrow{a}^2(u_{a2}) + \underrightarrow{a}^3(u_{a3})$, and

$$\begin{aligned}
\underrightarrow{\nabla} f(t) &= \underrightarrow{a}^1 \frac{\partial f(t)}{\partial u_{a1}} + \underrightarrow{a}^2 \frac{\partial f(t)}{\partial u_{a2}} + \underrightarrow{a}^3 \frac{\partial f(t)}{\partial u_{a3}} \\
&= \begin{bmatrix} \frac{\partial f(t)}{\partial u_{a1}} & \frac{\partial f(t)}{\partial u_{a2}} & \frac{\partial f(t)}{\partial u_{a3}} \end{bmatrix} \underrightarrow{\mathcal{F}}_a \\
&= \frac{\partial f(t)}{\partial \mathbf{u}_a} \underrightarrow{\mathcal{F}}_a \\
&= \boldsymbol{\nabla}_a f(t) \underrightarrow{\mathcal{F}}_a
\end{aligned}$$

is the *gradient*. Notice that $\boldsymbol{\nabla}_a f(t)$ is a row matrix so that it's consistent with the way a *Jacobian* is defined.

3.3 Angular Velocity and Transport Theorem

Consider a physical vector $\underrightarrow{r}$ and two reference frames, $\mathcal{F}_a$ and $\mathcal{F}_b$. Consider $\underrightarrow{r}$ resolved in each frame,

$$\begin{aligned}\underrightarrow{r} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{r}_a = r_{a1}\underrightarrow{a}^1 + r_{a2}\underrightarrow{a}^2 + r_{a3}\underrightarrow{a}^3 \\ &= \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}}\mathbf{r}_b = r_{b1}\underrightarrow{b}^1 + r_{b2}\underrightarrow{b}^2 + r_{b3}\underrightarrow{b}^3.\end{aligned}$$

The time-derivative of $\underrightarrow{r}$ with respect to (w.r.t.) $\mathcal{F}_a$ is

$$\begin{aligned}\underrightarrow{r}^{\bullet a} &= \dot{r}_{b1}\underrightarrow{b}^1 + \dot{r}_{b2}\underrightarrow{b}^2 + \dot{r}_{b3}\underrightarrow{b}^3 + r_{b1}\underrightarrow{b}^{1\bullet a} + r_{b2}\underrightarrow{b}^{2\bullet a} + r_{b3}\underrightarrow{b}^{3\bullet a} \\ &= \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}}\dot{\mathbf{r}}_b + \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\bullet a}\mathbf{r}_b \\ &= \underrightarrow{r}^{\bullet b} + \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\bullet a}\mathbf{r}_b. \end{aligned} \tag{3.5}$$

Wait ... what on Earth is $\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\bullet a}$?!

3.3.1 A Geometric Approach to the Transport Theorem

3.3.1.1 Angular Velocity

In order to find $\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\bullet a}$, how each $\underrightarrow{b}^{i\bullet a}$, $i = 1, 2, 3$ moves with respect to time will be considered. Consider Figure 3.2 depicting the $\underrightarrow{b}^i$ physical vector moving $\mathrm{d}\underrightarrow{b}^i$ in $\mathrm{d}t$. Specifically,

$$\underrightarrow{b}^i(t + \mathrm{d}t) = \underrightarrow{b}^i(t) + \mathrm{d}\underrightarrow{b}^i(t).$$

The angle $\mathrm{d}\theta$ is

$$\mathrm{d}\theta = \left\|\underrightarrow{\omega}^{ba}\right\|_2 \mathrm{d}t, \tag{3.6}$$

where the physical vector $\underrightarrow{\omega}^{ba}$ is the *angular velocity* of $\mathcal{F}_b$ relative to $\mathcal{F}_a$. The physical vector $\mathrm{d}\underrightarrow{b}^i$ is

$$\begin{aligned}\mathrm{d}\underrightarrow{b}^i &= \mathrm{d}\theta \sin\psi \left\|\underrightarrow{b}^i\right\|_2 \underrightarrow{n} \\ &= \sin\psi \left\|\underrightarrow{\omega}^{ba}\right\|_2 \left\|\underrightarrow{b}^i\right\|_2 \mathrm{d}t\,\underrightarrow{n},\end{aligned} \tag{3.7}$$

where Equation (3.6) has been used, and $\underrightarrow{n}$ is a unit-length physical vector that is perpendicular to both $\underrightarrow{b}^i$ and $\underrightarrow{\omega}^{ba}$ at the instant under consideration. To be clear, $\underrightarrow{n}$ is parallel to, and points in the same direction as, $\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^i$, which in turn means that $\mathrm{d}\underrightarrow{b}^i$ is perpendicular to the plane created by $\underrightarrow{\omega}^{ba}$ and $\underrightarrow{b}^i$. Recalling that $\underrightarrow{u} \times \underrightarrow{v} = \sin\phi \left\|\underrightarrow{u}\right\|_2 \left\|\underrightarrow{v}\right\|_2 \underrightarrow{n}$, Equation (3.7) can be written as

$$\underrightarrow{b}^{i\bullet a} = \frac{\mathrm{d}\underrightarrow{b}^i}{\mathrm{d}t} = \sin\psi \left\|\underrightarrow{\omega}^{ba}\right\|_2 \left\|\underrightarrow{b}^i\right\|_2 \underrightarrow{n} = \underrightarrow{\omega}^{ba} \times \underrightarrow{b}^i.$$

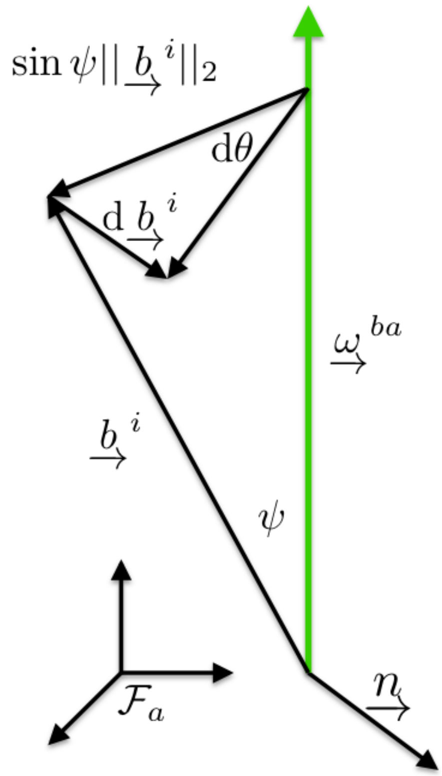


Figure 3.2: The change of $\underrightarrow{b}^{i}$ over an instant $\mathrm{d}t$.

This holds for all $i = 1, 2, 3$, and hence

$$\begin{aligned} \begin{bmatrix} \underrightarrow{b}^{1\,\bullet a} & \underrightarrow{b}^{2\,\bullet a} & \underrightarrow{b}^{3\,\bullet a} \end{bmatrix} &= \begin{bmatrix} \underrightarrow{\omega}^{ba} \times \underrightarrow{b}^{1} & \underrightarrow{\omega}^{ba} \times \underrightarrow{b}^{2} & \underrightarrow{\omega}^{ba} \times \underrightarrow{b}^{3} \end{bmatrix} \\ &= \underrightarrow{\omega}^{ba} \times \begin{bmatrix} \underrightarrow{b}^{1} & \underrightarrow{b}^{2} & \underrightarrow{b}^{3} \end{bmatrix}, \\ \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\,\bullet a} &= \underrightarrow{\omega}^{ba} \times \underrightarrow{\boldsymbol{\mathcal{F}}}_b. \end{aligned}$$

Therefore,

$$\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\,\bullet a} = \underrightarrow{\omega}^{ba} \times \underrightarrow{\boldsymbol{\mathcal{F}}}_b. \tag{3.8}$$

Although a relationship between $\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\,\bullet a}$, $\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}}$, and the angular velocity physical vector $\underrightarrow{\omega}^{ba}$ has been established, it will prove useful to write Equation (3.8) as

$$\underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}\,\bullet a} = \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}} \boldsymbol{\omega}_b^{ba^{\times}}, \tag{3.9}$$

where $\boldsymbol{\omega}_b^{ba} = [\omega_{b1}^{ba} \;\; \omega_{b2}^{ba} \;\; \omega_{b3}^{ba}]^{\mathsf{T}}$ are the components of $\underrightarrow{\omega}^{ba}$ resolved in $\mathcal{F}_b$. To show the equivalence of Equation (3.8) and Equation (3.9), consider

$$\underrightarrow{\omega}^{ba} = \omega_{b1}^{ba} \underrightarrow{b}^{1} + \omega_{b2}^{ba} \underrightarrow{b}^{2} + \omega_{b3}^{ba} \underrightarrow{b}^{3}$$

and $\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^1$, $\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^2$, and $\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^3$:

$$\begin{aligned}
\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^1 &= \omega_{b1}^{ba} \underrightarrow{b}^1 \times \underrightarrow{b}^1 + \omega_{b2}^{ba} \underrightarrow{b}^2 \times \underrightarrow{b}^1 + \omega_{b3}^{ba} \underrightarrow{b}^3 \times \underrightarrow{b}^1 = -\omega_{b2}^{ba} \underrightarrow{b}^3 + \omega_{b3}^{ba} \underrightarrow{b}^2, \\
\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^2 &= \omega_{b1}^{ba} \underrightarrow{b}^1 \times \underrightarrow{b}^2 + \omega_{b2}^{ba} \underrightarrow{b}^2 \times \underrightarrow{b}^2 + \omega_{b3}^{ba} \underrightarrow{b}^3 \times \underrightarrow{b}^2 = \omega_{b1}^{ba} \underrightarrow{b}^3 - \omega_{b3}^{ba} \underrightarrow{b}^1, \\
\underrightarrow{\omega}^{ba} \times \underrightarrow{b}^3 &= \omega_{b1}^{ba} \underrightarrow{b}^1 \times \underrightarrow{b}^3 + \omega_{b2}^{ba} \underrightarrow{b}^2 \times \underrightarrow{b}^3 + \omega_{b3}^{ba} \underrightarrow{b}^3 \times \underrightarrow{b}^3 = -\omega_{b1}^{ba} \underrightarrow{b}^2 + \omega_{b2}^{ba} \underrightarrow{b}^1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\underrightarrow{\omega}^{ba} \times \underbrace{\begin{bmatrix} \underrightarrow{b}^1 & \underrightarrow{b}^2 & \underrightarrow{b}^3 \end{bmatrix}}_{\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}} &= \begin{bmatrix} \left(-\omega_{b2}^{ba} \underrightarrow{b}^3 + \omega_{b3}^{ba} \underrightarrow{b}^2\right) & \left(\omega_{b1}^{ba} \underrightarrow{b}^3 - \omega_{b3}^{ba} \underrightarrow{b}^1\right) & \left(-\omega_{b1}^{ba} \underrightarrow{b}^2 + \omega_{b2}^{ba} \underrightarrow{b}^1\right) \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \underrightarrow{b}^1 & \underrightarrow{b}^2 & \underrightarrow{b}^3 \end{bmatrix}}_{\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}} \underbrace{\begin{bmatrix} 0 & -\omega_{b3}^{ba} & \omega_{b2}^{ba} \\ \omega_{b3}^{ba} & 0 & -\omega_{b1}^{ba} \\ -\omega_{b2}^{ba} & \omega_{b1}^{ba} & 0 \end{bmatrix}}_{\boldsymbol{\omega}_b^{ba^{\times}}}.
\end{aligned}$$

From Equation (3.8) and the above it follows that

$$\underrightarrow{\mathcal{F}}_b^{\mathsf{T}^{\bullet a}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \boldsymbol{\omega}_b^{ba^{\times}},$$

which is Equation (3.9).

3.3.1.2 The Transport Theorem

The reason an expression for $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}^{\bullet a}}$ is of interest, as given in Equation (3.8) and Equation (3.9), goes back to Equation (3.5), that being

$$\underrightarrow{r}^{\bullet a} = \underrightarrow{r}^{\bullet b} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}^{\bullet a}} \mathbf{r}_b.$$

Substitution of Equation (3.9) into Equation (3.5) yields

$$\begin{aligned}
\underrightarrow{r}^{\bullet a} &= \underrightarrow{r}^{\bullet b} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}^{\bullet a}} \mathbf{r}_b \\
&= \underrightarrow{r}^{\bullet b} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \boldsymbol{\omega}_b^{ba^{\times}} \mathbf{r}_b \\
&= \underrightarrow{r}^{\bullet b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}.
\end{aligned}$$

The equation

$$\underrightarrow{r}^{\bullet a} = \underrightarrow{r}^{\bullet b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r} \tag{3.10}$$

is called *the Transport Theorem*. More explicitly,

$$\underrightarrow{r}^{\bullet a} = \underrightarrow{r}^{\bullet b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}$$

This is a very important equation that will be used heavily. Note that $\underrightarrow{r}$ can be *any* physical vector!

3.3.2 A Matrix Mathematics Approach to the Transport Theorem

3.3.2.1 Poisson's Equation

Returning to Equation (3.5) once more, an alternative expression for $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}\,\bullet a}$ will be derived. Recall that

$$\begin{aligned} \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{C}_{ba}, \\ \underrightarrow{\mathcal{F}}_a &= \mathbf{C}_{ba}^{\mathsf{T}} \underrightarrow{\mathcal{F}}_b, \\ \underrightarrow{\mathcal{F}}_b &= \mathbf{C}_{ba} \underrightarrow{\mathcal{F}}_a. \end{aligned}$$

The last of these three equations, $\underrightarrow{\mathcal{F}}_b = \mathbf{C}_{ba}\underrightarrow{\mathcal{F}}_a$, can be verbosely written as

$$\begin{bmatrix} \underrightarrow{b}^1 \\ \underrightarrow{b}^2 \\ \underrightarrow{b}^3 \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^3 c_{ba\,1k} \underrightarrow{a}^k \\ \sum_{k=1}^3 c_{ba\,2k} \underrightarrow{a}^k \\ \sum_{k=1}^3 c_{ba\,3k} \underrightarrow{a}^k \end{bmatrix},$$

or individually as

$$\underrightarrow{b}^{\jmath} = \sum_{k=1}^{3} c_{ba\,\jmath k} \underrightarrow{a}^k, \quad \jmath = 1, 2, 3,$$

where

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

Taking the time derivative of each $\underrightarrow{b}^{\jmath}$ with respect to $\mathcal{F}_a$ yields

$$\begin{aligned} \underrightarrow{b}^{1\,\bullet a} &= \sum_{k=1}^{3} \dot{c}_{ba\,1k} \underrightarrow{a}^k, \\ \underrightarrow{b}^{2\,\bullet a} &= \sum_{k=1}^{3} \dot{c}_{ba\,2k} \underrightarrow{a}^k, \\ \underrightarrow{b}^{3\,\bullet a} &= \sum_{k=1}^{3} \dot{c}_{ba\,3k} \underrightarrow{a}^k. \end{aligned}$$

Columnizing each $\underrightarrow{b}^{\jmath\,\bullet a}$ into $\underrightarrow{\mathcal{F}}_b^{\bullet a}$ gives

$$\underrightarrow{\mathcal{F}}_b^{\bullet a} = \dot{\mathbf{C}}_{ba} \underrightarrow{\mathcal{F}}_a, \quad \text{or} \quad \underrightarrow{\mathcal{F}}_b^{\mathsf{T}\,\bullet a} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \dot{\mathbf{C}}_{ba}^{\mathsf{T}}. \tag{3.11}$$

Wait ... what on Earth is $\dot{\mathbf{C}}_{ba}^{\mathsf{T}}$?

To find $\dot{\mathbf{C}}_{ba}^{\mathsf{T}}$, or alternatively $\dot{\mathbf{C}}_{ba}$, consider the orthonormality constraint

$$\mathbf{C}_{ba}^{\mathsf{T}} \mathbf{C}_{ba} = \mathbf{1}.$$

Taking the time derivative of $\mathbf{C}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} = \mathbf{1}$ and rearranging yields

$$\begin{aligned}
\dot{\mathbf{C}}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} + \mathbf{C}_{ba}^{\mathsf{T}}\dot{\mathbf{C}}_{ba} &= \mathbf{0}, \\
\dot{\mathbf{C}}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} &= -\mathbf{C}_{ba}^{\mathsf{T}}\dot{\mathbf{C}}_{ba} \\
&= -\left(\dot{\mathbf{C}}_{ba}^{\mathsf{T}}\mathbf{C}_{ba}\right)^{\mathsf{T}}.
\end{aligned}$$

Note that this time derivative is the time derivative of the scalars composing $\mathbf{C}_{ba}$. Defining $\mathbf{\Omega}_a = \dot{\mathbf{C}}_{ba}^{\mathsf{T}}\mathbf{C}_{ba}$ notice that

$$\mathbf{\Omega}_a = -\mathbf{\Omega}_a^{\mathsf{T}},$$

indicating that $\mathbf{\Omega}_a$ is skew symmetric. Let $\mathbf{\Omega}_a = \boldsymbol{\omega}_a^{ba^\times}$ and $\boldsymbol{\omega}_a^{ba} = \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba}$ where $\boldsymbol{\omega}_a^{ba}$ is the *angular velocity* of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ resolved in $\mathcal{F}_a$, and $\boldsymbol{\omega}_b^{ba}$ is the *angular velocity* of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ resolved in $\mathcal{F}_b$. It follows that

$$\begin{aligned}
\dot{\mathbf{C}}_{ba}^{\mathsf{T}}\mathbf{C}_{ba} &= \boldsymbol{\omega}_a^{ba^\times} \\
&= \left(\mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba}\right)^\times \\
&= \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba^\times}\mathbf{C}_{ba},
\end{aligned}$$

where the identity $\left(\mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba}\right)^\times = \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba^\times}\mathbf{C}_{ba}$ has been used. Right multiplying both sides by $\mathbf{C}_{ba}^{\mathsf{T}}$ gives

$$\begin{aligned}
\dot{\mathbf{C}}_{ba}^{\mathsf{T}}\overbrace{\mathbf{C}_{ba}\mathbf{C}_{ba}^{\mathsf{T}}}^{\mathbf{1}} &= \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba^\times}\overbrace{\mathbf{C}_{ba}\mathbf{C}_{ba}^{\mathsf{T}}}^{\mathbf{1}}, \\
\dot{\mathbf{C}}_{ba}^{\mathsf{T}} &= \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba^\times}.
\end{aligned}$$

Transposing both sides leads to

$$\dot{\mathbf{C}}_{ba} = -\boldsymbol{\omega}_b^{ba^\times}\mathbf{C}_{ba},$$

or

$$\dot{\mathbf{C}}_{ba} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{C}_{ba} = \mathbf{0} \tag{3.12}$$

Equation (3.12) is called *Poisson's Equation*. Returning to Equation (3.11) and using Poisson's Equation

$$\begin{aligned}
\underrightarrow{\mathcal{F}}_b^{\mathsf{T}^{\bullet a}} &= \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\dot{\mathbf{C}}_{ba}^{\mathsf{T}} \\
&= \underbrace{\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{C}_{ab}}_{\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}}\boldsymbol{\omega}_b^{ba^\times} \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{ba^\times},
\end{aligned}$$

which is Equation (3.9) derived previously.

3.3.2.2 Back to the Transport Theorem

Again, the reason an expression for $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}\,\bullet a}$, as given in Equation (3.8) and Equation (3.9), is of interest goes back to Equation (3.5),

$$\underrightarrow{r}^{\bullet a} = \underrightarrow{r}^{\bullet b} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}\,\bullet a} \mathbf{r}_b.$$

Substitution of Equation (3.9) into Equation (3.5) yields the Transport Theorem, that being

$$\begin{aligned} \underrightarrow{r}^{\bullet a} &= \underrightarrow{r}^{\bullet b} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \boldsymbol{\omega}_b^{ba^\times} \mathbf{r}_b \\ &= \underrightarrow{r}^{\bullet b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}. \end{aligned}$$

3.4 More on Angular Velocity

3.4.1 The Addition of Angular Velocity Physical Vectors

Consider $\mathcal{F}_a$, $\mathcal{F}_q$, $\mathcal{F}_b$ along with the relationship $\mathbf{C}_{ba} = \mathbf{C}_{bq}\mathbf{C}_{qa}$. The angular velocity of $\mathcal{F}_q$ relative to $\mathcal{F}_a$ is $\underrightarrow{\omega}^{qa}$, the angular velocity of $\mathcal{F}_b$ relative to $\mathcal{F}_q$ is $\underrightarrow{\omega}^{bq}$, and the angular velocity of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ is $\underrightarrow{\omega}^{ba}$. Is it true that

$$\underrightarrow{\omega}^{ba} = \underrightarrow{\omega}^{bq} + \underrightarrow{\omega}^{qa},$$

which is to ask the question "do physical angular velocity vectors add and yield another physical angular velocity vector?" This answer is "yes" as shown in the following proof.

Theorem 3.1. Given $\mathcal{F}_a$, $\mathcal{F}_q$, $\mathcal{F}_b$ and the relationship $\mathbf{C}_{ba} = \mathbf{C}_{bq}\mathbf{C}_{qa}$, $\underrightarrow{\omega}^{ba} = \underrightarrow{\omega}^{bq} + \underrightarrow{\omega}^{qa}$, which is to say physical angular velocity vectors add and yield another physical angular velocity vector.

Proof. Consider $\mathcal{F}_a$, $\mathcal{F}_q$, $\mathcal{F}_b$, and $\mathbf{C}_{ba} = \mathbf{C}_{bq}\mathbf{C}_{qa}$. From Poisson's equation,

$$\begin{aligned} \dot{\mathbf{C}}_{qa} &= -\boldsymbol{\omega}_q^{qa^\times} \mathbf{C}_{qa}, \\ \dot{\mathbf{C}}_{bq} &= -\boldsymbol{\omega}_b^{bq^\times} \mathbf{C}_{bq}, \\ \dot{\mathbf{C}}_{ba} &= -\boldsymbol{\omega}_b^{ba^\times} \mathbf{C}_{ba}. \end{aligned}$$

Taking the time derivative of $\mathbf{C}_{ba} = \mathbf{C}_{bq}\mathbf{C}_{qa}$ and using the expressions for $\dot{\mathbf{C}}_{qa}$ and $\dot{\mathbf{C}}_{bq}$ given above yields

$$\begin{aligned} \dot{\mathbf{C}}_{ba} &= \dot{\mathbf{C}}_{bq}\mathbf{C}_{qa} + \mathbf{C}_{bq}\dot{\mathbf{C}}_{qa} \\ &= -\boldsymbol{\omega}_b^{bq^\times} \underbrace{\mathbf{C}_{bq}\mathbf{C}_{qa}}_{\mathbf{C}_{ba}} - \mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa^\times}\mathbf{C}_{qa} \\ &= -\boldsymbol{\omega}_b^{bq^\times}\mathbf{C}_{ba} - \mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa^\times} \underbrace{\mathbf{C}_{bq}^{\mathsf{T}}\mathbf{C}_{bq}}_{\mathbf{1}} \mathbf{C}_{qa} \\ &= -\boldsymbol{\omega}_b^{bq^\times}\mathbf{C}_{ba} - \underbrace{\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa^\times}\mathbf{C}_{bq}^{\mathsf{T}}}_{\left(\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa}\right)^\times} \underbrace{\mathbf{C}_{bq}\mathbf{C}_{qa}}_{\mathbf{C}_{ba}} \\ &= -\boldsymbol{\omega}_b^{bq^\times}\mathbf{C}_{ba} - \left(\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa}\right)^\times \mathbf{C}_{ba}, \end{aligned}$$

where the relationships $\mathbf{C}_{bq}^{\mathsf{T}}\mathbf{C}_{bq} = \mathbf{1}$ and $\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa^\times}\mathbf{C}_{bq}^{\mathsf{T}} = (\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa})^\times$ have been used. Factoring out $\mathbf{C}_{ba}$ on the left-hand-side results in

$$\begin{aligned}\dot{\mathbf{C}}_{ba} &= -\left(\boldsymbol{\omega}_b^{ba^\times} + \left(\mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa}\right)^\times\right)\mathbf{C}_{ba} \\ &= -\left(\boldsymbol{\omega}_b^{ba} + \mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa}\right)^\times \mathbf{C}_{ba},\end{aligned}$$

where the fact that the cross operator is linear has been exploited. Substitution of $\dot{\mathbf{C}}_{ba} = -\boldsymbol{\omega}_b^{ba^\times}\mathbf{C}_{ba}$ on the left-hand-side and simplifying gives

$$\begin{aligned}\boldsymbol{\omega}_b^{ba^\times} &= \left(\boldsymbol{\omega}_b^{bq} + \mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa}\right)^\times, \\ \boldsymbol{\omega}_b^{ba} &= \boldsymbol{\omega}_b^{bq} + \mathbf{C}_{bq}\boldsymbol{\omega}_q^{qa} \\ &= \boldsymbol{\omega}_b^{bq} + \boldsymbol{\omega}_b^{qa}.\end{aligned}$$

Pre-multiplying by $\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}$ results in

$$\begin{aligned}\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{ba} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{bq} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{qa}, \\ \underrightarrow{\omega}^{ba} &= \underrightarrow{\omega}^{bq} + \underrightarrow{\omega}^{qa}.\end{aligned}$$

This completes the proof.

$\square$

3.4.2 The Relationship Between Angular Velocity and Rates of Axis/Angle Parameters

Recall that $\mathbf{C}_{ba}(\mathbf{a},\phi) = \cos\phi\mathbf{1} + (1-\cos\phi)\mathbf{a}\mathbf{a}^{\mathsf{T}} - \sin\phi\mathbf{a}^\times$ where $\mathbf{a}\in\mathbb{S}^2$. The relationship between $\mathbf{a}$, ϕ, $\dot{\mathbf{a}}$, $\dot{\phi}$, and $\boldsymbol{\omega}_b^{ba}$ is [1, pp. 25]

$$\boldsymbol{\omega}_b^{ba} = \underbrace{\left[\begin{array}{cc} \sin\phi\mathbf{1} - (1-\cos\phi)\mathbf{a}^\times & \mathbf{a}\end{array}\right]}_{\mathbf{S}_b^{ba}(\mathbf{a},\phi)}\left[\begin{array}{c}\dot{\mathbf{a}} \\ \dot{\phi}\end{array}\right], \tag{3.13}$$

or

$$\left[\begin{array}{c}\dot{\mathbf{a}} \\ \dot{\phi}\end{array}\right] = \underbrace{\left[\begin{array}{c}\frac{1}{2}(\mathbf{a}^\times - \cot\frac{\phi}{2}\mathbf{a}^\times\mathbf{a}^\times) \\ \mathbf{a}^{\mathsf{T}}\end{array}\right]}_{\boldsymbol{\Gamma}_b^{ba}(\mathbf{a},\phi)}\boldsymbol{\omega}_b^{ba}. \tag{3.14}$$

In holonomic form the constraint is $\mathbf{a}^{\mathsf{T}}\mathbf{a} - 1 = 0$. Taking the time-derivative of $\mathbf{a}^{\mathsf{T}}\mathbf{a} - 1 = 0$ the constraint can also be written

$$\underbrace{\left[\begin{array}{cc} 2\mathbf{a}^{\mathsf{T}} & 0\end{array}\right]}_{\boldsymbol{\Xi}_b^{ba}(\mathbf{a},\phi)}\left[\begin{array}{c}\dot{\mathbf{a}} \\ \dot{\phi}\end{array}\right] = 0. \tag{3.15}$$

From Equations (3.14) and (3.15) notice that

$$\boldsymbol{\Xi}_b^{ba}(\mathbf{a},\phi)\boldsymbol{\Gamma}_b^{ba}(\mathbf{a},\phi) = \mathbf{0}. \tag{3.16}$$

As such, the columns of $\boldsymbol{\Gamma}_b^{ba}(\cdot,\cdot)$ lie in the null space of $\boldsymbol{\Xi}_b^{ba}(\cdot,\cdot)$ [33]. Also, using Equations (3.14) and (3.13)

$$\mathbf{S}_b^{ba}(\mathbf{a},\phi)\boldsymbol{\Gamma}_b^{ba}(\mathbf{a},\phi) = \mathbf{1}. \tag{3.17}$$

3.4.2.1 Kinematic Singularity

When $\phi = n2\pi$, $n \in \mathbb{Z}$, the axis $\mathbf{a}$ is undefined. Because $\mathbf{a}$ is undefined when $\phi = n2\pi$, $n \in \mathbb{Z}$, the matrix $\mathbf{S}_b^{ba}(\cdot,\cdot)$ in Equation (3.13) is not defined. Additionally, the term $\cot(\phi/2) \to \infty$ as $\phi \to n2\pi$ for any $n \in \mathbb{Z}$ resulting in $||\dot{\mathbf{a}}||_2 \to \infty$. This situation is referred to as a kinematic singularity.

3.4.3 The Relationship Between Angular Velocity and Quaternion Rates

In terms of quaternions, the direction cosine matrix is $\mathbf{C}_{ba}(\boldsymbol{\epsilon},\eta) = (\eta^2 - \boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon})\mathbf{1} + 2\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\mathsf{T} - 2\eta\boldsymbol{\epsilon}^\times = \mathbf{1} + 2\boldsymbol{\epsilon}^\times\boldsymbol{\epsilon}^\times - 2\eta\boldsymbol{\epsilon}^\times$ [1, pp. 18]. The relationship between $\boldsymbol{\epsilon}$, η, $\dot{\boldsymbol{\epsilon}}$, $\dot{\eta}$, and $\boldsymbol{\omega}_b^{ba}$ is [1, pp. 26, 31]

$$\boldsymbol{\omega}_b^{ba} = \underbrace{\begin{bmatrix} 2(\eta\mathbf{1} - \boldsymbol{\epsilon}^\times) & -2\boldsymbol{\epsilon} \end{bmatrix}}_{\mathbf{S}_b^{ba}(\boldsymbol{\epsilon},\eta)} \begin{bmatrix} \dot{\boldsymbol{\epsilon}} \\ \dot{\eta} \end{bmatrix}, \tag{3.18}$$

or

$$\begin{bmatrix} \dot{\boldsymbol{\epsilon}} \\ \dot{\eta} \end{bmatrix} = \underbrace{\frac{1}{2}\begin{bmatrix} \eta\mathbf{1} + \boldsymbol{\epsilon}^\times \\ -\boldsymbol{\epsilon}^\mathsf{T} \end{bmatrix}}_{\boldsymbol{\Gamma}_b^{ba}(\boldsymbol{\epsilon},\eta)} \boldsymbol{\omega}_b^{ba}. \tag{3.19}$$

Taking the time derivative of the constraint $\boldsymbol{\epsilon}^\mathsf{T}\boldsymbol{\epsilon} + \eta^2 = 1$ yields

$$\underbrace{\begin{bmatrix} 2\boldsymbol{\epsilon}^\mathsf{T} & 2\eta \end{bmatrix}}_{\boldsymbol{\Xi}_b^{ba}(\boldsymbol{\epsilon},\eta)} \begin{bmatrix} \dot{\boldsymbol{\epsilon}} \\ \dot{\eta} \end{bmatrix} = 0. \tag{3.20}$$

Using Equations (3.19) and (3.20) it follows that

$$\boldsymbol{\Xi}_b^{ba}(\boldsymbol{\epsilon},\eta)\boldsymbol{\Gamma}_b^{ba}(\boldsymbol{\epsilon},\eta) = \mathbf{0}. \tag{3.21}$$

From Equations (3.21) it is clear that $\boldsymbol{\Xi}_b^{ba}(\cdot,\cdot)$ and $\boldsymbol{\Gamma}_b^{ba}(\cdot,\cdot)$ are orthogonal complements [33]. Also, Equations (3.19) and (3.18) give

$$\mathbf{S}_b^{ba}(\boldsymbol{\epsilon},\eta)\boldsymbol{\Gamma}_b^{ba}(\boldsymbol{\epsilon},\eta) = \mathbf{1}. \tag{3.22}$$

3.4.3.1 No Kinematic Singularity

The matrix $\mathbf{S}_b^{ba}(\cdot,\cdot)$ in Equation (3.18) is full rank for all $\mathbf{q} \in \mathbb{S}^3$ where $\mathbf{q} = [\boldsymbol{\epsilon}^\mathsf{T} \quad \eta]^\mathsf{T}$. The matrix $\boldsymbol{\Gamma}_b^{ba}(\cdot,\cdot)$ is also full rank for all $\mathbf{q} \in \mathbb{S}^3$. As such, although quaternions double-cover $SO(3)$, quaternions do not suffer from a kinematic singularity.

3.4.4 The Relationship Between Angular Velocity and Rates of Rodrigues Parameters

In terms of Rodrigues parameters, the direction cosine matrix can be written as $\mathbf{C}_{qp}(\mathbf{p}) = \mathbf{1} + \frac{2}{1+\mathbf{p}^\mathsf{T}\mathbf{p}}\left(\mathbf{p}^\times\mathbf{p}^\times - \mathbf{p}^\times\right)$ [1, pp. 30-31] [29, pp. 48-50]. The relationship between $\mathbf{p}$, $\dot{\mathbf{p}}$, and $\boldsymbol{\omega}_b^{ba}$ is [1, pp. 30-31] [29, pp. 72] [34–36]

$$\boldsymbol{\omega}_b^{ba} = \underbrace{\frac{2}{1+\mathbf{p}^\mathsf{T}\mathbf{p}}\left(\mathbf{1} - \mathbf{p}^\times\right)}_{\mathbf{S}_b^{ba}(\mathbf{p})} \dot{\mathbf{p}}, \tag{3.23}$$

or

$$\dot{\mathbf{p}} = \underbrace{\frac{1}{2}\left(\mathbf{1} + \mathbf{p}^\times + \mathbf{p}\mathbf{p}^\mathsf{T}\right)}_{\mathbf{\Gamma}_b^{ba}(\mathbf{p})} \boldsymbol{\omega}_b^{ba}. \tag{3.24}$$

3.4.4.1 Kinematic Singularity

Recall that $\mathbf{p} = \boldsymbol{\epsilon}/\eta = \mathbf{a}\tan(\phi/2)$. As $\eta \to 0$ the two-norm of the Rodrigues column matrix goes to infinity, that is $||\mathbf{p}||_2 \to \infty$, resulting in $||\dot{\mathbf{p}}||_2 \to \infty$. This is a kinematic singularity.

3.4.5 The Relationship Between Angular Velocity and Rates of MRPs

The direction cosine matrix can be parameterized using MRPs: $\mathbf{C}_{qp}(\mathbf{s}) = \mathbf{1} + \frac{4}{(1+\mathbf{s}^\mathsf{T}\mathbf{s})^2}\left(2\mathbf{s}^\times\mathbf{s}^\times - (1 - \mathbf{s}^\mathsf{T}\mathbf{s})\mathbf{s}^\times\right)$ [29, pp. 51]. The relationship between $\mathbf{s}$, $\dot{\mathbf{s}}$, and $\boldsymbol{\omega}_b^{ba}$ is [29, pp. 72] [34–36]

$$\boldsymbol{\omega}_b^{ba} = \underbrace{\frac{4}{1 + \mathbf{s}^\mathsf{T}\mathbf{s}}\left(\mathbf{1} + 2\left(\frac{\mathbf{s}^\times\mathbf{s}^\times - \mathbf{s}^\times}{1 + \mathbf{s}^\mathsf{T}\mathbf{s}}\right)\right)}_{\mathbf{S}_b^{ba}(\mathbf{s})} \dot{\mathbf{s}}, \tag{3.25}$$

or

$$\dot{\mathbf{s}} = \underbrace{\frac{1}{2}\left(\left(\frac{1 - \mathbf{s}^\mathsf{T}\mathbf{s}}{2}\right)\mathbf{1} + \mathbf{s}^\times + \mathbf{s}\mathbf{s}^\mathsf{T}\right)}_{\mathbf{\Gamma}_b^{ba}(\mathbf{s})} \boldsymbol{\omega}_b^{ba}. \tag{3.26}$$

3.4.5.1 Kinematic Singularity

Recall that $\mathbf{s} = \boldsymbol{\epsilon}/(1+\eta) = \mathbf{a}\tan(\phi/4)$. As $\phi \to \pm 2\pi$ then $||\mathbf{s}||_2 \to \infty$ and $||\dot{\mathbf{s}}||_2 \to \infty$. This is a kinematic singularity.

3.4.6 The Relationship Between Angular Velocity and Euler Angles Rates

Consider a $1-2-3$ Euler angle sequence parameterizing $\mathbf{C}_{ba}$:

$$\mathbf{C}_{ba} = \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{C}_1(\delta).$$

How is $\boldsymbol{\omega}_b^{ba}$ related to $\dot{\delta}$, $\dot{\epsilon}$, and $\dot{\phi}$? To answer this question, define two intermediate frames, $\mathcal{F}_k$ and $\mathcal{F}_\ell$. The relationship between $\mathcal{F}_a$, $\mathcal{F}_k$, $\mathcal{F}_\ell$, and $\mathcal{F}_b$ is

$$\mathcal{F}_a \text{ — (Rot. about } \underrightarrow{a}^1 \text{ or } \underrightarrow{k}^1) \to \mathcal{F}_k \text{ — (Rot. about } \underrightarrow{k}^2 \text{ or } \underrightarrow{\ell}^2) \to \mathcal{F}_l \text{ — (Rot. about } \underrightarrow{\ell}^3 \text{ or } \underrightarrow{b}^3) \to \mathcal{F}_b,$$

which in terms of $\mathbf{C}_3(\phi)$, $\mathbf{C}_2(\epsilon)$, and $\mathbf{C}_1(\delta)$ is

$$\mathcal{F}_a \overset{\mathbf{C}_{ka}=\mathbf{C}_1(\delta)}{\longrightarrow} \mathcal{F}_k \overset{\mathbf{C}_{\ell k}=\mathbf{C}_2(\epsilon)}{\longrightarrow} \mathcal{F}_l \overset{\mathbf{C}_{b\ell}=\mathbf{C}_3(\phi)}{\longrightarrow} \mathcal{F}_b.$$

The approach to finding a relationship between $\boldsymbol{\omega}_b^{ba}$ and $\dot{\delta}$, $\dot{\epsilon}$, and $\dot{\phi}$ is to exploit the fact that angular velocity physical vectors add as shown in Section 3.4.1. First, however, an expressions for $\dot{\mathbf{C}}_{ka} = \dot{\mathbf{C}}_1(\delta)$, $\dot{\mathbf{C}}_{\ell k} =$

$\dot{\mathbf{C}}_2(\epsilon)$, and $\dot{\mathbf{C}}_{b\ell} = \dot{\mathbf{C}}_3(\phi)$ is needed. Starting with

$$\mathbf{C}_{b\ell} = \mathbf{C}_3(\phi) = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$\dot{\mathbf{C}}_3(\phi)$ is

$$\begin{aligned}
\dot{\mathbf{C}}_3(\phi) &= \begin{bmatrix} -\dot{\phi}\sin\phi & \dot{\phi}\cos\phi & 0 \\ -\dot{\phi}\cos\phi & -\dot{\phi}\sin\phi & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= -\begin{bmatrix} 0 & -\dot{\phi} & 0 \\ \dot{\phi} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 0 \\ \dot{\phi} \end{bmatrix}^{\times} \mathbf{C}_3(\phi) = -\dot{\phi}\mathbf{1}_3^{\times}\mathbf{C}_3(\phi),
\end{aligned}$$

where $\mathbf{1}_3 = [0 \;\; 0 \;\; 1]^{\mathsf{T}}$. In a similar fashion it can be shown that

$$\dot{\mathbf{C}}_2(\epsilon) = -\begin{bmatrix} 0 \\ \dot{\epsilon} \\ 0 \end{bmatrix}^{\times} \mathbf{C}_2(\epsilon) = -\dot{\epsilon}\mathbf{1}_2^{\times}\mathbf{C}_2(\epsilon), \quad \text{and} \quad \dot{\mathbf{C}}_1(\delta) = -\begin{bmatrix} \dot{\delta} \\ 0 \\ 0 \end{bmatrix}^{\times} \mathbf{C}_1(\delta) = -\dot{\delta}\mathbf{1}_1^{\times}\mathbf{C}_1(\delta),$$

where $\mathbf{1}_2 = [0 \;\; 1 \;\; 0]^{\mathsf{T}}$ and $\mathbf{1}_1 = [1 \;\; 0 \;\; 0]^{\mathsf{T}}$. Given that

$$\begin{aligned}
\dot{\mathbf{C}}_{ka} = -\boldsymbol{\omega}_k^{ka^{\times}}\mathbf{C}_{ka} \quad &\Leftrightarrow \quad \dot{\mathbf{C}}_1(\delta) = -\dot{\delta}\mathbf{1}_1^{\times}\mathbf{C}_1(\delta), \\
\dot{\mathbf{C}}_{\ell k} = -\boldsymbol{\omega}_\ell^{\ell k^{\times}}\mathbf{C}_{\ell k} \quad &\Leftrightarrow \quad \dot{\mathbf{C}}_2(\epsilon) = -\dot{\epsilon}\mathbf{1}_2^{\times}\mathbf{C}_2(\epsilon), \\
\dot{\mathbf{C}}_{b\ell} = -\boldsymbol{\omega}_b^{b\ell^{\times}}\mathbf{C}_{b\ell} \quad &\Leftrightarrow \quad \dot{\mathbf{C}}_3(\phi) = -\dot{\phi}\mathbf{1}_3^{\times}\mathbf{C}_3(\phi),
\end{aligned}$$

it follows that

$$\underrightarrow{\omega}^{ka} = \underrightarrow{\mathcal{F}}_k^{\mathsf{T}}\boldsymbol{\omega}_k^{ka} = \underrightarrow{\mathcal{F}}_k^{\mathsf{T}}\mathbf{1}_1\dot{\delta}, \quad \boldsymbol{\omega}_k^{ka} = \mathbf{1}_1\dot{\delta}, \tag{3.27}$$

$$\underrightarrow{\omega}^{\ell k} = \underrightarrow{\mathcal{F}}_\ell^{\mathsf{T}}\boldsymbol{\omega}_\ell^{\ell k} = \underrightarrow{\mathcal{F}}_\ell^{\mathsf{T}}\mathbf{1}_2\dot{\epsilon}, \quad \boldsymbol{\omega}_\ell^{\ell k} = \mathbf{1}_2\dot{\epsilon}, \tag{3.28}$$

$$\underrightarrow{\omega}^{bl} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{b\ell} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{1}_3\dot{\phi}, \quad \boldsymbol{\omega}_b^{bl} = \mathbf{1}_3\dot{\phi}. \tag{3.29}$$

Next, using the fact that that physical angular velocity vectors add to yield another physical angular velocity vector (see Section 3.4.1), it follows that

$$\begin{aligned}
\underrightarrow{\omega}^{ba} &= \underrightarrow{\omega}^{bl} + \underrightarrow{\omega}^{\ell k} + \underrightarrow{\omega}^{ka}, \\
\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{ba} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{1}_3\dot{\phi} + \underrightarrow{\mathcal{F}}_\ell^{\mathsf{T}}\mathbf{1}_2\dot{\epsilon} + \underrightarrow{\mathcal{F}}_k^{\mathsf{T}}\mathbf{1}_1\dot{\delta} \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\left(\mathbf{1}_3\dot{\phi} + \mathbf{C}_{b\ell}\mathbf{1}_2\dot{\epsilon} + \mathbf{C}_{bk}\mathbf{1}_1\dot{\delta}\right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\underbrace{\left(\mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_1\dot{\delta} + \mathbf{C}_3(\phi)\mathbf{1}_2\dot{\epsilon} + \mathbf{1}_3\dot{\phi}\right)}_{\boldsymbol{\omega}_b^{ba}} \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\boldsymbol{\omega}_b^{ba},
\end{aligned}$$

where $\mathbf{C}_{b\ell} = \mathbf{C}_3(\phi)$ and $\mathbf{C}_{bk} = \mathbf{C}_{b\ell}\mathbf{C}_{\ell k} = \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)$. Further, $\boldsymbol{\omega}_b^{ba}$ can be written as

$$\boldsymbol{\omega}_b^{ba} = \underbrace{\begin{bmatrix} \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_1 & \mathbf{C}_3(\phi)\mathbf{1}_2 & \mathbf{1}_3 \end{bmatrix}}_{\mathbf{S}_b^{ba}(\epsilon,\phi)} \underbrace{\begin{bmatrix} \dot{\delta} \\ \dot{\epsilon} \\ \dot{\phi} \end{bmatrix}}_{\dot{\boldsymbol{\theta}}^{ba}} = \mathbf{S}_b^{ba}(\epsilon,\phi)\dot{\boldsymbol{\theta}}^{ba},$$

where $\mathbf{S}_b^{ba}(\cdot,\cdot)$ is the mapping matrix from the Euler angle rates, $\dot{\boldsymbol{\theta}}^{ba}$, to the angular velocity of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ resolved in $\mathcal{F}_b$. It is straightforward to show that

$$\mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_3 = \begin{bmatrix} \cos\phi\cos\epsilon \\ -\sin\phi\cos\epsilon \\ \sin\epsilon \end{bmatrix}, \quad \mathbf{C}_3(\phi)\mathbf{1}_2 = \begin{bmatrix} \sin\phi \\ \cos\phi \\ 0 \end{bmatrix},$$

and therefore

$$\begin{bmatrix} \omega_{b1}^{ba} \\ \omega_{b2}^{ba} \\ \omega_{b3}^{ba} \end{bmatrix} = \begin{bmatrix} \cos\phi\cos\epsilon & \sin\phi & 0 \\ -\sin\phi\cos\epsilon & \cos\phi & 0 \\ \sin\epsilon & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\delta} \\ \dot{\epsilon} \\ \dot{\phi} \end{bmatrix}$$

3.4.6.1 Why is $\mathbf{S}_b^{ba}(\cdot,\cdot)$ Only a Function of Two of the Three Euler Angles?

Notice that $\mathbf{S}_b^{ba}(\cdot,\cdot)$ is a function of ϵ and ϕ, not ϵ, ϕ, and δ. To understand why, rather than resolving $\underrightarrow{\omega}^{ka}$, $\underrightarrow{\omega}^{\ell a}$, and $\underrightarrow{\omega}^{b\ell}$ in reference frames $\mathcal{F}_k$, $\mathcal{F}_\ell$, and $\mathcal{F}_b$, respectively, as is done in Equations (3.27), (3.28), and (3.29), resolve $\underrightarrow{\omega}^{ka}$, $\underrightarrow{\omega}^{\ell a}$, and $\underrightarrow{\omega}^{b\ell}$ in reference frames $\mathcal{F}_a$, $\mathcal{F}_k$, and $\mathcal{F}_\ell$, respectively, that being

$$\begin{aligned}
\underrightarrow{\omega}^{ka} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\boldsymbol{\omega}_a^{ka} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{C}_{ka}^{\mathsf{T}}\mathbf{1}_1\dot{\delta} = \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{C}_1^{\mathsf{T}}(\delta)\mathbf{1}_1\dot{\delta}, \\
\underrightarrow{\omega}^{\ell k} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_k^{\mathsf{T}}\boldsymbol{\omega}_k^{\ell k} = \underrightarrow{\boldsymbol{\mathcal{F}}}_k^{\mathsf{T}}\mathbf{C}_{\ell k}^{\mathsf{T}}\mathbf{1}_2\dot{\epsilon} = \underrightarrow{\boldsymbol{\mathcal{F}}}_k^{\mathsf{T}}\mathbf{C}_2^{\mathsf{T}}(\epsilon)\mathbf{1}_2\dot{\epsilon}, \\
\underrightarrow{\omega}^{bl} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_\ell^{\mathsf{T}}\boldsymbol{\omega}_\ell^{b\ell} = \underrightarrow{\boldsymbol{\mathcal{F}}}_\ell^{\mathsf{T}}\mathbf{C}_{b\ell}^{\mathsf{T}}\mathbf{1}_3\dot{\phi} = \underrightarrow{\boldsymbol{\mathcal{F}}}_\ell^{\mathsf{T}}\mathbf{C}_3^{\mathsf{T}}(\phi)\mathbf{1}_3\dot{\phi},
\end{aligned}$$

where $\mathbf{C}_{ka} = \mathbf{C}_1(\delta)$, $\mathbf{C}_{\ell k} = \mathbf{C}_2(\epsilon)$, and $\mathbf{C}_{b\ell} = \mathbf{C}_3(\phi)$. Notice that $\mathbf{C}_1^{\mathsf{T}}(\delta)\mathbf{1}_1 = \mathbf{1}_1$, $\mathbf{C}_2^{\mathsf{T}}(\epsilon)\mathbf{1}_2 = \mathbf{1}_2$, and $\mathbf{C}_3^{\mathsf{T}}(\phi)\mathbf{1}_3 = \mathbf{1}_3$. Therefore,

$$\begin{aligned}
\underrightarrow{\omega}^{ka} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{1}_1\dot{\delta}, \\
\underrightarrow{\omega}^{\ell k} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_k^{\mathsf{T}}\mathbf{1}_2\dot{\epsilon}, \\
\underrightarrow{\omega}^{bl} &= \underrightarrow{\boldsymbol{\mathcal{F}}}_\ell^{\mathsf{T}}\mathbf{1}_3\dot{\phi}.
\end{aligned}$$

Exploiting the fact that angular velocity physical vectors can be added,

$$\begin{aligned}
\underrightarrow{\omega}^{ba} &= \underrightarrow{\omega}^{bl} + \underrightarrow{\omega}^{\ell k} + \underrightarrow{\omega}^{ka} \\
&= \underrightarrow{\boldsymbol{\mathcal{F}}}_\ell^{\mathsf{T}}\mathbf{1}_3\dot{\phi} + \underrightarrow{\boldsymbol{\mathcal{F}}}_k^{\mathsf{T}}\mathbf{1}_2\dot{\epsilon} + \underrightarrow{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}}\mathbf{1}_1\dot{\delta} \\
&= \underrightarrow{\boldsymbol{\mathcal{F}}}_b^{\mathsf{T}}\left(\mathbf{C}_{b\ell}\mathbf{1}_3\dot{\phi} + \mathbf{C}_{bk}\mathbf{1}_2\dot{\epsilon} + \mathbf{C}_{ba}\mathbf{1}_1\dot{\delta}\right).
\end{aligned}$$

Recalling that $\mathbf{C}_{b\ell}\mathbf{1}_3 = \mathbf{C}_3(\phi)\mathbf{1}_3 = \mathbf{1}_3$, $\mathbf{C}_{bk}\mathbf{1}_2 = \mathbf{C}_{b\ell}\mathbf{C}_{\ell k}\mathbf{1}_2 = \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_2 = \mathbf{C}_3(\phi)\mathbf{1}_2$, and $\mathbf{C}_{ba}\mathbf{1}_1 = \mathbf{C}_{b\ell}\mathbf{C}_{\ell k}\mathbf{C}_{ka}\mathbf{1}_1 = \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{C}_1(\delta)\mathbf{1}_1 = \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_1$ it follows that

$$\begin{aligned} \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \boldsymbol{\omega}_b^{ba} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{C}_3(\phi)\mathbf{1}_3\dot{\phi} + \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_2\dot{\epsilon} + \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{C}_1(\delta)\mathbf{1}_1\dot{\delta} \right) \\ &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{1}_3\dot{\phi} + \mathbf{C}_3(\phi)\mathbf{1}_2\dot{\epsilon} + \mathbf{C}_3(\phi)\mathbf{C}_2(\epsilon)\mathbf{1}_1\dot{\delta} \right). \end{aligned}$$

Notice that the direction cosine matrix $\mathbf{C}_1(\delta)$ drops out owing to the fact that $\mathbf{C}_1(\delta)\mathbf{1}_1 = \mathbf{1}_1$. This is why $\mathbf{S}_b^{ba}(\cdot,\cdot)$ is not a function of ϵ, ϕ, and δ, but rather only a function of ϵ and ϕ.

3.4.6.2 Kinematic Singularity

Often it is necessary to solve for the Euler angle rates, $\dot{\boldsymbol{\theta}}^{ba}$, given $\boldsymbol{\omega}_b^{ba}$ and ϵ and δ, which means inverting the mapping matrix $\mathbf{S}_b^{ba}(\cdot,\cdot)$. A natural question to ask is if or if not $\mathbf{S}_b^{ba}(\cdot,\cdot)$ is always invertible. Recall that a matrix is invertible if it is nonsingular, and a matrix is nonsingular if its determinant does not equal zero. Taking the determinant of $\mathbf{S}_b^{ba}(\cdot,\cdot)$ results in

$$\det \mathbf{S}_b^{ba}(\epsilon,\phi) = \cos^2\phi\cos\epsilon + \sin^2\phi\cos\epsilon = \cos\epsilon.$$

Thus, $\det \mathbf{S}_b^{ba}(\epsilon,\phi) = 0$ for any $\epsilon = \pi/2 \pm n\pi$, $n \in \mathbb{Z}$. When $\epsilon = \pi/2 \pm n\pi$, $n \in \mathbb{Z}$, given $\boldsymbol{\omega}_b^{ba}$ and ϵ and δ, $\dot{\boldsymbol{\theta}}^{ba}$ cannot be solved for. This is a kinematic singularity.

3.4.7 Deficiencies of Direction Cosine Matrix Parameterizations - Kinematic Singularites

Recall that the direction cosine matrix is a global and unique representation of attitude. All parameterizations of the direction cosine matrix are deficient in some way. In Section 2.4.8.8 deficiencies of direction cosine matrix parameterizations were discussed. In addition to these deficiencies, direction cosine matrix parameterizations often have kinematic singularities, as summarized in Table 3.1.

Parameterization	Global	Unique	Kinematic Singularity
Direction Cosine Matrix	Yes	Yes	No
Axis/Angle Parameters	Yes	No	Yes
Quaternions	Yes	No	No
Rodrigues Parameters	No	Yes	Yes
MRPs	Yes	No	Yes
Euler Angle Sequence	Yes	No	Yes

Table 3.1: Attitude parameterizations and if they possess a kinematic singularity.

Given that all direction cosine matrix parameterizations except the quaternion parameterization possess singularities, it seems logical to use quaternions over any other parameterization. Alternatively, one could not use any parameterization at all, and just use the direction cosine matrix directly.

3.5 Velocity and Acceleration of a Point

As shown in Figure (3.3) consider points w, y, and z, along with the two reference frames $\mathcal{F}_a$ and $\mathcal{F}_b$. The position of point z relative to point w is denoted $\underrightarrow{r}^{zw}$ and is equal to

$$\underrightarrow{r}^{zw} = \underrightarrow{r}^{zy} + \underrightarrow{r}^{yw} = \underrightarrow{r}^{yw} + \underrightarrow{r}^{zy}.$$

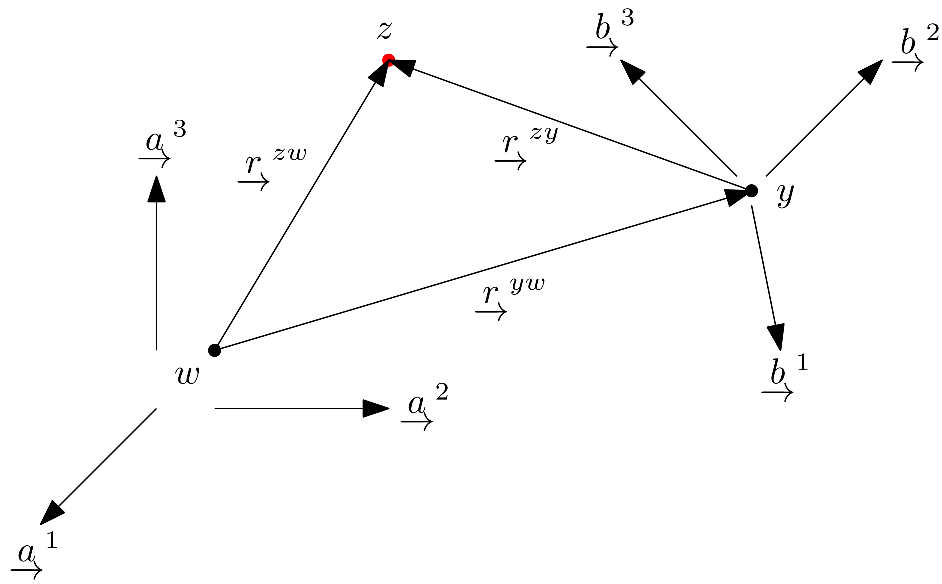


Figure 3.3: The point z and frames $\mathcal{F}_a$ and $\mathcal{F}_b$.

Using the Transport Theorem the *velocity of point z relative to point w with respect to $\mathcal{F}_a$* is

$$
\begin{aligned}
\underset{\to}{r}^{zw\,\bullet a} &= \underset{\to}{r}^{yw\,\bullet a} + \underset{\to}{r}^{zy\,\bullet a} \\
&= \underset{\to}{r}^{yw\,\bullet a} + \underset{\to}{r}^{zy\,\bullet b} + \underset{\to}{\omega}^{ba} \times \underset{\to}{r}^{zy}.
\end{aligned}
$$

To be explicit, write

$$
\begin{aligned}
\underset{\to}{v}^{zw/a} &= \underset{\to}{r}^{zw\,\bullet a} = \text{the velocity of point } z \text{ relative to point } w \text{ w.r.t. } \mathcal{F}_a, \\
\underset{\to}{v}^{yw/a} &= \underset{\to}{r}^{yw\,\bullet a} = \text{the velocity of point } y \text{ relative to point } w \text{ w.r.t. } \mathcal{F}_a, \\
\underset{\to}{v}^{zy/b} &= \underset{\to}{r}^{zy\,\bullet b} = \text{the velocity of point } z \text{ relative to point } y \text{ w.r.t. } \mathcal{F}_b, \\
\underset{\to}{v}^{zy/a} &= \underset{\to}{v}^{zy/b} + \underset{\to}{\omega}^{ba} \times \underset{\to}{r}^{zy} = \text{the velocity of point } z \text{ relative to point } y \text{ w.r.t. } \mathcal{F}_a.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\underset{\to}{v}^{zw/a} &= \underset{\to}{v}^{yw/a} + \underset{\to}{v}^{zy/a} \\
&= \underset{\to}{v}^{yw/a} + \underbrace{\underset{\to}{v}^{zy/b} + \underset{\to}{\omega}^{ba} \times \underset{\to}{r}^{zy}}_{\underset{\to}{v}^{zy/a}}.
\end{aligned}
$$

Remember that the Transport Theorem is applicable to any physical vector. Using the Transport Theorem on $\underrightarrow{v}^{zw/a}$ the *the acceleration of point z relative to point w w.r.t. $\mathcal{F}_a$* is

$$\begin{aligned}
\underrightarrow{v}^{zw/a\,\bullet a} &= \underrightarrow{v}^{yw/a\,\bullet a} + \underrightarrow{v}^{zy/a\,\bullet a} \\
&= \underrightarrow{v}^{yw/a\,\bullet a} + \underrightarrow{v}^{zy/a\,\bullet b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zy/a} \\
&= \underrightarrow{v}^{yw/a\,\bullet a} + \left(\underrightarrow{v}^{zy/b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zy}\right)^{\bullet b} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{v}^{zy/b} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zy}\right) \\
&= \underrightarrow{v}^{yw/a\,\bullet a} + \underrightarrow{v}^{zy/b\,\bullet b} + \underrightarrow{\omega}^{ba\,\bullet b} \times \underrightarrow{r}^{zy} + 2\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zy/b} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zy}\right) \\
&= \underrightarrow{a}^{yw/a} + \underrightarrow{a}^{zy/b} + 2\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zy/b} + \underrightarrow{\omega}^{ba\,\bullet b} \times \underrightarrow{r}^{zy} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zy}\right),
\end{aligned}$$

where

$$\begin{aligned}
\underrightarrow{a}^{zw/a/a} &= \underrightarrow{v}^{zw/a\,\bullet a} = \underrightarrow{r}^{zw\,\bullet a\bullet a} = \text{acceleration of point } z \text{ relative to point } w \text{ w.r.t. } \mathcal{F}_a, \\
\underrightarrow{a}^{yw/a/a} &= \underrightarrow{v}^{yw/a\,\bullet a} = \underrightarrow{r}^{yw\,\bullet a\bullet a} = \text{acceleration of point } y \text{ relative to point } w \text{ w.r.t. } \mathcal{F}_a, \\
2\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zy/b} &= \text{the Coriolis acceleration}, \\
\underrightarrow{\omega}^{ba\,\bullet b} \times \underrightarrow{r}^{zy} &= \text{the angular acceleration}, \\
\underrightarrow{\omega}^{ba} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zy}\right) &= \text{centripetal acceleration}.
\end{aligned}$$

The word *centripetal* means "center seeking", while the word *centrifugal* means "center fleeing".

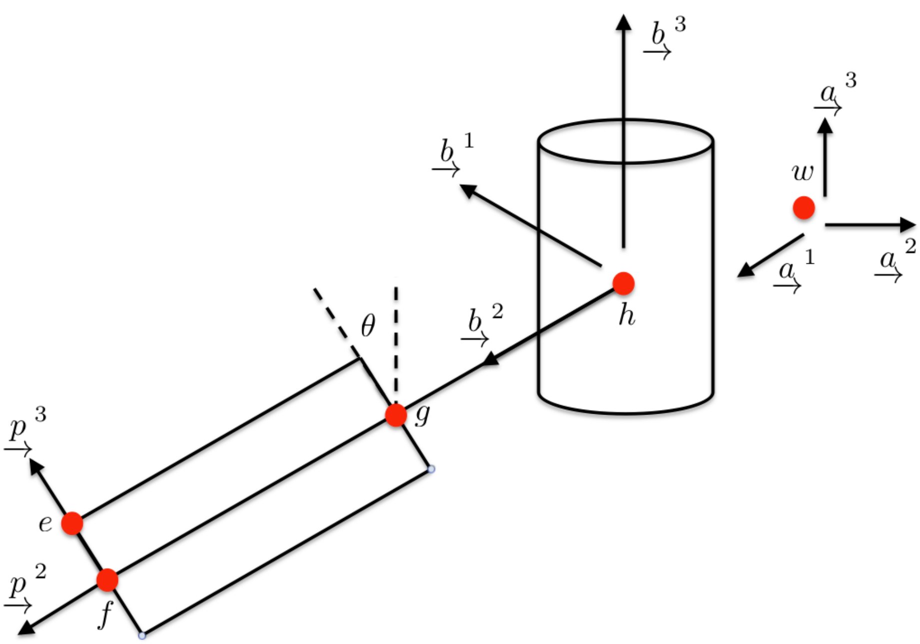


Figure 3.4: A spacecraft with a solar panel. Frame $\mathcal{F}_b$ is attached to the spacecraft bus (the "hub"), while $\mathcal{F}_p$ is at the tip of the solar panel.

Example 3.1. [37, pp. 491-492] Consider Figure 3.4. The spacecraft bus rotates about the $\underrightarrow{b}^3$ axis at a rate $\dot{\gamma}$, while the solar panel rotates about the $\underrightarrow{b}^2$ axis at a constant rate $\dot{\theta}$. Find $\underrightarrow{v}^{ew/a}$ and $\underrightarrow{a}^{ew/a/a}$, which is to say, find the velocity of point e relative to point w w.r.t. $\mathcal{F}_a$ and the acceleration of point e relative to point w w.r.t. $\mathcal{F}_a$. Assume $\underrightarrow{r}^{hw}$ is constant.

Solution. First, find the angular velocity of $\mathcal{F}_p$ relative to $\mathcal{F}_a$, $\underrightarrow{\omega}^{pa}$. The relationship between $\mathcal{F}_a$, $\mathcal{F}_b$, and $\mathcal{F}_p$ is

$$\mathcal{F}_a \text{ — (Rot. about } \underrightarrow{a}^3 \text{ or } \underrightarrow{b}^3) \rightarrow \mathcal{F}_b \text{ — (Rot. about } \underrightarrow{b}^2 \text{ or } \underrightarrow{p}^2) \rightarrow \mathcal{F}_p,$$

which in terms of $\mathbf{C}_{ba} = \mathbf{C}_3(\gamma)$ and $\mathbf{C}_{pb} = \mathbf{C}_2(\theta)$. It follows that

$$\mathbf{C}_{pa} = \mathbf{C}_{pb}\mathbf{C}_{ba} = \mathbf{C}_2(\theta)\mathbf{C}_3(\gamma).$$

The angular velocity $\underrightarrow{\omega}^{pa}$ is then

$$\begin{aligned}
\underrightarrow{\omega}^{pa} &= \underrightarrow{\omega}^{pb} + \underrightarrow{\omega}^{ba} \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \mathbf{1}_2 \dot{\theta} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \mathbf{1}_3 \dot{\gamma} \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \left(\begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \underbrace{\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}}_{\mathbf{C}_{pb}=\mathbf{C}_2(\theta)} \begin{bmatrix} 0 \\ 0 \\ \dot{\gamma} \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} -\sin\theta\dot{\gamma} \\ \dot{\theta} \\ \cos\theta\dot{\gamma} \end{bmatrix}.
\end{aligned}$$

Next, the position of point e relative to point w is

$$\underrightarrow{r}^{ew} = \underrightarrow{r}^{ef} + \underrightarrow{r}^{fg} + \underrightarrow{r}^{gh} + \underrightarrow{r}^{hw}.$$

In particular, letting

$$\begin{aligned}
\underrightarrow{r}^{ef} &= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} 0 \\ 0 \\ \ell^{ef} \end{bmatrix}, \\
\underrightarrow{r}^{fh} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} 0 \\ \ell^{fh} \\ 0 \end{bmatrix} = \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} 0 \\ \ell^{fh} \\ 0 \end{bmatrix},
\end{aligned}$$

it follows that $\underrightarrow{r}^{eh}$ is

$$\underrightarrow{r}^{eh} = \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} 0 \\ \ell^{fh} \\ \ell^{ef} \end{bmatrix},$$

and

$$\underrightarrow{r}^{ew} = \underrightarrow{r}^{eh} + \underrightarrow{r}^{hw}.$$

Next, taking the time-derivative w.r.t. $\mathcal{F}_a$ of the expression for $\underrightarrow{r}^{ew}$ and using the Transport Theorem yields

$$\begin{aligned}
{\underrightarrow{r}^{ew}}^{\bullet a} &= {\underrightarrow{r}^{eh}}^{\bullet a} + \cancelto{\underrightarrow{0}}{{\underrightarrow{r}^{hw}}^{\bullet a}} \\
&= \cancelto{\underrightarrow{0}}{{\underrightarrow{r}^{eh}}^{\bullet p}} + \underrightarrow{\omega}^{pa} \times \underrightarrow{r}^{eh} \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} 0 & -\cos\theta\dot{\gamma} & \dot{\theta} \\ \cos\theta\dot{\gamma} & 0 & \sin\theta\dot{\gamma} \\ -\dot{\theta} & -\sin\theta\dot{\gamma} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \ell^{fh} \\ \ell^{ef} \end{bmatrix} \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \begin{bmatrix} \dot{\theta}\ell^{ef} - \cos\theta\dot{\gamma}\ell^{fh} \\ \sin\theta\dot{\gamma}\ell^{ef} \\ -\sin\theta\dot{\gamma}\ell^{fh} \end{bmatrix},
\end{aligned}$$

which is the velocity of point e relative to point w w.r.t. $\mathcal{F}_a$.

To find ${\underrightarrow{v}^{ew/a}}^{\bullet a}$, where $\underrightarrow{v}^{ew/a} = {\underrightarrow{r}^{ew}}^{\bullet a}$, apply the Transport Theorem once more, resulting in

$$\begin{aligned}
{\underrightarrow{v}^{ew/a}}^{\bullet a} &= {\underrightarrow{v}^{ew/a}}^{\bullet p} + \underrightarrow{\omega}^{pa} \times \underrightarrow{v}^{ew/a} \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \left(\begin{bmatrix} \ell^{ef}\ddot{\theta} + \dot{\theta}\sin\theta\dot{\gamma}\ell^{fh} \\ \dot{\theta}\cos\theta\dot{\gamma}\ell^{ef} \\ -\dot{\theta}\cos\theta\dot{\gamma}\ell^{fh} \end{bmatrix} + \begin{bmatrix} 0 & -\cos\theta\dot{\gamma} & \dot{\theta} \\ \cos\theta\dot{\gamma} & 0 & \sin\theta\dot{\gamma} \\ -\dot{\theta} & -\sin\theta\dot{\gamma} & 0 \end{bmatrix} \begin{bmatrix} \ell^{ef}\dot{\theta} - \cos\theta\dot{\gamma}\ell^{fh} \\ \sin\theta\dot{\gamma}\ell^{ef} \\ -\sin\theta\dot{\gamma}\ell^{fh} \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_p^{\mathsf{T}} \left(\begin{bmatrix} \ell^{ef}(\ddot{\theta} - \cos\theta\sin\theta\dot{\gamma}^2) \\ 2\dot{\theta}\dot{\gamma}\cos\theta\ell^{ef} - \dot{\gamma}^2\ell^{fh} \\ -\ell^{ef}(\dot{\theta}^2 + \sin^2\theta\dot{\gamma}^2) \end{bmatrix} \right).
\end{aligned}$$

□

3.6 Coordinate Systems

Recall that

$$\underrightarrow{r}^{zw} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{r}_a^{zw} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{r}_b^{zw}.$$

Often the components $\mathbf{r}_a^{zw}$ or $\mathbf{r}_b^{zw}$ are parameterized using a *coordinate system*. Cartesian, cylindrical, spherical, and normal-tangential (a.k.a., Serret-Frenet or path-rate) coordinates will now be discussed. The presentation herein follows that of [2].

It should be clear that any one coordinate system can be used to parameterize $\mathbf{r}_a^{zw}$ or $\mathbf{r}_b^{zw}$, and a coordinate system is *not* a reference frame. In some books and papers a reference frame is referred to as a coordinate system, which is not a correct use of the term coordinate system.

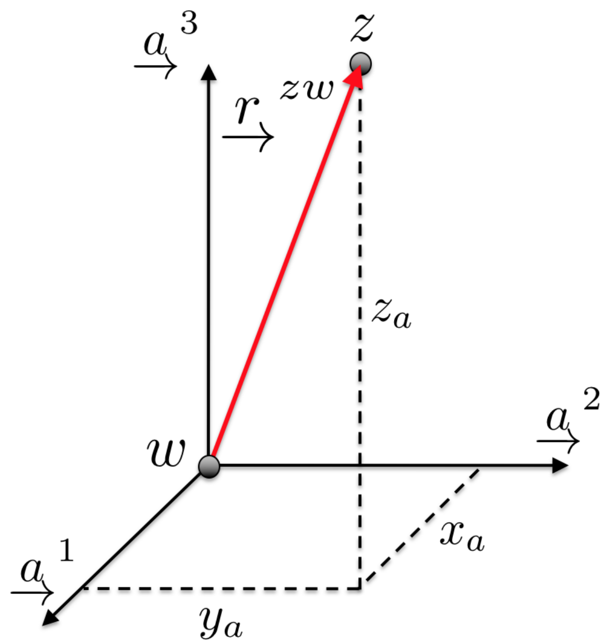


Figure 3.5: Cartesian coordinate system.

3.6.1 Cartesian Coordinates

Position

Starting from Figure 3.5, the physical vector $\underrightarrow{r}^{zw}$ can be written as

$$\underrightarrow{r}^{zw} = \underrightarrow{a}^1 x_a + \underrightarrow{a}^2 y_a + \underrightarrow{a}^3 z_a = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix},$$

where

$$\mathbf{r}_a^{zw} = \begin{bmatrix} r_{a1}^{zw} \\ r_{a2}^{zw} \\ r_{a3}^{zw} \end{bmatrix} = \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix}.$$

Together x_a, y_a, and z_a make up a cartesian coordinate system.

Velocity

The velocity of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$\underrightarrow{r}^{zw^{\bullet a}} = \underrightarrow{v}^{zw/a} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix}.$$

Acceleration

The acceleration of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$\underrightarrow{v}^{zw/a^{\bullet a}} = \underrightarrow{a}^{zw/a/a} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix}.$$

3.6.2 Cylindrical Coordinates

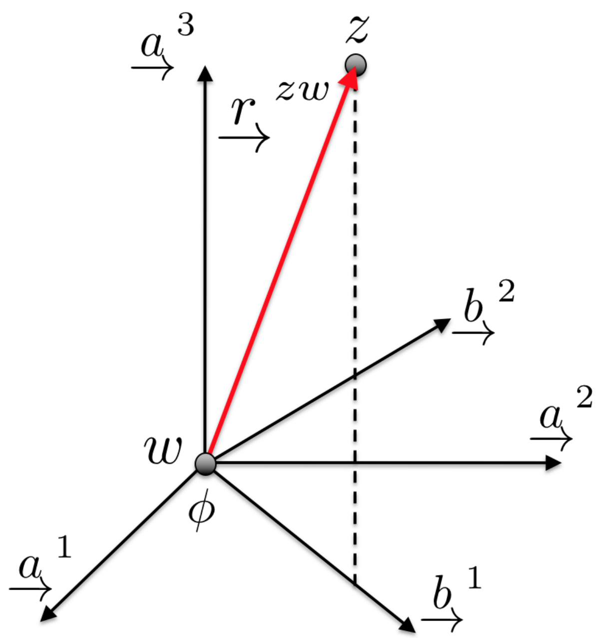


Figure 3.6: Cylindrical coordinate system.

Position

As shown in Figure 3.6, to go from $\mathcal{F}_a$ to $\mathcal{F}_b$, the frame $\mathcal{F}_b$ is rotated about the $\underrightarrow{a}^1 = \underrightarrow{b}^1$ axis through the angle ϕ. As such, $\mathbf{C}_{ba} = \mathbf{C}_3(\phi)$. Then, in terms of either $\mathcal{F}_a$ or $\mathcal{F}_b$, the physical vector $\underrightarrow{r}^{zw}$ can be written as

$$\underrightarrow{r}^{zw} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} \rho_b \\ 0 \\ z_b \end{bmatrix},$$

where $\rho_b = r_{b1}^{zw}$, $z_a = z_b$, and

$$\mathbf{r}_b^{zw} = \begin{bmatrix} \rho_b \\ 0 \\ z_b \end{bmatrix}.$$

Together ρ, z_b, and ϕ compose a cylindrical coordinate system.

Velocity

To compute the velocity of point z relative to point w w.r.t. $\mathcal{F}_a$, the Transport Theorem will be employed. As such, first $\underrightarrow{\omega}^{ba}$ will be computed. Recall that $\mathbf{C}_{ba} = \mathbf{C}_3(\phi)$, and therefore

$$\boldsymbol{\omega}_b^{ba} = \boldsymbol{\omega}_a^{ba} = \begin{bmatrix} 0 \\ 0 \\ \dot{\phi} \end{bmatrix}.$$

Then, using the Transport Theorem, the velocity of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$\begin{aligned}
\underrightarrow{r}^{zw^{\bullet a}} = \underrightarrow{v}^{zw/a} &= \underrightarrow{r}^{zw^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zw}, \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{v}_a^{zw/a} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{v}_b^{zw/b} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{r}_b^{zw} \right), \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\begin{bmatrix} \dot{\rho}_b \\ 0 \\ \dot{z}_b \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dot{\phi} \end{bmatrix}^\times \begin{bmatrix} \rho_b \\ 0 \\ \dot{z}_b \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\begin{bmatrix} \dot{\rho}_b \\ 0 \\ \dot{z}_b \end{bmatrix} + \begin{bmatrix} 0 & -\dot{\phi} & 0 \\ \dot{\phi} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho_b \\ 0 \\ \dot{z}_b \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} \dot{\rho}_b \\ \rho_b \dot{\phi} \\ \dot{z}_b \end{bmatrix}.
\end{aligned}$$

It follows that

$$\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix}}_{\mathbf{v}_a^{zw/a}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \underbrace{\begin{bmatrix} \dot{\rho}_b \\ \rho_b \dot{\phi} \\ \dot{z}_b \end{bmatrix}}_{\mathbf{v}_b^{zw/a}},$$

which in referential form is

$$\mathbf{v}_a^{zw/a} = \mathbf{C}_{ba}^{\mathsf{T}} \mathbf{v}_b^{zw/a}.$$

Acceleration

Using the Transport Theorem again, the acceleration of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$\begin{aligned}
\underrightarrow{v}^{zw/a^{\bullet a}} = \underrightarrow{a}^{zw/a/a} &= \underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a}, \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{a}_a^{zw/a/a} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{a}_b^{zw/a/b} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} \right), \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\begin{bmatrix} \ddot{\rho}_b \\ \dot{\rho}_b \dot{\phi} + \rho_b \ddot{\phi} \\ \ddot{z}_b \end{bmatrix} + \begin{bmatrix} 0 & -\dot{\phi} & 0 \\ \dot{\phi} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\rho}_b \\ \rho_b \dot{\phi} \\ \dot{z}_b \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} \ddot{\rho}_b - \rho_b \dot{\phi}^2 \\ \rho_b \ddot{\phi} + 2 \dot{\rho}_b \dot{\phi} \\ \ddot{z}_b \end{bmatrix}.
\end{aligned}$$

Therefore,

$$\underset{\to}{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix}}_{\mathbf{a}_a^{zw/a/a}} = \underset{\to}{\mathcal{F}}_b^{\mathsf{T}} \underbrace{\begin{bmatrix} \ddot{\rho}_b - \rho_b \dot{\phi}^2 \\ \rho_b \ddot{\phi} + 2\dot{\rho}_b \dot{\phi} \\ \ddot{z}_b \end{bmatrix}}_{\mathbf{a}_b^{zw/a/a}},$$

which in referential form is

$$\mathbf{a}_a^{zw/a/a} = \mathbf{C}_{ba}^{\mathsf{T}} \mathbf{a}_b^{zw/a/a}.$$

3.6.3 Spherical Coordinates

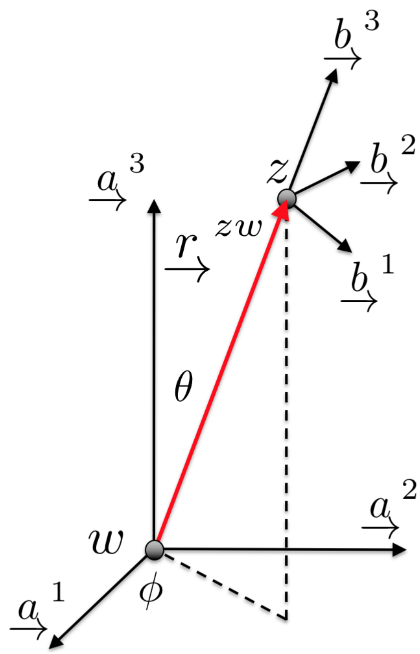


Figure 3.7: Spherical coordinate system.

Position

Spherical coordinates involves three frames, $\mathcal{F}_a$, $\mathcal{F}_\ell$, and $\mathcal{F}_b$, where $\mathcal{F}_\ell$ is an intermediate frame. To go from $\mathcal{F}_a$ to $\mathcal{F}_\ell$, rotate about the $\underset{\to}{a}^3 = \underset{\to}{\ell}^3$ axis by ϕ, and hence $\mathbf{C}_{\ell a} = \mathbf{C}_3(\phi)$. To go from $\mathcal{F}_\ell$ to $\mathcal{F}_b$, rotate about the $\underset{\to}{\ell}^2 = \underset{\to}{b}^2$ axis by θ, meaning that $\mathbf{C}_{b\ell} = \mathbf{C}_2(\theta)$. Thus,

$$\mathbf{C}_{ba} = \mathbf{C}_{b\ell}\mathbf{C}_{\ell a} = \mathbf{C}_2(\theta)\mathbf{C}_3(\phi).$$

Then, the physical vector $\underset{\to}{r}^{zw}$ can be written as

$$\underset{\to}{r}^{zw} = \underset{\to}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} = \underset{\to}{\mathcal{F}}_b^{\mathsf{T}} \underbrace{\begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}}_{\mathbf{r}_b^{zw}},$$

where $r = \left\| \underrightarrow{r}^{zw} \right\|_2$. A cylindrical coordinate system is composed of r, θ, and ϕ.

Velocity

To compute the velocity of point z relative to point w w.r.t. $\mathcal{F}_a$, the Transport Theorem will be employed, necessitating an expression for $\underrightarrow{\omega}^{ba}$ resolved in $\mathcal{F}_b$. Recall that the orientation of $\mathcal{F}_b$ relative to $\mathcal{F}_a$ is

$$\mathbf{C}_{ba} = \mathbf{C}_{b\ell}\mathbf{C}_{\ell a} = \mathbf{C}_2(\theta)\mathbf{C}_3(\phi).$$

Therefore,

$$\begin{aligned}
\underrightarrow{\omega}^{ba} &= \underrightarrow{\mathcal{F}}_\ell^{\mathsf{T}}\mathbf{1}_3\dot{\phi} + \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\mathbf{1}_2\dot{\theta}, \\
\boldsymbol{\omega}_b^{ba} &= \mathbf{C}_{b\ell}\mathbf{1}_3\dot{\phi} + \mathbf{1}_2\dot{\theta} \\
&= \begin{bmatrix} c_\theta & 0 & -s_\theta \\ 0 & 1 & 0 \\ s_\theta & 0 & c_\theta \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \dot{\phi} \end{bmatrix} + \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -s_\theta\dot{\phi} \\ \dot{\theta} \\ c_\theta\dot{\phi} \end{bmatrix}.
\end{aligned}$$

Using the Transport Theorem, the velocity of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$\begin{aligned}
\underrightarrow{r}^{zw^{\bullet a}} = \underrightarrow{v}^{zw/a} &= \underrightarrow{r}^{zw^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{zw}, \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\mathbf{v}_a^{zw/a} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\left(\mathbf{v}_b^{zw/b} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{r}_b^{zw}\right), \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\left(\begin{bmatrix} 0 \\ 0 \\ \dot{r} \end{bmatrix} + \begin{bmatrix} -s_\theta\dot{\phi} \\ \dot{\theta} \\ c_\theta\dot{\phi} \end{bmatrix}^\times \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}\right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\left(\begin{bmatrix} 0 \\ 0 \\ \dot{r} \end{bmatrix} + \begin{bmatrix} 0 & -c_\theta\dot{\phi} & \dot{\theta} \\ c_\theta\dot{\phi} & 0 & -(-s_\theta\dot{\phi}) \\ -\dot{\theta} & -s_\theta\dot{\phi} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}\right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\begin{bmatrix} r\dot{\theta} \\ r\dot{\phi}s_\theta \\ \dot{r} \end{bmatrix}.
\end{aligned}$$

It follows that

$$\underrightarrow{\mathcal{F}}_a^{\mathsf{T}}\underbrace{\begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix}}_{\mathbf{v}_a^{zw/a}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}}\underbrace{\begin{bmatrix} r\dot{\theta} \\ r\dot{\phi}s_\theta \\ \dot{r} \end{bmatrix}}_{\mathbf{v}_b^{zw/a}}.$$

In referential form,

$$\mathbf{v}_a^{zw/a} = \mathbf{C}_{ba}^{\mathsf{T}}\mathbf{v}_b^{zw/a}.$$

Acceleration

Using the Transport Theorem again, the acceleration of point z relative to point w w.r.t. $\mathcal{F}_a$ is

$$
\begin{aligned}
\underrightarrow{v}^{zw/a^{\bullet a}} = \underrightarrow{a}^{zw/a/a} &= \underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a}, \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \mathbf{a}_a^{zw/a/a} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\mathbf{a}_b^{zw/a/b} + \boldsymbol{\omega}_b^{ba^{\times}} \mathbf{v}_b^{zw/a} \right), \\
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \ddot{x}_a \\ \ddot{y}_a \\ \ddot{z}_a \end{bmatrix} &= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \left(\begin{bmatrix} \dot{r}\dot{\theta} + r\ddot{\theta} \\ \dot{r}\dot{\phi}s_\theta + r\ddot{\phi}s_\theta + r\dot{\phi}\dot{\theta}c_\theta \\ \ddot{r} \end{bmatrix} + \begin{bmatrix} -s_\theta\dot{\phi} \\ \dot{\theta} \\ c_\theta\dot{\phi} \end{bmatrix}^{\times} \begin{bmatrix} r\dot{\theta} \\ r\dot{\phi}s_\theta \\ \dot{r} \end{bmatrix} \right) \\
&= \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} 2\dot{r}\dot{\theta} + r\ddot{\theta} - r\dot{\phi}^2 s_\theta c_\theta \\ r\ddot{\phi}s_\theta + 2r\dot{\phi}\dot{\theta}c_\theta + 2\dot{r}\dot{\phi}s_\theta \\ \ddot{r} - r\dot{\theta}^2 - r\dot{\theta}^2 s_\theta^2 \end{bmatrix}.
\end{aligned}
$$

Therefore,

$$
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \underbrace{\begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix}}_{\mathbf{a}_a^{zw/a/a}} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \underbrace{\begin{bmatrix} 2\dot{r}\dot{\theta} + r\ddot{\theta} - r\dot{\phi}^2 s_\theta c_\theta \\ r\ddot{\phi}s_\theta + 2r\dot{\phi}\dot{\theta}c_\theta + 2\dot{r}\dot{\phi}s_\theta \\ \ddot{r} - r\dot{\theta}^2 - r\dot{\theta}^2 s_\theta^2 \end{bmatrix}}_{\mathbf{a}_b^{zw/a/a}}.
$$

In referential form,

$$
\mathbf{a}_a^{zw/a/a} = \mathbf{C}_{ba}^{\mathsf{T}} \mathbf{a}_b^{zw/a/a}.
$$

3.6.4 Normal-Tangential Coordinates

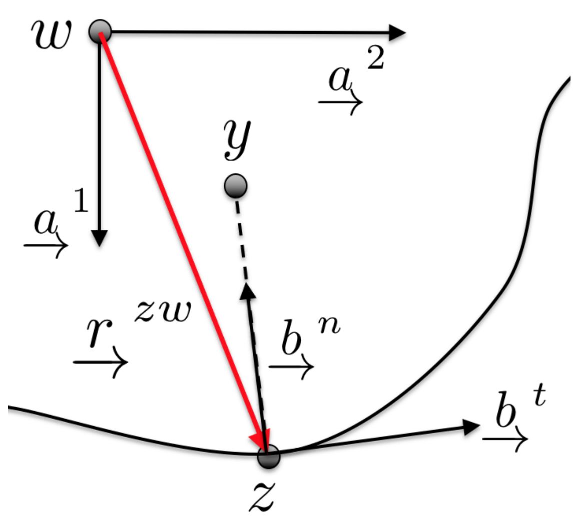


Figure 3.8: Normal-tangential coordinates.

Position

When the path of point z relative to point w is known, physical vectors normal and tangent to the path can be used to define a reference frame. As shown in Figure 3.8, the physical vector $\underset{\rightarrow}{b}^{t}$ is tangent to the path at point z, while the physical vector $\underset{\rightarrow}{b}^{t}$ is parallel to $\underset{\rightarrow}{r}^{yz}$ where y denotes the instantaneous center of curvature. The instantaneous radius of curvature magnitude is denoted $\rho = \left\| \underset{\rightarrow}{r}^{yz} \right\|_2$, and s is the distance along the path. Together ρ and s are the normal and tangential coordinates. With these definitions,

$$\underset{\rightarrow}{r}^{yz} = \rho\, \underset{\rightarrow}{b}^{n} = \underbrace{\begin{bmatrix} \underset{\rightarrow}{b}^{t} & \underset{\rightarrow}{b}^{n} & \underset{\rightarrow}{b}^{b} \end{bmatrix}}_{\underset{\rightarrow}{\mathcal{F}}_b^{\mathsf{T}}} \underbrace{\begin{bmatrix} 0 \\ \rho \\ 0 \end{bmatrix}}_{\mathbf{r}_b^{zw}},$$

where $\underset{\rightarrow}{b}^{b} = \underset{\rightarrow}{b}^{t} \times \underset{\rightarrow}{b}^{n}$ is the bi-normal physical vector.

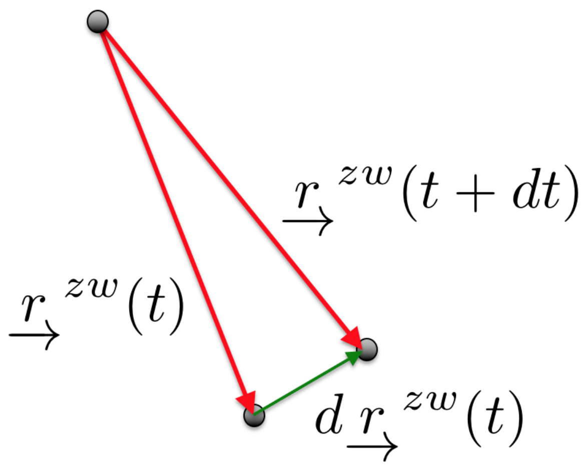


Figure 3.9: Differential associated with position vector.

Velocity

With regard to Figures 3.8 and 3.9, the tangent physical vector $\underset{\rightarrow}{b}^{t}$ is tangent to the path at point z. Moreover, as shown in Figure 3.9,

$$\underset{\rightarrow}{r}^{zw}(t + \mathrm{d}t) = \underset{\rightarrow}{r}^{zw}(t) + \mathrm{d}\,\underset{\rightarrow}{r}^{zw}(t),$$

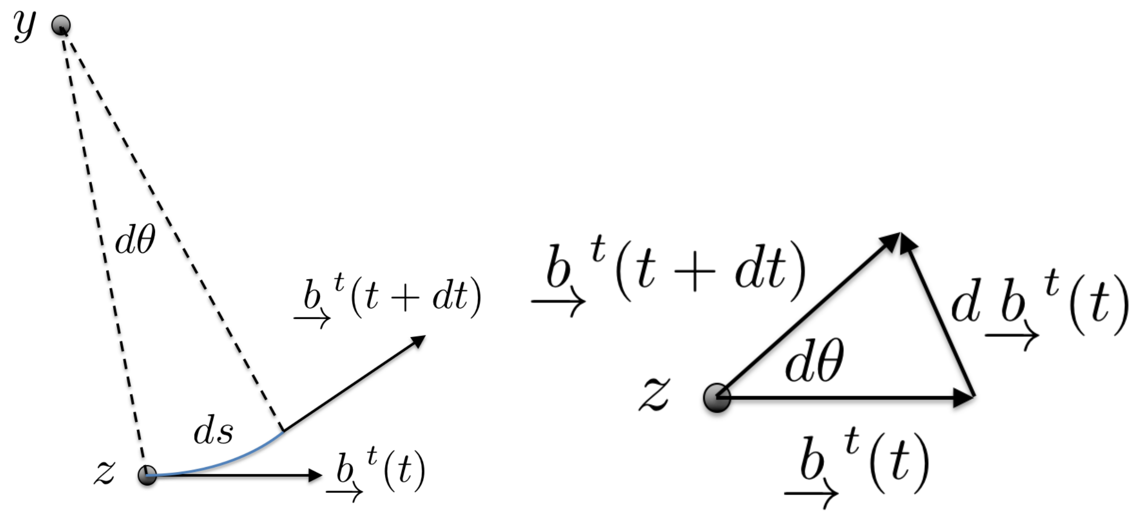
where $\underset{\rightarrow}{b}^{t}$ and $\mathrm{d}\,\underset{\rightarrow}{r}^{zw}(t)$ are parallel. It follows that

$$\mathrm{d}\,\underset{\rightarrow}{r}^{zw} \quad = \quad \mathrm{d}s\, \underset{\rightarrow}{b}^{t}.$$

Dividing by $\mathrm{d}t$ results in

$$
\begin{aligned}
\frac{\mathrm{d}\,\underrightarrow{r}^{zw}}{\mathrm{d}t} &= \frac{\mathrm{d}s}{\mathrm{d}t}\underrightarrow{b}^{t}, \\
{\underrightarrow{r}^{zw}}^{\bullet a} &= \dot{s}\,\underrightarrow{b}^{t}, = \underbrace{\begin{bmatrix} \underrightarrow{b}^{t} & \underrightarrow{b}^{n} & \underrightarrow{b}^{b} \end{bmatrix}}_{\underrightarrow{\mathcal{F}}_b^{\mathsf{T}}} \underbrace{\begin{bmatrix} \dot{s} \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{v}_b^{zw/a}},
\end{aligned}
$$

meaning the velocity of point z relative to point w w.r.t. $\mathcal{F}_a$ is tangent to the path.



(a) (b)

Figure 3.10: How to simply the acceleration expression.

Acceleration

The acceleration of point z relative to point w w.r.t. $\mathcal{F}_a$, ${\underrightarrow{v}^{zw/a}}^{\bullet a}$, is

$$
{\underrightarrow{v}^{zw/a}}^{\bullet a} = \ddot{s}\,\underrightarrow{b}^{t} + \dot{s}\,{\underrightarrow{b}^{t}}^{\bullet a}.
$$

To find an expression for ${\underrightarrow{b}^{t}}^{\bullet a}$, recall that (arc length) = (angle)(radius). Referring to Figures 3.10(a) and 3.10(b), it follows that

$$
\mathrm{d}\,\underrightarrow{b}^{t} = \mathrm{d}\theta(1)\,\underrightarrow{b}^{n},
$$

where the length of $\underrightarrow{b}^{t}$ is one, $\underrightarrow{b}^{t}(t)$ and $\underrightarrow{b}^{t}(t+\mathrm{d}t)$ have the same length, and $\mathrm{d}\,\underrightarrow{b}^{t}$ points in the same direction as $\underrightarrow{b}^{n}$. Recall that $\mathrm{d}s = \mathrm{d}\theta\rho$, where the radius of curvature is $\rho = \left\|\underrightarrow{r}^{yz}\right\|_2$. Using this relation with $\mathrm{d}\underset{\rightarrow}{b}^{t} = \mathrm{d}\theta\,\underset{\rightarrow}{b}^{n}$ gives

$$\begin{aligned}
\mathrm{d}s\,\underset{\rightarrow}{b}^{n} &= \rho\mathrm{d}\underset{\rightarrow}{b}^{t}, \\
\frac{\mathrm{d}s}{\mathrm{d}t}\underset{\rightarrow}{b}^{n} &= \rho\frac{\mathrm{d}\underset{\rightarrow}{b}^{t}}{\mathrm{d}t}, \\
\dot{s}\,\underset{\rightarrow}{b}^{n} &= \rho\,\underset{\rightarrow}{b}^{t^{\bullet a}}, \\
\underset{\rightarrow}{b}^{t^{\bullet a}} &= \frac{\dot{s}}{\rho}\underset{\rightarrow}{b}^{n}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\underset{\rightarrow}{v}^{zw/a^{\bullet a}} &= \ddot{s}\,\underset{\rightarrow}{b}^{t} + \dot{s}\,\underset{\rightarrow}{b}^{t^{\bullet a}} \\
&= \ddot{s}\,\underset{\rightarrow}{b}^{t} + \frac{\dot{s}^2}{\rho}\underset{\rightarrow}{b}^{n} \\
&= \left[\begin{array}{ccc} \underset{\rightarrow}{b}^{t} & \underset{\rightarrow}{b}^{n} & \underset{\rightarrow}{b}^{b} \end{array}\right] \left[\begin{array}{c} \ddot{s} \\ \dot{s}^2\rho^{-1} \\ 0 \end{array}\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\underset{\rightarrow}{a}^{zw/a/a} = \underset{\rightarrow}{v}^{zw/a^{\bullet a}} &= \underbrace{\left[\begin{array}{ccc} \underset{\rightarrow}{b}^{t} & \underset{\rightarrow}{b}^{n} & \underset{\rightarrow}{b}^{b} \end{array}\right]}_{\underset{\rightarrow}{\mathcal{F}}_b^{\mathsf{T}}} \left[\begin{array}{c} \ddot{s} \\ \dot{s}^2\rho^{-1} \\ 0 \end{array}\right] \\
&= \underset{\rightarrow}{\mathcal{F}}_b^{\mathsf{T}} \left[\begin{array}{c} a_{b\,t}^{zw/a/a} \\ a_{b\,n}^{zw/a/a} \\ 0 \end{array}\right].
\end{aligned}$$

The triad $\underset{\rightarrow}{b}^{t}$, $\underset{\rightarrow}{b}^{n}$, $\underset{\rightarrow}{b}^{b}$ is shown in Figure 3.11. The plane normal to $\underset{\rightarrow}{b}^{b}$ is referred to as the osculating plane.

Some special cases are worth discussing. To begin, if the path of point z approaches a straight line, $\rho \to \infty$. In this case,

$$\underset{\rightarrow}{a}^{zw/a/a} = \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{\mathcal{F}}_b^{\mathsf{T}} \left[\begin{array}{c} \ddot{s} \\ 0 \\ 0 \end{array}\right].$$

Next, if the point z moves at a constant speed along the path, meaning $\ddot{s} = 0$, then

$$\underset{\rightarrow}{a}^{zw/a/a} = \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{\mathcal{F}}_b^{\mathsf{T}} \left[\begin{array}{c} 0 \\ \dot{s}^2\rho^{-1} \\ 0 \end{array}\right] = \underset{\rightarrow}{b}^{t}\frac{\dot{s}^2}{\rho}.$$

The normal component of the acceleration is toward y; this is a centripetal (center seeking) acceleration.

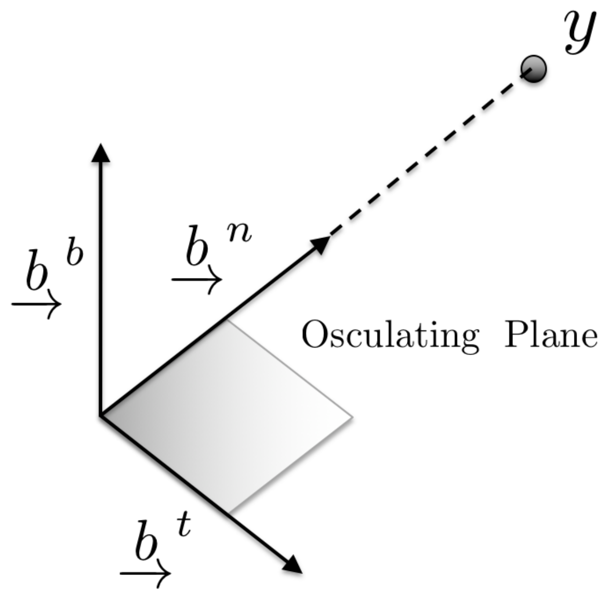


Figure 3.11: The triad associated with normal-tangential coordinates.

3.7 Rolling Without Slip

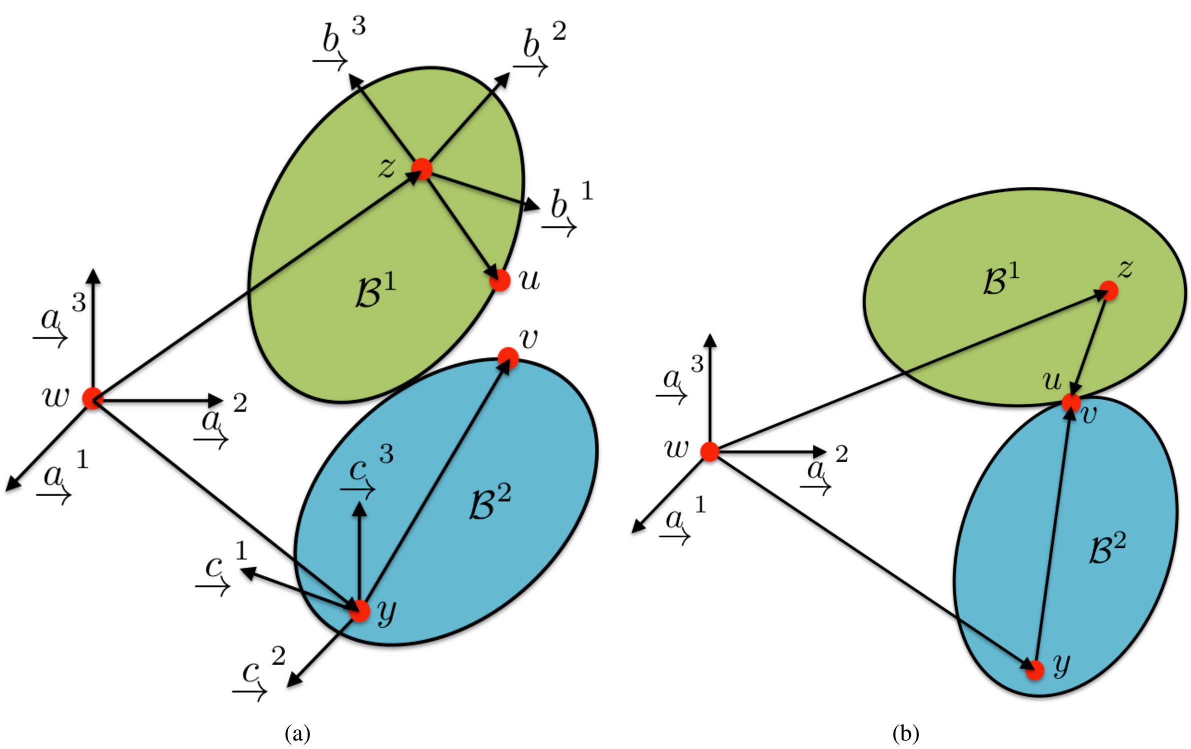


Figure 3.12: Two bodies rolling without slip.

This section is based on [4, pp. 107-110] and [3]. The definition of *rolling without slipping* is first given.

Definition 3.3. [3] Consider two rigid bodies[1], $\mathcal{B}^1$ and $\mathcal{B}^2$, and a reference frame $\mathcal{F}_a$, as shown in Figure 3.12(a). Let u and v be points fixed on the boundary of $\mathcal{B}^1$ and $\mathcal{B}^2$, respectively. Assume that $\mathcal{B}^1$ and $\mathcal{B}^2$ are in contact, and that points u and v are collocated, as shown in Figure 3.12(b). Then, $\mathcal{B}^1$ and $\mathcal{B}^2$ are *rolling without slipping* if $\underrightarrow{v}^{uv/a} = \underrightarrow{0}$. Otherwise, $\mathcal{B}^1$ and $\mathcal{B}^2$ are *slipping*.

An expression for $\mathcal{B}^1$ rolling without slip on $\mathcal{B}^2$ when $\mathcal{B}^2$ is stationary will now be derived. By stationary we mean that, referring to Figure 3.12, $\underrightarrow{r}^{yw^{\bullet a}} = \underrightarrow{0}$ and $\underrightarrow{\omega}^{ca} = \underrightarrow{0}$. However, to derive an expression for $\mathcal{B}^1$ rolling without slip on $\mathcal{B}^2$, first assume that $\mathcal{B}^2$ is permitted to translate and rotate.

Consider Figure 3.12(a). The position of point u relative to point w, and the velocity of point u relative to point w w.r.t. $\mathcal{F}_a$, are

$$\begin{aligned}
\underrightarrow{r}^{uw} &= \underrightarrow{r}^{uz} + \underrightarrow{r}^{zw}, \\
\underrightarrow{r}^{uw^{\bullet a}} &= \underrightarrow{r}^{uz^{\bullet a}} + \underrightarrow{r}^{zw^{\bullet a}} \\
&= \underrightarrow{r}^{zw^{\bullet a}} + \cancelto{\underrightarrow{0}}{\underrightarrow{r}^{uz^{\bullet b}}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{uz} \\
&= \underrightarrow{r}^{zw^{\bullet a}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{uz},
\end{aligned}$$

where $\underrightarrow{r}^{uz^{\bullet b}} = \underrightarrow{0}$ because $\mathcal{B}^1$ is a rigid body. Similarly, the position of point v relative to point w, and the velocity of point v relative to point w w.r.t. $\mathcal{F}_a$, are

$$\begin{aligned}
\underrightarrow{r}^{vw} &= \underrightarrow{r}^{vy} + \underrightarrow{r}^{yw}, \\
\underrightarrow{r}^{vw^{\bullet a}} &= \underrightarrow{r}^{vy^{\bullet a}} + \underrightarrow{r}^{yw^{\bullet a}} \\
&= \underrightarrow{r}^{yw^{\bullet a}} + \cancelto{\underrightarrow{0}}{\underrightarrow{r}^{vy^{\bullet c}}} + \underrightarrow{\omega}^{ca} \times \underrightarrow{r}^{vy} \\
&= \underrightarrow{r}^{yw^{\bullet a}} + \underrightarrow{\omega}^{ca} \times \underrightarrow{r}^{vy}.
\end{aligned}$$

where $\underrightarrow{r}^{vy^{\bullet c}} = \underrightarrow{0}$ because $\mathcal{B}^2$ is a rigid body.

Next, consider when points u and v are collocated, as shown in Figure 3.12(b). When points u and v are collocated they have the same velocity w.r.t. $\mathcal{F}_a$, that is $\underrightarrow{r}^{uw^{\bullet a}} = \underrightarrow{r}^{vw^{\bullet a}}$. If $\mathcal{B}^2$ is stationary, meaning $\underrightarrow{r}^{yw^{\bullet a}} = \underrightarrow{0}$ and $\underrightarrow{\omega}^{ca} = \underrightarrow{0}$, it follows that when points u and v are collocated

$$\begin{aligned}
\underbrace{\underrightarrow{r}^{uw^{\bullet a}}}_{\underrightarrow{0}} &= \underrightarrow{r}^{zw^{\bullet a}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{uz}, \\
\underrightarrow{0} &= \underrightarrow{r}^{zw^{\bullet a}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{uz}, \\
\underrightarrow{r}^{zw^{\bullet a}} &= \underrightarrow{r}^{uz} \times \underrightarrow{\omega}^{ba}. \qquad (3.30)
\end{aligned}$$

[1] Yes, the definition of a rigid body has yet to be defined. A rigid body is a body where every point in the body is fixed relative to all other points in the body.

Equation (3.30) describes $\mathcal{B}^1$ rolling without slip on $\mathcal{B}^2$.

Notice that the previous derivation leading to Equation (3.30) can be repeated for any points u and v on the boundary of $\mathcal{B}^1$ and $\mathcal{B}^2$. As such, points u and v are not special in that the derivation leading to Equation (3.30) holds for any general collocated points u on $\mathcal{B}^1$ and v on $\mathcal{B}^2$.

3.7.0.1 Example — Planar Rolling of a Coin

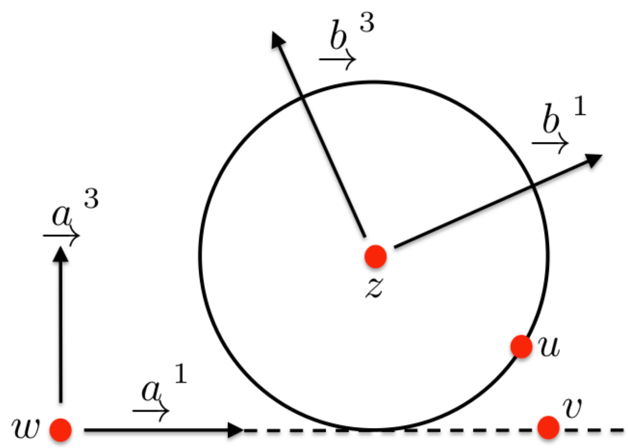


Figure 3.13: A coin or cylinder rolling in the $\underrightarrow{a}^1 - \underrightarrow{a}^3$ plane about the $\underrightarrow{a}^2$ physical basis vector.

Consider Figure 3.13 depicting a coin of radius r rolling without slip in the $\underrightarrow{a}^1 - \underrightarrow{a}^3$ plane. Fixed to the coin is frame $\mathcal{F}_b$ where the $\underrightarrow{b}^2$ axis is parallel to the $\underrightarrow{a}^2$ axis. Therefore, $\mathbf{C}_{ba} = \mathbf{C}_2(\theta)$ and

$$\underrightarrow{\omega}^{ba} = \underrightarrow{\mathcal{F}}_b^{\mathsf{T}} \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix}.$$

Let

$$\underrightarrow{r}^{zw} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} r_{a1}^{zw} \\ 0 \\ r_{a3}^{zw} \end{bmatrix} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ 0 \\ r \end{bmatrix}.$$

A relationship between $\underrightarrow{r}^{zw^{\bullet a}}$ and $\underrightarrow{\omega}^{ba}$ is desired. To begin, the position of point u relative to point w is given by

$$\underrightarrow{r}^{uw} = \underrightarrow{r}^{uz} + \underrightarrow{r}^{zw} = \underrightarrow{r}^{zw} + \underrightarrow{r}^{uz}.$$

Using the Transport Theorem the velocity of point u relative to point w w.r.t. $\mathcal{F}_a$ is

$$\begin{aligned} \underrightarrow{r}^{uw^{\bullet a}} &= \underrightarrow{r}^{zw^{\bullet a}} + \underrightarrow{r}^{uz^{\bullet a}} \\ &= \underrightarrow{r}^{zw^{\bullet a}} + \cancelto{\underrightarrow{0}}{\underrightarrow{r}^{uz^{\bullet b}}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{uz}, \end{aligned}$$

where $\underrightarrow{r}^{uz^{\bullet b}} = \underrightarrow{0}$ owing to the fact that the point u is fixed to the circumference of the coin and the coin is a rigid body. Next, when points u and v collocate $\underrightarrow{r}^{uw^{\bullet a}} = \underrightarrow{0}$ as a result of the fact that $\underrightarrow{r}^{vw^{\bullet a}} = \underrightarrow{0}$,

$\underset{\rightarrow}{r}^{uw^{\bullet a}} = \underset{\rightarrow}{0}$. Therefore,

$$\underset{\rightarrow}{0} = \underset{\rightarrow}{r}^{zw^{\bullet a}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{uz}.$$

In particular, when point u collocates with point v

$$\underset{\rightarrow}{r}^{uz} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} 0 \\ 0 \\ -r \end{bmatrix}.$$

Thus,

$$\begin{aligned} \underset{\rightarrow}{0} &= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \dot{x}_a \\ 0 \\ 0 \end{bmatrix} + \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} 0 & 0 & \dot{\theta} \\ 0 & 0 & 0 \\ -\dot{\theta} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -r \end{bmatrix} \\ &= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} \dot{x}_a \\ 0 \\ 0 \end{bmatrix} + \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} -r\dot{\theta} \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Therefore,

$$\dot{x}_a = r\dot{\theta}.$$

3.7.0.2 Example — General Rolling of a Coin

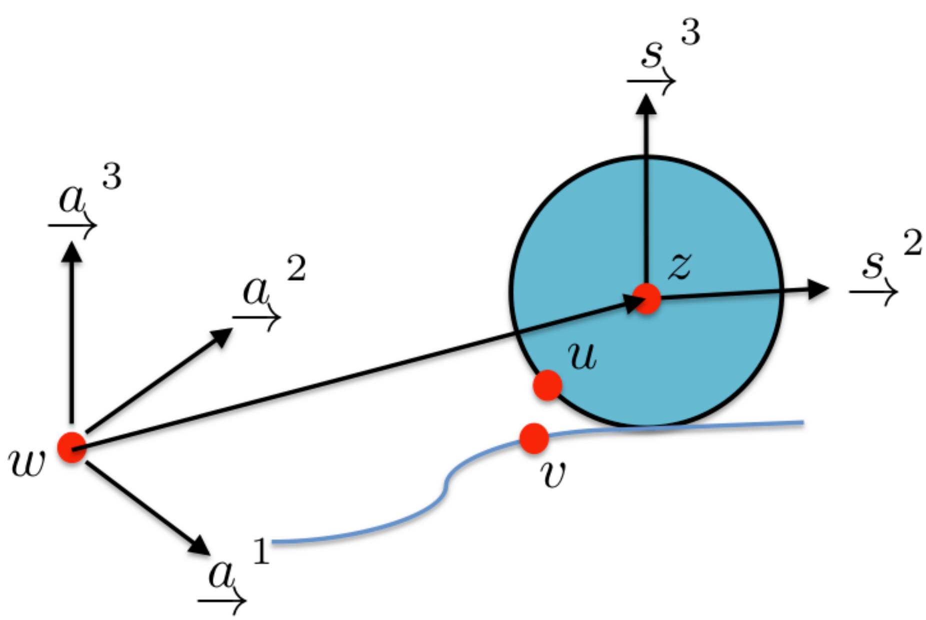


Figure 3.14: A coin rolling. The $\underset{\rightarrow}{s}^1$ physical basis vector is not parallel to the $\underset{\rightarrow}{a}^1 - \underset{\rightarrow}{a}^2$ plane, while the $\underset{\rightarrow}{s}^2$ physical basis vector does. The $\underset{\rightarrow}{s}^3$ physical basis vector does not remain parallel to the $\underset{\rightarrow}{a}^3$ physical basis vector.

This section is based on [2] and [19, pp. 162-164].

Consider Figure 3.14 depicting a coin rolling. The coin does not remain perfectly upright, but can "twist" and "lean" as it rolls. In particular, let $\mathcal{F}_a$ be a frame, let $\mathcal{F}_p$ be a fame with basis vectors fixed to the coin, and let $\mathcal{F}_s$ be a frame where the $\underrightarrow{s}^2$ axis is always parallel to the $\underrightarrow{a}^1 - \underrightarrow{a}^2$ plane. The attitude of $\mathcal{F}_p$ relative to $\mathcal{F}_a$ can be parameterized using an Euler angle sequence as

$$\mathbf{C}_{pa} = \mathbf{C}_{ps}\mathbf{C}_{sq}\mathbf{C}_{qa} = \mathbf{C}_1(\phi)\mathbf{C}_2(\theta)\mathbf{C}_3(\psi).$$

The angle ψ captures the "twist" of the coin about the $\underrightarrow{a}^3 = \underrightarrow{q}^3$ axis, the angle θ captures the "lean" of the coin about the $\underrightarrow{q}^2 = \underrightarrow{s}^2$ axis, and the angle ϕ captures the "rolling" of coin about the $\underrightarrow{s}^1 = \underrightarrow{p}^1$ axis. The angular velocity of $\mathcal{F}_p$ relative to $\mathcal{F}_a$ is then

$$\begin{aligned}
\underrightarrow{\omega}^{pa} &= \underrightarrow{\mathcal{F}}_q^{\mathsf{T}}\mathbf{1}_3\dot{\psi} + \underrightarrow{\mathcal{F}}_s^{\mathsf{T}}\mathbf{1}_2\dot{\theta} + \underrightarrow{\mathcal{F}}_p^{\mathsf{T}}\mathbf{1}_1\dot{\phi} \\
&= \underrightarrow{\mathcal{F}}_q^{\mathsf{T}}\left(\mathbf{1}_1\dot{\phi} + \mathbf{C}_{ps}\mathbf{1}_2\dot{\theta} + \mathbf{C}_{ps}\mathbf{C}_{sq}\mathbf{1}_3\dot{\psi}\right) \\
&= \underrightarrow{\mathcal{F}}_q^{\mathsf{T}}\left(\mathbf{1}_1\dot{\phi} + \mathbf{C}_1(\phi)\mathbf{1}_2\dot{\theta} + \mathbf{C}_1(\phi)\mathbf{C}_2(\theta)\mathbf{1}_3\dot{\psi}\right).
\end{aligned}$$

Noting that

$$\mathbf{C}_1(\phi)\mathbf{1}_2 = \begin{bmatrix} 0 \\ c_\phi \\ -s_\phi \end{bmatrix} \quad \text{and} \quad \mathbf{C}_1(\phi)\mathbf{C}_2(\theta)\mathbf{1}_3 = \begin{bmatrix} -s_\theta \\ s_\phi c_\theta \\ c_\phi c_\theta \end{bmatrix}$$

it follows that the angular velocity of $\mathcal{F}_p$ relative to $\mathcal{F}_a$ resolved in $\mathcal{F}_p$ is

$$\boldsymbol{\omega}_p^{pa} = \begin{bmatrix} -s_\theta & 0 & 1 \\ s_\phi c_\theta & c_\phi & 0 \\ c_\phi c_\theta & -s_\phi & 0 \end{bmatrix} \begin{bmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{bmatrix}.$$

Momentarily, it will prove useful to resolve $\underrightarrow{\omega}^{ba}$ in $\mathcal{F}_s$. To this end,

$$\begin{aligned}
\boldsymbol{\omega}_s^{pa} &= \mathbf{C}_{sp}\boldsymbol{\omega}_p^{pa} \\
&= \mathbf{C}_1^{\mathsf{T}}(\phi)\left(\mathbf{1}_1\dot{\phi} + \mathbf{C}_1(\phi)\mathbf{1}_2\dot{\theta} + \mathbf{C}_1(\phi)\mathbf{C}_2(\theta)\mathbf{1}_3\dot{\psi}\right) \\
&= \mathbf{1}_1\dot{\phi} + \mathbf{1}_2\dot{\theta} + \mathbf{C}_2(\theta)\mathbf{1}_3\dot{\psi} \\
&= \begin{bmatrix} \dot{\phi} - s_\theta\dot{\psi} \\ \dot{\theta} \\ c_\theta\dot{\psi} \end{bmatrix}.
\end{aligned}$$

A relationship between $\underrightarrow{r}^{zw}$ and $\underrightarrow{\omega}^{pa}$ is desired. Let z be a point at the geometric center of the coin. The position of point u relative to point w is given by

$$\underrightarrow{r}^{uw} = \underrightarrow{r}^{uz} + \underrightarrow{r}^{zw} = \underrightarrow{r}^{zw} + \underrightarrow{r}^{uz}.$$

Using the Transport Theorem the velocity of point u relative to point w w.r.t. $\mathcal{F}_a$ is

$$\begin{aligned}
\underrightarrow{r}^{uw^{\bullet a}} &= \underrightarrow{r}^{zw^{\bullet a}} + \underrightarrow{r}^{uz^{\bullet a}} \\
&= \underrightarrow{r}^{zw^{\bullet a}} + \cancelto{\underrightarrow{0}}{\underrightarrow{r}^{uz^{\bullet p}}} + \underrightarrow{\omega}^{pa} \times \underrightarrow{r}^{uz},
\end{aligned}$$

where ${\underrightarrow{r}^{uz}}^{\bullet p} = \underrightarrow{0}$ owing to the fact that the point u is fixed to the circumference of the coin and the coin is a rigid body. Next, when points u and v collocate ${\underrightarrow{r}^{uw}}^{\bullet a} = \underrightarrow{0}$ because ${\underrightarrow{r}^{vw}}^{\bullet a} = \underrightarrow{0}$. Therefore,

$$\begin{aligned}
\underrightarrow{0} &= {\underrightarrow{r}^{zw}}^{\bullet a} + \underrightarrow{\omega}^{pa} \times \underrightarrow{r}^{uz}, \\
{\underrightarrow{r}^{zw}}^{\bullet a} &= \underrightarrow{r}^{uz} \times \underrightarrow{\omega}^{pa}.
\end{aligned}$$

In particular, when point u collocates with point v

$$\underrightarrow{r}^{uz} = \underrightarrow{\mathcal{F}}_s^{\mathsf{T}} \underbrace{\begin{bmatrix} 0 \\ 0 \\ -r \end{bmatrix}}_{\mathbf{r}_s^{uz}}.$$

Letting

$$\underrightarrow{r}^{zw} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} r_{a1}^{zw} \\ r_{a2}^{zw} \\ r_{a3}^{zw} \end{bmatrix} = \underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix},$$

results in

$$\begin{aligned}
\underrightarrow{\mathcal{F}}_a^{\mathsf{T}} \dot{\mathbf{r}}_a^{zw} &= \underrightarrow{\mathcal{F}}_s^{\mathsf{T}} \mathbf{r}_s^{uz\times} \boldsymbol{\omega}_s^{pa} \\
&= \underrightarrow{\mathcal{F}}_s^{\mathsf{T}} \begin{bmatrix} 0 & -(-r) & 0 \\ -r & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\phi} - s_\theta \dot{\psi} \\ \dot{\theta} \\ c_\theta \dot{\psi} \end{bmatrix} \\
&= \underrightarrow{\mathcal{F}}_s^{\mathsf{T}} \begin{bmatrix} r\dot{\theta} \\ -r\dot{\phi} + r s_\theta \dot{\psi} \\ 0 \end{bmatrix}.
\end{aligned}$$

In referential form the above is

$$\dot{\mathbf{r}}_a^{zw} = \mathbf{C}_{sa}^{\mathsf{T}} \begin{bmatrix} r\dot{\theta} \\ -r\dot{\phi} + r s_\theta \dot{\psi} \\ 0 \end{bmatrix}.$$

The direction cosine matrix $\mathbf{C}_{sa}$ is

$$\begin{aligned}
\mathbf{C}_{sa} &= \mathbf{C}_2(\theta)\mathbf{C}_3(\psi) \\
&= \begin{bmatrix} c_\theta & 0 & -s_\theta \\ 0 & 1 & 0 \\ s_\theta & 0 & c_\theta \end{bmatrix} \begin{bmatrix} c_\psi & s_\psi & 0 \\ -s_\psi & c_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} c_\theta c_\psi & c_\theta s_\psi & -s_\theta \\ -s_\psi & c_\psi & 0 \\ s_\theta c_\psi & s_\theta s_\psi & c_\theta \end{bmatrix}.
\end{aligned}$$

Therefore,

$$\begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix} = \begin{bmatrix} c_\theta c_\psi & -s_\psi & s_\theta c_\psi \\ c_\theta s_\psi & c_\psi & s_\theta s_\psi \\ -s_\theta & 0 & c_\theta \end{bmatrix} \begin{bmatrix} r\dot{\theta} \\ -r\dot{\phi} + r s_\theta \dot{\psi} \\ 0 \end{bmatrix} = \begin{bmatrix} c_\theta c_\psi r\dot{\theta} + s_\psi r\dot{\phi} - s_\psi s_\theta r\dot{\psi} \\ c_\theta s_\psi r\dot{\theta} - c_\psi r\dot{\phi} + c_\psi s_\theta r\dot{\psi} \\ -s_\theta r\dot{\theta} \end{bmatrix}.$$

This is the relationship between $\underset{\rightarrow}{r}^{zw^{\bullet a}}$ and $\underset{\rightarrow}{\omega}^{pa}$ as the coin rolls without slip on the $\underset{\rightarrow}{a}^1 - \underset{\rightarrow}{a}^2$ plane.

3.8 To Do

- Show that, given $\underset{\rightarrow}{r}^{zx} = \underset{\rightarrow}{r}^{zy} + \underset{\rightarrow}{r}^{yx}$ and two frames $\mathcal{F}_a$ and $\mathcal{F}_b$, that

$$\begin{aligned} \underset{\rightarrow}{r}^{zx^{\bullet a}} &= \underset{\rightarrow}{r}^{zy^{\bullet a}} + \underset{\rightarrow}{r}^{yx^{\bullet a}} \\ &= \underset{\rightarrow}{r}^{zy^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{zy} + \underset{\rightarrow}{r}^{yx^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{yx} \end{aligned}$$

  is equal to

$$\underset{\rightarrow}{r}^{zx^{\bullet a}} = \underset{\rightarrow}{r}^{zx^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{zx}.$$

Chapter 4

Newton-Euler Approach to the Dynamics of Particles and Rigid Bodies

Having discussed how to precisely describe the position of points, as well as how points move in time with respect to different frames of reference, the laws that govern the motion of masses and rigid bodies can be discussed.

4.1 The Principle of Inertia and Newton's First Law

Before stating Newton's three laws, Galileo's Principle of Inertia [2] will be discussed. Consider Figure 4.1(a). A ball starting from rest rolls down an inclined plane in a vacuum under the influence of gravity, across a flat frictionless surface, and then up another inclined plane until it stops. (Note, gravity, nor the notion of a frictionless surface, has yet to be defined. But, in order to follow Galileo's thought experiment, it is assumed that readers are familiar with "gravity" and the notion of a "frictionless surface".) When $\alpha = \beta_1$, the ball reach a height h on the second incline, where h is the height the ball initially started at on the fist incline. Now consider Figure 4.1(b) where the second incline is at an angle $\beta_2 > \alpha$. The ball will still go up the second incline to a height h, but this time, the ball must travel a farther distance up the incline until it stops. Similarly, if the angle of the second incline is increased further, the ball will travel father up the second incline than before, as shown in Figure 4.1(c). When the angle of the second incline reaches π the ball will continue rolling on the flat frictionless surface forever; there's no reason why it would either speed up or slow down. This though experiment leads to the Principle of Inertia.

Definition 4.1 (Principle of Inertia)**.** Every body continues in its state of rest or uniform motion unless it is compelled to change that state by a force impressed on it.

At this point readers are hopefully wondering if or if not the principle of inertia holds w.r.t. any reference frame. The answer is *no*; the principle of inertia only holds w.r.t. an inertial reference frame. Well, what is an inertial reference frame?

Before defining an inertial reference frame, a particle and an unforced particle will be defined.

Definition 4.2 (A Particle)**.** A *particle* has zero size and strictly positive mass.

Definition 4.3 (An Unforced Particle)**.** An *unforced particle* has no force applied to it.

It will also serve useful to define a *massive particle* [3, Sec. 1.1].

Definition 4.4 (A massive particle)**.** A *massive particle* is a particle with infinite mass.

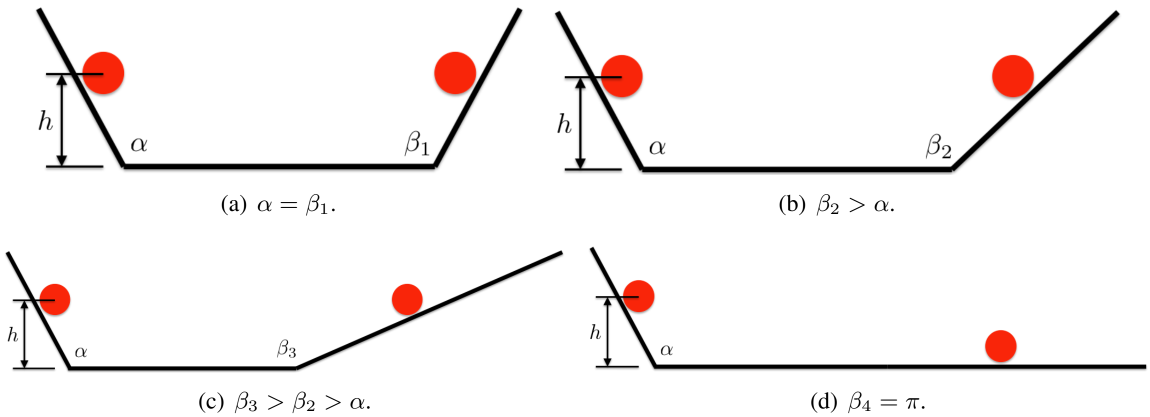


Figure 4.1: Galileo's experiment. A ball rolls down an inclined plane and up another inclined plane.

Massive particles are unaffected by forces. As such, a massive particle is in effect an unforced particle. The definition of an inertial reference frame, or an inertial frame for short, can now be given.

Definition 4.5 (Inertial Reference Frame)**.** The reference frame $\mathcal{F}_a$ is an *inertial reference frame* if for all unforced particles y and w [3, Sec. 8.1]

$$ {\underrightarrow{v}^{yw/a}}^{\bullet a} = \underrightarrow{a}^{yw/a/a} = \underrightarrow{0}. $$

Consider Figure 4.2 depicting the pairs of unforced particles y_1, w_1 and y_2, w_2. The unforced particles are permitted to move relative to each other in time. The position of y_1 relative to w_1 is $\underrightarrow{r}^{y_1 w_1}$, and the position of y_2 relative to w_2 is $\underrightarrow{r}^{y_2 w_2}$. The reference frame $\mathcal{F}_a$ is an inertial frame if

$$ {\underrightarrow{r}^{y_1 w_1}}^{\bullet a} = \underrightarrow{v}^{y_1 w_1/a} = \underrightarrow{\alpha}^{1}, \quad {\underrightarrow{r}^{y_2 w_2}}^{\bullet a} = \underrightarrow{v}^{y_2 w_2/a} = \underrightarrow{\alpha}^{2}, $$

where $\underrightarrow{\alpha}^{1}$ and $\underrightarrow{\alpha}^{2}$ are constant w.r.t. $\mathcal{F}_a$. Also, recall that the physical basis vectors $\underrightarrow{a}^{1}$, $\underrightarrow{a}^{2}$, and $\underrightarrow{a}^{3}$ do not have to be drawn together with their tails collocated at a point. The physical basis vectors do not have a physical location; only points and particles have physical locations relative to each other.

The principle of inertia looks awfully similar to Newton's first law (N1L). In fact, Newton himself attributes "his" first law to Galileo [2]. However, Galileo's principle of inertia is not all that precise. As such, N1L will be precisely defined, just as an inertial frame was precisely defined.

Definition 4.6 (Newton's First Law (N1L) [38])**.** There exists an inertial reference frame, that is, for all unforced particles y and w there exists a reference frame $\mathcal{F}_a$ called an *inertial reference frame* such that

$$ {\underrightarrow{v}^{yw/a}}^{\bullet a} = \underrightarrow{a}^{yw/a/a} = \underrightarrow{0}. $$

Often N1L is stated as something like "an object remains in uniform motion unless acted upon by an external force". Such statements are imprecise.

A natural question to ask is if there is only one single inertial frame. The answer to this question is "no", as proven next.

Theorem 4.1. Let $\mathcal{F}_b$ be an inertial frame and let $\mathcal{F}_a$ be another frame. Then $\mathcal{F}_a$ is an inertial frame if and only if $\underrightarrow{\omega}^{ba} = \underrightarrow{0}$ [3, Sec. 8.1].

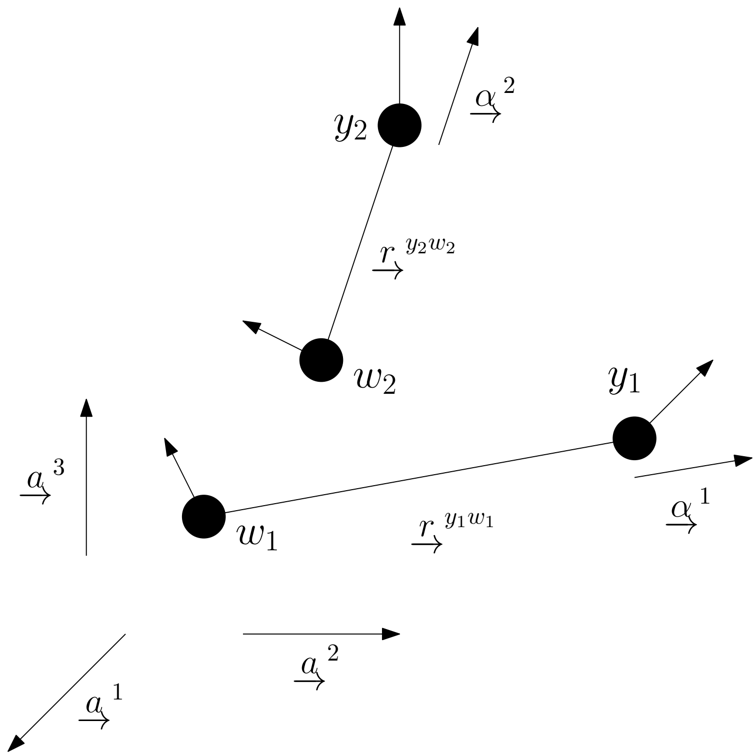


Figure 4.2: An inertial frame $\mathcal{F}_a$.

Proof. Sufficiency, the "if" part, will be proven, meaning $\underrightarrow{\omega}^{ba} = \underrightarrow{0} \Rightarrow \mathcal{F}_a$ is an inertial frame will be proved. To begin, recall that

$$\underrightarrow{v}^{yw/a^{\bullet a}} = \underrightarrow{v}^{yw/b^{\bullet b}} + 2\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{yw/b} + \underrightarrow{\omega}^{ba^{\bullet b}} \times \underrightarrow{r}^{yw} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{r}^{yw}\right).$$

If $\underrightarrow{\omega}^{ba} = \underrightarrow{0}$ then $\underrightarrow{v}^{yw/a^{\bullet a}} = \underrightarrow{v}^{yw/b^{\bullet b}}$. If $\mathcal{F}_b$ is an inertial frame then $\underrightarrow{v}^{yw/b^{\bullet b}} = \underrightarrow{0}$, which implies $\underrightarrow{v}^{yw/a^{\bullet a}} = \underrightarrow{0}$, which in turn implies that $\mathcal{F}_a$ is an inertial frame.

Necessity, the "only if" part, is more difficult to prove, and is omitted. $\square$

Theorem 4.1 says that if $\mathcal{F}_b$ is an inertial frame and $\mathcal{F}_a$ is an inertial frame then $\underrightarrow{\omega}^{ba} = \underrightarrow{0}$. As such, $\mathcal{F}_b$ cannot rotate relative to $\mathcal{F}_a$. Additionally, recall that the physical basis vectors $\underrightarrow{a}^1$, $\underrightarrow{a}^2$, $\underrightarrow{a}^3$ and $\underrightarrow{b}^1$, $\underrightarrow{b}^2$, $\underrightarrow{b}^3$ do not have a physical location. As such, one cannot say $\mathcal{F}_a$ and $\mathcal{F}_b$ can "translate".

4.2 The Dynamics of a Single Particle

Before discussing the dynamics of a single particle, forces and moments will be discussed.

4.2.1 Forces

Linear Spring Force

Consider two particles, y and w, and a spring connected between the two particles. Let the unstretched length of the spring be $0 \leq d < \infty$ and the spring stiffness be $0 < k < \infty$. The force on y due to the spring

s is

$$\underset{\rightarrow}{f}^{y} = -k(\ell - d)\frac{\underset{\rightarrow}{r}^{yw}}{||\underset{\rightarrow}{r}^{yw}||_2},$$

where $\ell = ||\underset{\rightarrow}{r}^{yw}||_2$.

As an example, consider when particles y and w move only along the "1" axis of a frame, say $\mathcal{F}_a$, that being

$$\begin{aligned}\underset{\rightarrow}{f}^{y} &= -k(\ell - d)\frac{\underset{\rightarrow}{r}^{yw}}{||\underset{\rightarrow}{r}^{yw}||_2} \\ &= \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} -k(x_a - d) \\ 0 \\ 0 \end{bmatrix},\end{aligned}$$

where

$$\underset{\rightarrow}{r}^{yw} = \underset{\rightarrow}{\mathcal{F}}_a^{\mathsf{T}} \begin{bmatrix} x_a \\ 0 \\ 0 \end{bmatrix}.$$

Linear Viscous Damping

Consider two particles, y and w, and a viscous damper connected between the two particles. Let the damping coefficient be $0 < c < \infty$. The force on y due to the viscous damper d, also called a dashpot, is

$$\underset{\rightarrow}{f}^{y} = -c\,\underset{\rightarrow}{v}^{yw/a}.$$

Gravitational Force

Consider two particles, y and x, who's masses are m_y and m_x, respectively. The gravitational force on y due to x is

$$\underset{\rightarrow}{f}^{y} = -\frac{Gm_x m_y}{||\underset{\rightarrow}{r}^{yx}||_2^3}\underset{\rightarrow}{r}^{yx}$$

where G is the universal gravitational constant. If the motion of y is small in comparison to $\underset{\rightarrow}{r}^{yx}$ the force can be approximated as

$$\underset{\rightarrow}{f}^{y} = m_y\,\underset{\rightarrow}{g}, \quad \underset{\rightarrow}{g} = -\frac{Gm_x}{||\underset{\rightarrow}{r}^{yx}||_2^3}\underset{\rightarrow}{r}^{yx},$$

where $\underset{\rightarrow}{r}^{yx}$ is constant.

4.2.2 Moments

Definition 4.7. Let w be a point, y be a particle, and $\underset{\rightarrow}{f}^{y}$ be the force applied to y, as shown in Figure 4.3. The *moment on y relative to w due to $\underset{\rightarrow}{f}^{y}$* is defined as

$$\underset{\rightarrow}{m}^{yw} = \underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{f}^{y}.$$

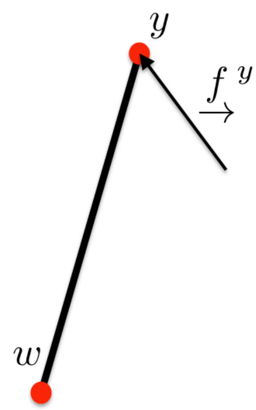


Figure 4.3: A point w, a particle y, and a force $\underset{\rightarrow}{f}^{\,y}$.

4.2.3 Newton's Second Law and Kinetic Energy for a Single Particle

Definition 4.8. Let $\mathcal{F}_a$ be frame, let w be a point, and let y be a particle of mass m. The *translational momentum of particle y relative to w w.r.t. $\mathcal{F}_a$* is defined as [3, Sec. 8.4]

$$\underset{\rightarrow}{p}^{\,yw/a} = m\,\underset{\rightarrow}{v}^{\,yw/a}.$$

Axiom 4.1 (Newton's Second Law (N2L))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let y be a particle of mass m, and let $\underset{\rightarrow}{f}^{\,y}$ be the force acting on particle y. Then [3, Sec. 8.4]

$$m\,{\underset{\rightarrow}{v}^{\,yw/a}}^{\bullet a} = \underset{\rightarrow}{f}^{\,y}.$$

Notice that N2L is composed of an inertial frame $\mathcal{F}_a$, the particle y with force $\underset{\rightarrow}{f}^{\,y}$ acting on it, and another unforced particle w. When the force $\underset{\rightarrow}{f}^{\,y}$ is zero the particle y becomes an unforced particle; naturally, y and w together satisfy ${\underset{\rightarrow}{v}^{\,yw/a}}^{\bullet a} = \underset{\rightarrow}{0}$ as in the definition of an inertial frame. This is an subtle but very important but point. When $\underset{\rightarrow}{f}^{\,y}$ is zero it must be that N1L is recovered, and hence the definition of an inertial frame is recovered.

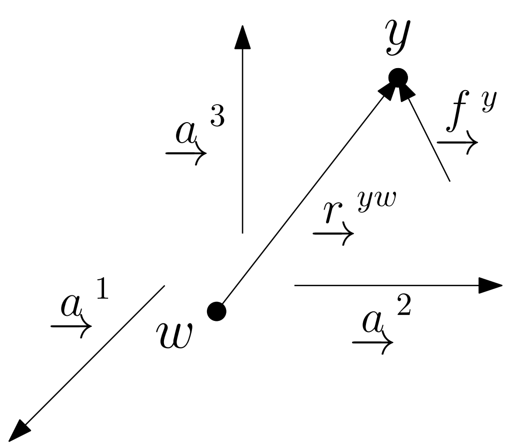


Figure 4.4: Force applied to one particle.

Definition 4.9. Let $\mathcal{F}_a$ be frame, let w be a point, and let y be a particle of mass m. The *angular momentum of particle y relative to w w.r.t. $\mathcal{F}_a$* is defined as [3, Sec. 8.6]

$$\underset{\rightarrow}{h}^{yw/a} = \underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{p}^{yw/a} = m\,\underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{v}^{yw/a}. \tag{4.1}$$

Axiom 4.2 (Newton's Second Law in Rotation (N2LR))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let y be a particle of mass m, and let $\underset{\rightarrow}{f}^{y}$ be the force acting on particle y. Then

$$\underset{\rightarrow}{h}^{yw/a^{\bullet a}} = \underset{\rightarrow}{m}^{yw}, \tag{4.2}$$

where $\underset{\rightarrow}{m}^{yw}$ is the moment on y relative to w.

Proof. Although N2LR is an axiom, using Equation (4.1), Equation (4.2) can be derived explicitly, that is

$$\begin{aligned}
\underset{\rightarrow}{h}^{yw/a^{\bullet a}} &= \underset{\rightarrow}{r}^{yw^{\bullet a}} \times \underset{\rightarrow}{p}^{yw/a} + \underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{p}^{yw/a^{\bullet a}} \\
&= m\,\underset{\rightarrow}{r}^{yw^{\bullet a}} \times \underset{\rightarrow}{r}^{yw^{\bullet a}} + \underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{p}^{yw/a^{\bullet a}} \\
&= \underset{\rightarrow}{r}^{yw} \times \underset{\rightarrow}{f}^{y} \\
&= \underset{\rightarrow}{m}^{yw}.
\end{aligned}$$

□

Definition 4.10 (Kinetic Energy of a Particle)**.** Let $\mathcal{F}_a$ be frame, let w be a point, and let y be a particle of mass m. The *kinetic energy of y relative to w w.r.t. $\mathcal{F}_a$* is defined as

$$T_{yw/a} = \tfrac{1}{2} m\,\underset{\rightarrow}{v}^{yw/a} \cdot \underset{\rightarrow}{v}^{yw/a}.$$

4.3 The Dynamics of Many Particles

The dynamics of many particles that interact with each other will now be considered. Newton's Third Law (N3L) will be stated first.

Definition 4.11 (Newton's Third Law (N3L))**.** Consider two particles, y and x, connected via a rigid massless link, a spring, or dashpot. Then,

$$\underset{\rightarrow}{f}^{yx} = -\underset{\rightarrow}{f}^{xy}.$$

4.3.1 Properties of Discrete-Rigid Bodies

Definition 4.12. A *discrete body* is a finite collection of particles.

Definition 4.13. A *discrete-rigid body* is a finite collection of particles such that the distance between each pair of particles is constant.

Definition 4.14. A *massive discrete-rigid body* is a discrete-rigid body composed of at least three massive particles that are not colinear.

Massive discrete-rigid bodies are unaffected by forces and moments, which in turn means each massive particle composing a massive discrete-rigid body are effectively unforced [3, Sec. 1.1].

Related to *massive discrete-rigid body* is an *inertially nonrotating massive discrete-rigid body* [3, Sec. 1.1].

Definition 4.15. Let $\mathcal{F}_a$ be an inertial fame, and $\mathcal{F}_b$ be a body-fixed frame affixed to a massive discrete-rigid body. The massive discrete-rigid body is an *inertially nonrotating massive discrete-rigid body* if $\underset{\rightarrow}{\omega}^{ba} = \underset{\rightarrow}{0}$.

An inertially nonrotating massive discrete-rigid body is important because any body-fixed frame associated with the massive discrete-rigid body is an inertial frame.

It will be helpful to define the mass and centre of mass.

Definition 4.16. The *zeroth moment of mass*, or the *mass*, of body $\mathcal{B}$ composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, is defined as [3, Sec. 7.1]

$$m_{\mathcal{B}} = \sum_{i=1}^{\ell} m_i. \tag{4.3}$$

Definition 4.17. Consider a body $\mathcal{B}$ composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. Let $m_{\mathcal{B}}$ be the mass of $\mathcal{B}$ and z be a point. The *centre of mass c of* $\mathcal{B}$ is defined as [3, Sec. 7.1]

$$\underset{\rightarrow}{r}^{cz} = \frac{1}{m_{\mathcal{B}}} \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z}. \tag{4.4}$$

Next, the first moment of mass is defined.

Definition 4.18. Consider a body $\mathcal{B}$ composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and a point z. The *first moment of mass of body $\mathcal{B}$ relative to point z* is defined as [3, Sec. 7.1]

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z}. \tag{4.5}$$

Using the definition of mass given in Equation (4.4), Equation (4.5) can be written

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} = m_{\mathcal{B}} \underset{\rightarrow}{r}^{cz}. \tag{4.6}$$

The second moment of mass, also called the moment of inertia or the mass-moment of inertia, will now be defined.

Definition 4.19. Consider a body $\mathcal{B}$ composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and a point z. The *second moment of mass of body $\mathcal{B}$ relative to point z* is defined as [1, pp. 43]

$$\underset{\rightarrow}{J}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \left(\underset{\rightarrow}{r}^{y_i z} \cdot \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{r}^{y_i z} \right), \tag{4.7}$$

where $\underset{\rightarrow}{1}$ is the identity tensor.

4.3.2 Forces and Moments

Definition 4.20. Let $\mathcal{B}$ be a body, z a point, y_i be a particle in $\mathcal{B}$, and $\underset{\rightarrow}{f}^{y_i}$ be a force applied to y_i, where $i = 1, 2, \ldots, \ell$. The *moment on $\mathcal{B}$ relative to z due to* $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ is [3, Sec. 7.2]

$$\underset{\rightarrow}{m}^{\mathcal{B}z} = \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i}.$$

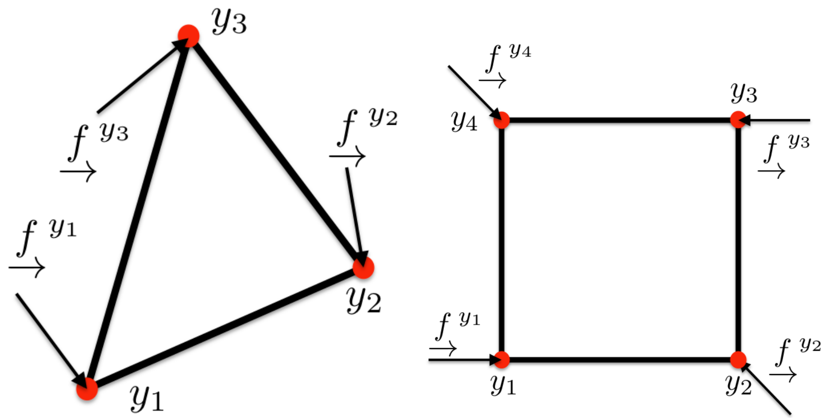
Note, point z does not have to be fixed in the body $\mathcal{B}$.

Definition 4.21. Let $\mathcal{B}$ be a body, y_i be a particle in $\mathcal{B}$, and $\underset{\rightarrow}{f}^{y_i}$ be a force applied to y_i, where $i = 1, 2, \ldots, \ell$. The *total force on $\mathcal{B}$ due to* $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ is [3, Sec. 7.2]

$$\underset{\rightarrow}{f}^{\mathcal{B}} = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i}.$$

The forces $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ are *balanced* if $\underset{\rightarrow}{f}^{\mathcal{B}} = \underset{\rightarrow}{0}$.

Shown in Figure 4.5(a) are forces that are not balanced, while in Figure 4.5(b) depicts balanced forces.



(a) Forces applied to points in the discrete-rigid body $\mathcal{B}$.

(b) Balanced forces.

Figure 4.5: Forces applied at points and balanced forces.

Theorem 4.2. Let $\mathcal{B}$ be a body, z and x be points, y_i be a particle in $\mathcal{B}$, and $\underset{\rightarrow}{f}^{y_i}$ be a force applied to y_i, where $i = 1, 2, \ldots, \ell$. Then [3, Sec. 7.2]

$$\underset{\rightarrow}{m}^{\mathcal{B}z} = \underset{\rightarrow}{m}^{\mathcal{B}x} + \underset{\rightarrow}{r}^{xz} \times \underset{\rightarrow}{f}^{\mathcal{B}}.$$

Proof.

$$\begin{aligned}
\underset{\rightarrow}{m}^{\mathcal{B}z} &= \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i} \\
&= \sum_{i=1}^{\ell} \left(\underset{\rightarrow}{r}^{y_i x} + \underset{\rightarrow}{r}^{xz} \right) \times \underset{\rightarrow}{f}^{y_i} \\
&= \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_i x} \times \underset{\rightarrow}{f}^{y_i} + \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{xz} \times \underset{\rightarrow}{f}^{y_i} \\
&= \underset{\rightarrow}{m}^{\mathcal{B}x} + \underset{\rightarrow}{r}^{xz} \times \underset{\rightarrow}{f}^{\mathcal{B}}
\end{aligned}$$

If $\underset{\rightarrow}{f}^{\mathcal{B}} = \underset{\rightarrow}{0}$ then $\underset{\rightarrow}{m}^{\mathcal{B}z} = \underset{\rightarrow}{m}^{\mathcal{B}x}$.

□

Theorem 4.3. Let $\mathcal{B}$ be a body, z a point, y_1 and y_2 be particles fixed in $\mathcal{B}$, and $\underset{\rightarrow}{f}^{y_1}$ and $\underset{\rightarrow}{f}^{y_2} = -\underset{\rightarrow}{f}^{y_1}$ be forces applied to y_1 and y_2. Assume that $\underset{\rightarrow}{f}^{y_1}$ and $\underset{\rightarrow}{f}^{y_2}$ are the only forces applied to $\mathcal{B}$. Then [3, Sec. 7.2]

$$\underset{\rightarrow}{m}^{\mathcal{B}z} = \underset{\rightarrow}{m}^{y_1y_2} = \underset{\rightarrow}{m}^{y_2y_1}$$

Proof. Note that the forces are balanced (i.e., $\underset{\rightarrow}{f}^{\mathcal{B}} = \underset{\rightarrow}{0}$). Consider $\underset{\rightarrow}{m}^{\mathcal{B}z}$:

$$\begin{aligned}
\underset{\rightarrow}{m}^{\mathcal{B}z} &= \underset{\rightarrow}{r}^{y_1z} \times \underset{\rightarrow}{f}^{y_1} + \underset{\rightarrow}{r}^{y_2z} \times \underset{\rightarrow}{f}^{y_2} \\
&= \left(\underset{\rightarrow}{r}^{y_1y_2} + \underset{\rightarrow}{r}^{y_2z}\right) \times \underset{\rightarrow}{f}^{y_1} + \underset{\rightarrow}{r}^{y_2z} \times \left(-\underset{\rightarrow}{f}^{y_1}\right) \\
&= \underset{\rightarrow}{r}^{y_1y_2} \times \underset{\rightarrow}{f}^{y_1} = \underset{\rightarrow}{m}^{y_1y_2} \\
&= -\underset{\rightarrow}{r}^{y_2y_1} \times \left(-\underset{\rightarrow}{f}^{y_2}\right) = \underset{\rightarrow}{m}^{y_2y_1}
\end{aligned}$$

□

The significance of this result is that the moment is independent of the point z. This leads to the definition of a *torque*.

Definition 4.22. Let $\mathcal{B}$ be a body, y_i be particles in $\mathcal{B}$, and $\underset{\rightarrow}{f}^{y_i}$ be forces applied to each y_i, where $i = 1, 2, \ldots, \ell$, and assume that $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ are balanced. The *torque on $\mathcal{B}$ due to* $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ is [3, Sec. 7.2]

$$\underset{\rightarrow}{\tau}^{\mathcal{B}} = \sum_{i=1,\, j\neq i}^{\ell} \underset{\rightarrow}{m}^{y_iy_j}.$$

Proof. To see why $\underset{\rightarrow}{\tau}^{\mathcal{B}} = \sum_{i=1,\, j\neq i}^{\ell} \underset{\rightarrow}{m}^{y_iy_j}$, consider a point z, let $j = \ell$, and compute $\underset{\rightarrow}{m}^{\mathcal{B}z}$, that is

$$\begin{aligned}
\underset{\rightarrow}{m}^{\mathcal{B}z} &= \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{f}^{y_i} \\
&= \sum_{i=1}^{\ell-1} \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{f}^{y_i} + \underset{\rightarrow}{r}^{y_\ell z} \times \underset{\rightarrow}{f}^{y_\ell} \\
&= \sum_{i=1}^{\ell-1} \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{f}^{y_i} + \underset{\rightarrow}{r}^{y_\ell z} \times \left(-\sum_{i=1}^{\ell-1} \underset{\rightarrow}{f}^{y_i}\right) \\
&= \sum_{i=1}^{\ell-1} \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{f}^{y_i} + \sum_{i=1}^{\ell-1} \underset{\rightarrow}{r}^{zy_\ell} \times \underset{\rightarrow}{f}^{y_i} \\
&= \sum_{i=1}^{\ell-1} \left(\underset{\rightarrow}{r}^{y_iz} + \underset{\rightarrow}{r}^{zy_\ell}\right) \times \underset{\rightarrow}{f}^{y_i} \\
&= \sum_{i=1}^{\ell-1} \left(\underset{\rightarrow}{r}^{y_iy_\ell}\right) \times \underset{\rightarrow}{f}^{y_i} = \sum_{i=1}^{\ell-1} \underset{\rightarrow}{m}^{y_iy_\ell} = \underset{\rightarrow}{\tau}^{\mathcal{B}}.
\end{aligned}$$

This is independent of the point z.

□

4.3.3 Newton's Second Law for Many Particles

Definition 4.23. Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. The *translational momentum of body $\mathcal{B}$ relative to w w.r.t. $\mathcal{F}_a$* is defined as [3, Sec. 8.4]

$$\underrightarrow{p}^{\mathcal{B}w/a} = \sum_{i=1}^{\ell} \underrightarrow{p}^{y_i w/a} = \sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_i w/a}.$$

Using the definition of the centre of mass given in Equation (4.4),

$$\underrightarrow{r}^{cw} = \frac{1}{m_{\mathcal{B}}} \sum_{i=1}^{\ell} m_i \underrightarrow{r}^{y_i w},$$

the expression for $\underrightarrow{p}^{\mathcal{B}w/a}$ can alternatively be written as

$$\begin{aligned} \underrightarrow{p}^{\mathcal{B}w/a} &= \sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_i w/a} \\ &= m_{\mathcal{B}} \underrightarrow{r}^{cw^{\bullet a}} \\ &= m_{\mathcal{B}} \underrightarrow{v}^{cw/a} \end{aligned} \tag{4.8}$$

Axiom 4.3 (Newton's Second Law For a System of Particles (N2L))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and let $\underrightarrow{f}^{y_i}$ be the external force applied to particle y_i. Then [3, Sec. 8.4]

$$\underrightarrow{p}^{\mathcal{B}w/a^{\bullet a}} = m_{\mathcal{B}} \underrightarrow{v}^{cw/a^{\bullet a}} = \underrightarrow{f}^{\mathcal{B}}, \tag{4.9}$$

Proof. The equations of motion of each individual particle y_i are

$$m_i \underrightarrow{v}^{y_i w/a^{\bullet a}} = \underrightarrow{f}^{y_i} + \sum_{j=1}^{\ell} \underrightarrow{f}^{y_i y_j}$$

where $\underrightarrow{f}^{y_i}$ are external forces acting on particle y_i and $\underrightarrow{f}^{y_i y_j}$ is the force acting on particle y_i due to particle y_j. Summing over all particles yields

$$\begin{aligned} \sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_i w/a^{\bullet a}} &= \sum_{i=1}^{\ell} \underrightarrow{f}^{y_i} + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \underrightarrow{f}^{y_i y_j}, \\ \underrightarrow{p}^{\mathcal{B}w/a^{\bullet a}} &= \underrightarrow{f}^{\mathcal{B}}, \end{aligned}$$

where

$$\underrightarrow{p}^{\mathcal{B}w/a^{\bullet a}} = \sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_i w/a^{\bullet a}}, \qquad \underrightarrow{f}^{\mathcal{B}} = \sum_{i=1}^{\ell} \underrightarrow{f}^{y_i},$$

and, from N3L,

$$\sum_{i=1}^{\ell}\sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_iy_j} = \underset{\rightarrow}{0}. \tag{4.10}$$

Furthermore, from Equation (4.8)

$$\underset{\rightarrow}{p}^{\mathcal{B}w/a^{\bullet a}} = m_{\mathcal{B}} \underset{\rightarrow}{v}^{cw/a^{\bullet a}},$$

it follows that

$$m_{\mathcal{B}} \underset{\rightarrow}{v}^{cw/a^{\bullet a}} = \underset{\rightarrow}{f}^{\mathcal{B}}.$$

□

Definition 4.24. Let $\mathcal{F}_a$ be frame, let z be a point, and let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. The *angular momentum of body $\mathcal{B}$ relative to z w.r.t. $\mathcal{F}_a$* is defined as [3, Sec. 8.6]

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a} = \sum_{i=1}^{\ell} \underset{\rightarrow}{h}^{y_iz/a} = \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{p}^{y_iz/a} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{v}^{y_iz/a}.$$

Axiom 4.4 (Newton's Second Law in Rotation for a System of Particles (N2LR))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and let $\underset{\rightarrow}{f}^{y_i}$ be the external force applied to particle y_i. Then

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a^{\bullet a}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{m}^{\mathcal{B}z}. \tag{4.11}$$

where $\underset{\rightarrow}{m}^{\mathcal{B}z}$ is the moment on $\mathcal{B}$ relative to z. When point z is collocated with point c,

$$\underset{\rightarrow}{h}^{\mathcal{B}c/a^{\bullet a}} = \underset{\rightarrow}{m}^{\mathcal{B}c}.$$

Proof. Again recall that the equations of motion of each individual particle y_i are

$$m_i \underset{\rightarrow}{v}^{y_iw/a^{\bullet a}} = \underset{\rightarrow}{f}^{y_i} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_iy_j}. \tag{4.12}$$

Using the fact that $\underset{\rightarrow}{v}^{y_iw/a^{\bullet a}} = \underset{\rightarrow}{v}^{y_iz/a^{\bullet a}} + \underset{\rightarrow}{v}^{zw/a^{\bullet a}}$ Equation (4.12) can be written

$$\begin{aligned} m_i \left(\underset{\rightarrow}{v}^{y_iz/a^{\bullet a}} + \underset{\rightarrow}{v}^{zw/a^{\bullet a}} \right) &= \underset{\rightarrow}{f}^{y_i} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_iy_j}, \\ m_i \underset{\rightarrow}{v}^{y_iz/a^{\bullet a}} &= \underset{\rightarrow}{f}^{y_i} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_iy_j} - m_i \underset{\rightarrow}{v}^{zw/a^{\bullet a}}. \end{aligned} \tag{4.13}$$

Taking the derivative w.r.t. $\mathcal{F}_a$ of $\underset{\rightarrow}{h}^{\mathcal{B}z/a} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_iz} \times \underset{\rightarrow}{v}^{y_iz/a}$ and employing the expression for

$m_i \underset{\rightarrow}{v}^{y_i z/a^{\bullet a}}$ given in Equation (4.13) gives

$$
\begin{aligned}
\underset{\rightarrow}{h}^{\mathcal{B}z/a^{\bullet a}} &= \sum_{i=1}^{\ell}\left(m_i \underbrace{\underset{\rightarrow}{r}^{y_i z^{\bullet a}} \times \underset{\rightarrow}{v}^{y_i z/a}}_{\underset{\rightarrow}{0}} + m_i \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{v}^{y_i z/a^{\bullet a}}\right) \\
&= \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{v}^{y_i z/a^{\bullet a}} \\
&= \sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \left(\underset{\rightarrow}{f}^{y_i} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_i y_j} - m_i \underset{\rightarrow}{v}^{zw/a^{\bullet a}}\right) \\
&= \sum_{i=1}^{\ell}\left(\underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i} + \sum_{j=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i y_j} - m_i \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}}\right) \\
&= \underbrace{\sum_{i=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i}}_{\underset{\rightarrow}{m}^{\mathcal{B}z}} + \underbrace{\sum_{i=1}^{\ell}\sum_{j=1}^{\ell} \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{f}^{y_i y_j}}_{\underset{\rightarrow}{0}} - \underbrace{\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}}}_{\underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}}} \\
&= \underset{\rightarrow}{m}^{\mathcal{B}z} - \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}}
\end{aligned}
$$

When point z collocates with point c, $\underset{\rightarrow}{c}^{\mathcal{B}c} = \underset{\rightarrow}{0}$ and therefore

$$\underset{\rightarrow}{h}^{\mathcal{B}c/a^{\bullet a}} = \underset{\rightarrow}{m}^{\mathcal{B}c}$$

□

4.3.3.1 Alternative Forms

Equations (4.9) and (4.11) can be written in two alternative forms.

Theorem 4.4. Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a discrete-rigid body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\mathcal{F}_b$ be a body-fixed frame, and let $\underset{\rightarrow}{f}^{y_i}$ be the external force applied to particle y_i. Then

$$m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{c}^{\mathcal{B}z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{c}^{\mathcal{B}z}\right) = \underset{\rightarrow}{f}^{\mathcal{B}} \tag{4.14}$$

where $m_{\mathcal{B}}$ is the mass of the body $\mathcal{B}$ and

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} = m_{\mathcal{B}} \underset{\rightarrow}{r}^{cz}.$$

is the first moment of mass of body $\mathcal{B}$ relative to point z. Equation (4.14) is an alternative form of Equation (4.9).

Proof. Recall that $\underset{\rightarrow}{r}^{y_i w} = \underset{\rightarrow}{r}^{y_i z} + \underset{\rightarrow}{r}^{zw}$ and therefore

$$\begin{aligned}
\underset{\rightarrow}{r}^{y_i w^{\bullet a}} &= \underset{\rightarrow}{r}^{y_i z^{\bullet a}} + \underset{\rightarrow}{r}^{zw^{\bullet a}}, \\
\underset{\rightarrow}{v}^{y_i w/a} &= \underset{\rightarrow}{v}^{y_i z/a} + \underset{\rightarrow}{v}^{zw/a}.
\end{aligned} \tag{4.15}$$

Next, write $\underset{\rightarrow}{p}^{\mathcal{B}w/a} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i w/a}$ as

$$\begin{aligned}
\underset{\rightarrow}{p}^{\mathcal{B}w/a} &= \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i w/a} \\
&= \sum_{i=1}^{\ell} m_i \left(\underset{\rightarrow}{v}^{y_i z/a} + \underset{\rightarrow}{v}^{zw/a} \right) \\
&= \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i z/a} + m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a},
\end{aligned} \tag{4.16}$$

where Equation (4.15) has been used. Taking the derivative w.r.t. $\mathcal{F}_a$ of Equation (4.16), recalling from Equation (4.9) that $\underset{\rightarrow}{p}^{\mathcal{B}w/a^{\bullet a}} = \underset{\rightarrow}{f}^{\mathcal{B}}$, and using

$$\begin{aligned}
\underset{\rightarrow}{v}^{y_i z/a^{\bullet a}} &= \underset{\rightarrow}{v}^{y_i z/b^{\bullet b}} + 2\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{y_i z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z} \right) \\
&= \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{y_i z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z} \right),
\end{aligned}$$

where $\underset{\rightarrow}{v}^{y_i z/b^{\bullet b}}$ and $2\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z^{\bullet b}}$ equal $\underset{\rightarrow}{0}$ because the body is rigid, results in

$$\begin{aligned}
\underset{\rightarrow}{f}^{\mathcal{B}} &= \underset{\rightarrow}{p}^{\mathcal{B}w/a^{\bullet a}} \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet a}} + \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i z/a^{\bullet a}} \\
&= m_{\mathcal{B}} \left(\underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} \right) + \sum_{i=1}^{\ell} m_i \left(\underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{y_i z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z} \right) \right) \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \left[\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \right] + \underset{\rightarrow}{\omega}^{ba} \times \left[\underset{\rightarrow}{\omega}^{ba} \times \left(\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \right) \right] \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{c}^{\mathcal{B}z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{c}^{\mathcal{B}z} \right).
\end{aligned}$$

$\square$

Theorem 4.5. Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a discrete-rigid body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\mathcal{F}_b$ be a body-fixed frame, and let $\underset{\rightarrow}{f}^{y_i}$ be the external force applied to particle y_i. Then

$$\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba} \right) + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} \right) = \underset{\rightarrow}{m}^{\mathcal{B}z}, \tag{4.17}$$

where

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} = m_{\mathcal{B}} \underset{\rightarrow}{r}^{cz}.$$

is the first moment of mass of body $\mathcal{B}$ relative to point z and

$$\underset{\rightarrow}{J}^{\mathcal{B}z} = \sum_{i=1}^{\ell} m_i \left(\underset{\rightarrow}{r}^{y_i z} \cdot \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{r}^{y_i z} \right)$$

is the second moment of mass of body $\mathcal{B}$ relative to point z. Equation (4.17) is an alternative form of Equation (4.11).

Proof. First, note that

$$\underset{\rightarrow}{v}^{y_i z/a} = \underset{\rightarrow}{r}^{y_i z^{\bullet a}} = \underset{\rightarrow}{r}^{y_i z^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z} = \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z},$$

where $\underset{\rightarrow}{r}^{y_i z^{\bullet b}} = \underset{\rightarrow}{0}$ owing to the fact the body is a discrete-rigid body. Therefore,

$$\begin{aligned}
\underset{\rightarrow}{h}^{\mathcal{B}z/a} &= \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{v}^{y_i z/a} \\
&= \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_i z} \right) \\
&= -\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z} \times \left(\underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{\omega}^{ba} \right).
\end{aligned}$$

Using the identity

$$-\underset{\rightarrow}{r}^{y_i z} \times \left(\underset{\rightarrow}{r}^{y_i z} \times \underset{\rightarrow}{\omega}^{ba} \right) = \left(\underset{\rightarrow}{r}^{y_i z} \cdot \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{r}^{y_i z} \right) \cdot \underset{\rightarrow}{\omega}^{ba},$$

it follows that

$$\begin{aligned}
\underset{\rightarrow}{h}^{\mathcal{B}z/a} &= \left(\sum_{i=1}^{\ell} m_i \left(\underset{\rightarrow}{r}^{y_i z} \cdot \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{y_i z} \underset{\rightarrow}{r}^{y_i z} \right) \right) \cdot \underset{\rightarrow}{\omega}^{ba} \\
&= \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}.
\end{aligned}$$

Next, returning to

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a^{\bullet a}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{m}^{\mathcal{B}z},$$

using

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a} = \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba} \quad \text{and} \quad \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}$$

results in

$$
\begin{aligned}
\underrightarrow{h}^{\mathcal{B}z/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{h}^{\mathcal{B}z/a} + \underrightarrow{c}^{\mathcal{B}z} \times \left(\underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a} \right) &= \underrightarrow{m}^{\mathcal{B}z}, \\
\left(\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba} \right)^{\bullet b} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba} \right) + \underrightarrow{c}^{\mathcal{B}z} \times \underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{c}^{\mathcal{B}z} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a} \right) &= \underrightarrow{m}^{\mathcal{B}z}, \\
\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba} \right) + \underrightarrow{c}^{\mathcal{B}z} \times \underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{c}^{\mathcal{B}z} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a} \right) &= \underrightarrow{m}^{\mathcal{B}z}.
\end{aligned}
$$

$\square$

4.3.3.2 Referential Form of Translational and Rotational Equations of Motion

Equations (4.14) and (4.17),

$$
\begin{aligned}
m_{\mathcal{B}} \underrightarrow{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a} + \underrightarrow{\omega}^{ba^{\bullet b}} \times \underrightarrow{c}^{\mathcal{B}z} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{c}^{\mathcal{B}z} \right) &= \underrightarrow{f}^{\mathcal{B}}, \\
\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \left(\underrightarrow{J}^{\mathcal{B}z} \cdot \underrightarrow{\omega}^{ba} \right) + \underrightarrow{c}^{\mathcal{B}z} \times \underrightarrow{v}^{zw/a^{\bullet b}} + \underrightarrow{c}^{\mathcal{B}z} \times \left(\underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{zw/a} \right) &= \underrightarrow{m}^{\mathcal{B}z},
\end{aligned}
$$

can be written in referential form, that being

$$
\begin{aligned}
m_{\mathcal{B}} \dot{\mathbf{v}}_b^{zw/a} + m_{\mathcal{B}} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} + \dot{\boldsymbol{\omega}}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z} + \boldsymbol{\omega}_b^{ba^\times} \boldsymbol{\omega}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z} &= \mathbf{f}_b^{\mathcal{B}}, && (4.18) \\
\mathbf{J}_b^{\mathcal{B}z} \dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{J}_b^{\mathcal{B}z} \boldsymbol{\omega}_b^{ba} + \mathbf{c}_b^{\mathcal{B}z^\times} \dot{\mathbf{v}}_b^{zw/a} + \mathbf{c}_b^{\mathcal{B}z^\times} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} &= \mathbf{m}_b^{\mathcal{B}z}. && (4.19)
\end{aligned}
$$

Rewriting the above as

$$
\begin{aligned}
m_{\mathcal{B}} \dot{\mathbf{v}}_b^{zw/a} - \mathbf{c}_b^{\mathcal{B}z^\times} \dot{\boldsymbol{\omega}}_b^{ba} + m_{\mathcal{B}} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} - \boldsymbol{\omega}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z^\times} \boldsymbol{\omega}_b^{ba} &= \mathbf{f}_b^{\mathcal{B}}, \\
\mathbf{c}_b^{\mathcal{B}z^\times} \dot{\mathbf{v}}_b^{zw/a} + \mathbf{J}_b^{\mathcal{B}z} \dot{\boldsymbol{\omega}}_b^{ba} - \mathbf{c}_b^{\mathcal{B}z^\times} \mathbf{v}_b^{zw/a^\times} \boldsymbol{\omega}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{J}_b^{\mathcal{B}z} \boldsymbol{\omega}_b^{ba} &= \mathbf{m}_b^{\mathcal{B}z},
\end{aligned}
$$

the coupled translational plus rotational equations of motion can be written concisely as

$$
\mathbf{M}_b^{\mathcal{B}z} \dot{\boldsymbol{\nu}}_b + \boldsymbol{\nu}_b^{\otimes} \mathbf{M}_b^{\mathcal{B}z} \boldsymbol{\nu}_b = \boldsymbol{f}_b^{\mathcal{B}z}, \tag{4.20}
$$

where

$$
\begin{aligned}
\mathbf{M}_b^{\mathcal{B}z} &= \begin{bmatrix} m_{\mathcal{B}} \mathbf{1} & -\mathbf{c}_b^{\mathcal{B}z^\times} \\ \mathbf{c}_b^{\mathcal{B}z^\times} & \mathbf{J}_b^{\mathcal{B}z} \end{bmatrix}, \\
\boldsymbol{\nu}_b &= \begin{bmatrix} \mathbf{v}_b^{zw/a} \\ \boldsymbol{\omega}_b^{ba} \end{bmatrix}, \\
\boldsymbol{\nu}_b^{\otimes} &= \begin{bmatrix} \boldsymbol{\omega}_b^{ba^\times} & \mathbf{0} \\ \mathbf{v}_b^{zw/a^\times} & \boldsymbol{\omega}_b^{ba^\times} \end{bmatrix}, \\
\boldsymbol{f}_b^{\mathcal{B}z} &= \begin{bmatrix} \mathbf{f}_b^{\mathcal{B}} \\ \mathbf{m}_b^{\mathcal{B}z} \end{bmatrix}.
\end{aligned}
$$

The complexity of the coupled translational and rotational equations of motion given in Equations (4.18) and (4.19), or the more concise form given in Equation (4.20), can be simplified collocating the point z with the centre of mass. When z collocated with the centre of mass $\mathbf{c}_b^{\mathcal{B}c} = \mathbf{0}$ and Equations (4.18) and (4.19) become

$$m_{\mathcal{B}}\left(\dot{\mathbf{v}}_b^{cw/a} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{v}_b^{cw/a}\right) = \mathbf{f}_b^{\mathcal{B}}, \tag{4.21}$$

$$\mathbf{J}_b^{\mathcal{B}c}\dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{J}_b^{\mathcal{B}c}\boldsymbol{\omega}_b^{ba} = \mathbf{m}_b^{\mathcal{B}c}. \tag{4.22}$$

Recalling that

$$\begin{aligned}
\underrightarrow{v}^{cw/a^{\bullet a}} &= \underrightarrow{v}^{cw/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{cw/a}, \\
\dot{\mathbf{v}}_a^{cw/a} &= \mathbf{C}_{ba}^{\mathsf{T}}\left(\dot{\mathbf{v}}_b^{cw/a} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{v}_b^{cw/a}\right),
\end{aligned}$$

Equation (4.21) can be written together with Equation (4.22) as

$$m_{\mathcal{B}}\dot{\mathbf{v}}_a^{cw/a} = \mathbf{f}_a^{\mathcal{B}}, \tag{4.23}$$

$$\mathbf{J}_b^{\mathcal{B}c}\dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{J}_b^{\mathcal{B}c}\boldsymbol{\omega}_b^{ba} = \mathbf{m}_b^{\mathcal{B}c}. \tag{4.24}$$

4.3.4 Kinetic and Potential Energy of a System of Particles

Definition 4.25 (Kinetic Energy of a System of Particles)**.** Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. The *kinetic energy of body $\mathcal{B}$ relative to w w.r.t. $\mathcal{F}_a$* is defined as [3, Sec. 8.4]

$$T_{\mathcal{B}w/a} = \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_iw/a} \cdot \underrightarrow{v}^{y_iw/a}. \tag{4.25}$$

Recalling that

$$\begin{aligned}
\underrightarrow{r}^{y_iw} &= \underrightarrow{r}^{y_iz} + \underrightarrow{r}^{zw}, \\
\underrightarrow{v}^{y_iw/a} &= \underrightarrow{v}^{y_iz/a} + \underrightarrow{v}^{zw/a},
\end{aligned}$$

Equation (4.25) can be written as [1, pp. 47]

$$\begin{aligned}
T_{\mathcal{B}w/a} &= \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_iw/a} \cdot \underrightarrow{v}^{y_iw/a} \\
&= \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \left(\underrightarrow{v}^{y_iz/a} + \underrightarrow{v}^{zw/a}\right) \cdot \left(\underrightarrow{v}^{y_iz/a} + \underrightarrow{v}^{zw/a}\right) \\
&= \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \left(\underrightarrow{v}^{y_iz/a} \cdot \underrightarrow{v}^{y_iz/a} + 2\underrightarrow{v}^{y_iz/a} \cdot \underrightarrow{v}^{zw/a} + \underrightarrow{v}^{zw/a} \cdot \underrightarrow{v}^{zw/a}\right) \\
&= \tfrac{1}{2}m_{\mathcal{B}} \underrightarrow{v}^{zw/a} \cdot \underrightarrow{v}^{zw/a} + \sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_iz/a} \cdot \underrightarrow{v}^{zw/a} + \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_iz/a} \cdot \underrightarrow{v}^{y_iz/a} \\
&= \tfrac{1}{2}m_{\mathcal{B}} \underrightarrow{v}^{zw/a} \cdot \underrightarrow{v}^{zw/a} + \underrightarrow{c}^{\mathcal{B}z^{\bullet a}} \cdot \underrightarrow{v}^{zw/a} + \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \underrightarrow{v}^{y_iz/a} \cdot \underrightarrow{v}^{y_iz/a}
\end{aligned}$$

where, from the definition of the first moment of mass given in Equation (4.5),

$$\underset{\rightarrow}{c}^{\mathcal{B}z^{\bullet a}} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i z^{\bullet a}} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i z/a}.$$

Definition 4.26 (Gravitational Potential Energy of a System of Particles)**.** Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. The *gravitational potential energy of body $\mathcal{B}$ relative to w* is defined as [3, Sec. 9.4]

$$V_{\mathcal{B}w} = \sum_{i=1}^{\ell} V_{y_i w}(\underset{\rightarrow}{r}^{y_i w}) \tag{4.26}$$

where

$$V_{y_i w}(\underset{\rightarrow}{r}^{y_i w}) = -m_i \underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{y_i w}.$$

Note that

$$\begin{aligned} V_{\mathcal{B}w} &= \sum_{i=1}^{\ell} V_{y_i w}(\underset{\rightarrow}{r}^{y_i w}) \\ &= \sum_{i=1}^{\ell} -m_i \underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{y_i w} \\ &= -\underset{\rightarrow}{g} \cdot \left(\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i w} \right) \\ &= -m_{\mathcal{B}} \underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{cw}. \end{aligned}$$

where $m_{\mathcal{B}} \underset{\rightarrow}{r}^{cw} = \sum_{i=1}^{\ell} m_i \underset{\rightarrow}{r}^{y_i w}$.

4.4 The Dynamics of Continuous-Rigid Bodies

In the preceding sections the dynamics of a single particle and then the dynamics of multi-particle systems (i.e., discrete bodies and discrete-rigid bodies) were discussed. These derivations are but stepping stones to the penultimate set of motion equations, that being the motion equations of a rigid body.

In the discrete particle case $\mathcal{B}$ composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully. The approach to deriving the motion equations of a continuous-rigid body, a continua, is to let the each particle y_i's mass, the m_i's, be a dm's, and let the summations in the definition of mass, first moment of mass, etc., be integrals.

4.4.1 Properties of Continuous-Rigid Bodies

Definition 4.27. A *continuous-rigid body* is a continuum in which the distance between any two points on the body is constant.

Definition 4.28. A *massive continuous-rigid body* is a continuum in which the distance between any two points on the body is constant, and the body is massive.

A massive continuous-rigid body is unaffected by forces and moments.

Definition 4.29. Let $\mathcal{F}_a$ be an inertial fame, and $\mathcal{F}_b$ be a body-fixed frame affixed to a massive continuous-rigid body. The massive continuous-rigid body is an *inertially nonrotating massive continuous-rigid body* if $\underset{\rightarrow}{\omega}^{ba} = \underset{\rightarrow}{0}$.

An inertially nonrotating massive continuous-rigid body is useful because any body-fixed frame associated with the massive continuous-rigid body is an inertial frame.

Definition 4.30. Consider a continuous-rigid body $\mathcal{B}$. The *zeroth moment of mass*, or the *mass*, of body $\mathcal{B}$ is defined as [1, pp. 56]

$$m_{\mathcal{B}} = \int_{\mathcal{B}} \mathrm{d}m = \int_{\mathcal{B}} \sigma(\underset{\rightarrow}{r}^{\mathrm{d}mz})\mathrm{d}V, \tag{4.27}$$

where $\mathrm{d}m = \sigma(\underset{\rightarrow}{r}^{\mathrm{d}mz})\mathrm{d}V$ is a mass element of $\mathcal{B}$, $\sigma(\cdot)$ is the density per unit volume, $\mathrm{d}V$ is a volume element, and z is a point within $\mathcal{B}$.

Definition 4.31. Consider a continuous-rigid body $\mathcal{B}$. Let $m_{\mathcal{B}}$ be the mass of $\mathcal{B}$ and z be a point within $\mathcal{B}$. The *centre of mass c of* $\mathcal{B}$ is defined as [1, pp. 56]

$$\underset{\rightarrow}{r}^{cz} = \frac{1}{m_{\mathcal{B}}} \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m. \tag{4.28}$$

Definition 4.32. Consider a continuous-rigid body $\mathcal{B}$ and a point z within $\mathcal{B}$. The *first moment of mass of body* $\mathcal{B}$ *relative to point* z is defined as

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m. \tag{4.29}$$

Using Equation (4.28) the first moment of mass definition (in Equation (4.29)) can be written

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m = m_{\mathcal{B}} \underset{\rightarrow}{r}^{cz}. \tag{4.30}$$

The second moment of mass, also called the moment of inertia or the mass-moment of inertia, will now be defined.

Definition 4.33. Consider a continuous-rigid body $\mathcal{B}$ and a point z within $\mathcal{B}$. The *second moment of mass of body* $\mathcal{B}$ *relative to point* z is defined as [1, pp. 56]

$$\underset{\rightarrow}{J}^{\mathcal{B}z} = \int_{\mathcal{B}} \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \mathrm{d}m, \tag{4.31}$$

where $\underset{\rightarrow}{1}$ is the identity tensor.

4.4.2 The Newton-Euler Approach for Continuous-Rigid Bodies

Definition 4.34. Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a continuous-rigid body. The *translational momentum of body* $\mathcal{B}$ *relative to* w *w.r.t.* $\mathcal{F}_a$ is defined as [1, pp. 56]

$$\underset{\rightarrow}{p}^{\mathcal{B}w/a} = \int_{\mathcal{B}} \mathrm{d}\,\underset{\rightarrow}{p}^{\mathrm{d}mw/a} = \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mw/a} \mathrm{d}m. \tag{4.32}$$

Using the definition of the centre of mass given in Equation (4.28),

$$\underset{\rightarrow}{r}^{cw} = \frac{1}{m_{\mathcal{B}}} \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mw} \mathrm{d}m,$$

the expression for $\underrightarrow{p}^{\mathcal{B}w/a}$ can alternatively be written as

$$\begin{aligned}
\underrightarrow{p}^{\mathcal{B}w/a} &= \int_{\mathcal{B}} \underrightarrow{v}^{\mathrm{d}mw/a}\mathrm{d}m \\
&= m_{\mathcal{B}} {\underrightarrow{r}^{cw}}^{\bullet a} \\
&= m_{\mathcal{B}} \underrightarrow{v}^{cw/a}. \qquad (4.33)
\end{aligned}$$

Axiom 4.5 (Euler's First Law (E1L))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a continuous-rigid body, let $\mathrm{d}\,\underrightarrow{f}^{\mathrm{d}m}$ be the external force per unit volume applied to the mass element $\mathrm{d}m = \sigma \mathrm{d}V$, and let σ be the mass per unit volume of $\mathcal{B}$. Then

$$ {\underrightarrow{p}^{\mathcal{B}w/a}}^{\bullet a} = m_{\mathcal{B}} {\underrightarrow{v}^{cw/a}}^{\bullet a} = \underrightarrow{f}^{\mathcal{B}}, \qquad (4.34)$$

where

$$\underrightarrow{f}^{\mathcal{B}} = \int_{\mathcal{B}} \mathrm{d}\,\underrightarrow{f}^{\mathrm{d}m}.$$

Definition 4.35. Let $\mathcal{F}_a$ be frame, let z be a point, and let $\mathcal{B}$ be a continuous-rigid body. The *angular momentum of body $\mathcal{B}$ relative to z w.r.t. $\mathcal{F}_a$* is defined as

$$\underrightarrow{h}^{\mathcal{B}z/a} = \int_{\mathcal{B}} \mathrm{d}\,\underrightarrow{h}^{\mathrm{d}mz/a} = \int_{\mathcal{B}} \underrightarrow{r}^{\mathrm{d}mz} \times \mathrm{d}\,\underrightarrow{p}^{\mathrm{d}mz/a} = \int_{\mathcal{B}} \underrightarrow{r}^{\mathrm{d}mz} \times \underrightarrow{v}^{\mathrm{d}mz/a}\mathrm{d}m.$$

Axiom 4.6 (Euler's Second Law (E2L))**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a continuous-rigid body, let $\mathrm{d}\,\underrightarrow{f}^{\mathrm{d}m}$ be the external force per unit volume applied to the mass element $\mathrm{d}m = \sigma \mathrm{d}V$, and let σ be the mass per unit volume of $\mathcal{B}$. Then

$$ {\underrightarrow{h}^{\mathcal{B}z/a}}^{\bullet a} + \underrightarrow{c}^{\mathcal{B}z} \times {\underrightarrow{v}^{zw/a}}^{\bullet a} = \underrightarrow{m}^{\mathcal{B}z}, \qquad (4.35)$$

where

$$\underrightarrow{m}^{\mathcal{B}z} = \int_{\mathcal{B}} \underrightarrow{r}^{\mathrm{d}mz} \times \mathrm{d}\,\underrightarrow{f}^{\mathrm{d}m}$$

is the moment on $\mathcal{B}$ relative to z. When point z is collocated with point c,

$$ {\underrightarrow{h}^{\mathcal{B}c/a}}^{\bullet a} = \underrightarrow{m}^{\mathcal{B}c}.$$

At this point it is worth pausing and discussing why Laws 4.5 and 4.6 are called Euler's first and second laws, respectively. Recall the proof of Theorem 4.3 N3L was used to show the double sum in Equation (4.10) equals zero. Similarly, in the proof of Theorem 4.4 N3L was again used to show that the a similar double sum equals zero. In the case of a continuous-rigid body one must make a further axiomatic assumptions over and above N3L about how the mass elements composing the continuous-rigid body interact. Euler realized this, and for this reason, abandon the multi-particle model and postulated his own laws, namely Euler's first and second laws [5, pp. 340-346]. It is for this reason that when driving the motion equations of a continuous-rigid body, or the motion equations of a system of rigid bodies and discrete particles, the approach is called a Newton-Euler approach, rather than just a N2L approach.

4.4.2.1 Alternative Forms

Equations (4.34) and (4.35) can be written in two alternative forms.

Theorem 4.6. Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a continuous-rigid body, let $\mathrm{d}\underset{\rightarrow}{f}^{\mathrm{d}m}$ be the external force per unit volume applied to the mass element $\mathrm{d}m = \sigma\mathrm{d}V$, and let σ be the mass per unit volume of $\mathcal{B}$. Then

$$m_{\mathcal{B}}\underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}}\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{c}^{\mathcal{B}z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{c}^{\mathcal{B}z}\right) = \underset{\rightarrow}{f}^{\mathcal{B}} \tag{4.36}$$

where $m_{\mathcal{B}}$ is the mass of the body $\mathcal{B}$ and

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz}\mathrm{d}m = m_{\mathcal{B}}\underset{\rightarrow}{r}^{cz}.$$

is the first moment of mass of body $\mathcal{B}$ relative to point z. Equation (4.36) is an alternative form of Equation (4.34).

Proof. Recall that $\underset{\rightarrow}{r}^{\mathrm{d}mw} = \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{r}^{zw}$ and therefore

$$\begin{aligned} \underset{\rightarrow}{r}^{\mathrm{d}mw} &= \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{r}^{zw}, \\ \underset{\rightarrow}{v}^{\mathrm{d}mw/a} &= \underset{\rightarrow}{v}^{\mathrm{d}mz/a} + \underset{\rightarrow}{v}^{zw/a}. \end{aligned} \tag{4.37}$$

Next, write $\underset{\rightarrow}{p}^{\mathcal{B}w/a} = \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mw/a}\mathrm{d}m$ as

$$\begin{aligned} \underset{\rightarrow}{p}^{\mathcal{B}w/a} &= \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mw/a}\mathrm{d}m \\ &= \int_{\mathcal{B}} \left(\underset{\rightarrow}{v}^{\mathrm{d}mz/a} + \underset{\rightarrow}{v}^{zw/a}\right)\mathrm{d}m \\ &= \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mz/a}\mathrm{d}m + m_{\mathcal{B}}\underset{\rightarrow}{v}^{zw/a}, \end{aligned} \tag{4.38}$$

where Equation (4.37) has been used. Taking the derivative w.r.t. $\mathcal{F}_a$ of Equation (4.38), recalling from Equation (4.34) that $\underset{\rightarrow}{p}^{\mathcal{B}w/a^{\bullet a}} = \underset{\rightarrow}{f}^{\mathcal{B}}$, and using

$$\begin{aligned} \underset{\rightarrow}{v}^{y_iz/a^{\bullet a}} &= \underset{\rightarrow}{v}^{y_iz/b^{\bullet b}} + 2\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_iz^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{y_iz} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_iz}\right) \\ &= \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{y_iz} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_iz}\right), \end{aligned}$$

where $\underset{\rightarrow}{v}^{y_iz/b^{\bullet b}}$ and $2\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{y_iz^{\bullet b}}$ equal $\underset{\rightarrow}{0}$ because the body is rigid, results in

$$\begin{aligned}
\underset{\rightarrow}{f}^{\mathcal{B}} &= \underset{\rightarrow}{p}^{\mathcal{B}w/a^{\bullet a}} \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet a}} + \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mz/a^{\bullet a}} \mathrm{d}m \\
&= m_{\mathcal{B}} \left(\underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} \right) + \int_{\mathcal{B}} \left(\underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \right) \mathrm{d}m \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \left(\int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m \right) + \underset{\rightarrow}{\omega}^{ba} \times \left[\underset{\rightarrow}{\omega}^{ba} \times \left(\int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m \right) \right] \\
&= m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{c}^{\mathcal{B}z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{c}^{\mathcal{B}z} \right).
\end{aligned}$$

□

Theorem 4.7. Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let z be a point, let $\mathcal{B}$ be a continuous-rigid body, let $\mathrm{d}\underset{\rightarrow}{f}^{\mathrm{d}m}$ be the external force per unit volume applied to the mass element $\mathrm{d}m = \sigma \mathrm{d}V$, and let σ be the mass per unit volume of $\mathcal{B}$. Then

$$\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba} \right) + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} \right) = \underset{\rightarrow}{m}^{\mathcal{B}z}, \quad (4.39)$$

where

$$\underset{\rightarrow}{c}^{\mathcal{B}z} = \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \mathrm{d}m = m_{\mathcal{B}} \underset{\rightarrow}{r}^{cz}.$$

is the first moment of mass of body $\mathcal{B}$ relative to point z and

$$\underset{\rightarrow}{J}^{\mathcal{B}z} = \int_{\mathcal{B}} \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \mathrm{d}m$$

is the second moment of mass of body $\mathcal{B}$ relative to point z. Equation (4.39) is an alternative form of Equation (4.35).

Proof. First, note that

$$\underset{\rightarrow}{v}^{\mathrm{d}mz/a} = \underset{\rightarrow}{r}^{\mathrm{d}mz^{\bullet a}} = \underset{\rightarrow}{r}^{\mathrm{d}mz^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} = \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz},$$

where $\underset{\rightarrow}{r}^{\mathrm{d}mz^{\bullet b}} = \underset{\rightarrow}{0}$ owing to the fact the body is a continuous-rigid body. Therefore,

$$\begin{aligned}
\underset{\rightarrow}{h}^{\mathcal{B}z/a} &= \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \times \underset{\rightarrow}{v}^{\mathrm{d}mz/a} \mathrm{d}m \\
&= \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \mathrm{d}m \\
&= -\int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mz} \times \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \times \underset{\rightarrow}{\omega}^{ba} \right) \mathrm{d}m.
\end{aligned}$$

Using the identity

$$-\underset{\rightarrow}{r}^{\mathrm{d}mz} \times \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \times \underset{\rightarrow}{\omega}^{ba} \right) = \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \cdot \underset{\rightarrow}{\omega}^{ba},$$

it follows that

$$\begin{aligned} \underset{\rightarrow}{h}^{\mathcal{B}z/a} &= \left(\int_{\mathcal{B}} \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{\mathrm{d}mz} \underset{\rightarrow}{r}^{\mathrm{d}mz}\right) \mathrm{d}m\right) \cdot \underset{\rightarrow}{\omega}^{ba} \\ &= \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}. \end{aligned}$$

Next, returning to

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a^{\bullet a}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{m}^{\mathcal{B}z},$$

using

$$\underset{\rightarrow}{h}^{\mathcal{B}z/a} = \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba} \quad \text{and} \quad \underset{\rightarrow}{v}^{zw/a^{\bullet a}} = \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}$$

results in

$$\begin{aligned} \underset{\rightarrow}{h}^{\mathcal{B}z/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{h}^{\mathcal{B}z/a} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}\right) &= \underset{\rightarrow}{m}^{\mathcal{B}z}, \\ \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}\right)^{\bullet b} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}\right) + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}\right) &= \underset{\rightarrow}{m}^{\mathcal{B}z}, \\ \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}\right) + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}\right) &= \underset{\rightarrow}{m}^{\mathcal{B}z}. \end{aligned}$$

□

4.4.2.2 Referential Form of Translational and Rotational Equations of Motion

Equations (4.36) and (4.39),

$$\begin{aligned} m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + m_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba^{\bullet b}} \times \underset{\rightarrow}{c}^{\mathcal{B}z} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{c}^{\mathcal{B}z}\right) &= \underset{\rightarrow}{f}^{\mathcal{B}}, \\ \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba^{\bullet b}} + \underset{\rightarrow}{\omega}^{ba} \times \left(\underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}\right) + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{v}^{zw/a^{\bullet b}} + \underset{\rightarrow}{c}^{\mathcal{B}z} \times \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{v}^{zw/a}\right) &= \underset{\rightarrow}{m}^{\mathcal{B}z}, \end{aligned}$$

can be written in referential form as

$$m_{\mathcal{B}} \dot{\mathbf{v}}_b^{zw/a} + m_{\mathcal{B}} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} + \dot{\boldsymbol{\omega}}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z} + \boldsymbol{\omega}_b^{ba^\times} \boldsymbol{\omega}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z} = \mathbf{f}_b^{\mathcal{B}}, \tag{4.40}$$

$$\mathbf{J}_b^{\mathcal{B}z} \dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{J}_b^{\mathcal{B}z} \boldsymbol{\omega}_b^{ba} + \mathbf{c}_b^{\mathcal{B}z^\times} \dot{\mathbf{v}}_b^{zw/a} + \mathbf{c}_b^{\mathcal{B}z^\times} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} = \mathbf{m}_b^{\mathcal{B}z}. \tag{4.41}$$

Rewriting the above as

$$\begin{aligned} m_{\mathcal{B}} \dot{\mathbf{v}}_b^{zw/a} - \mathbf{c}_b^{\mathcal{B}z^\times} \dot{\boldsymbol{\omega}}_b^{ba} + m_{\mathcal{B}} \boldsymbol{\omega}_b^{ba^\times} \mathbf{v}_b^{zw/a} - \boldsymbol{\omega}_b^{ba^\times} \mathbf{c}_b^{\mathcal{B}z^\times} \boldsymbol{\omega}_b^{ba} &= \mathbf{f}_b^{\mathcal{B}}, \\ \mathbf{c}_b^{\mathcal{B}z^\times} \dot{\mathbf{v}}_b^{zw/a} + \mathbf{J}_b^{\mathcal{B}z} \dot{\boldsymbol{\omega}}_b^{ba} - \mathbf{c}_b^{\mathcal{B}z^\times} \mathbf{v}_b^{zw/a^\times} \boldsymbol{\omega}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times} \mathbf{J}_b^{\mathcal{B}z} \boldsymbol{\omega}_b^{ba} &= \mathbf{m}_b^{\mathcal{B}z}, \end{aligned}$$

the coupled translational plus rotational equations of motion can be written concisely as

$$\mathbf{M}_b^{\mathcal{B}z} \dot{\boldsymbol{\nu}}_b + \boldsymbol{\nu}_b^{\otimes} \mathbf{M}_b^{\mathcal{B}z} \boldsymbol{\nu}_b = \boldsymbol{f}_b^{\mathcal{B}z}, \tag{4.42}$$

where

$$
\begin{aligned}
\mathbf{M}_b^{\mathcal{B}z} &= \begin{bmatrix} m_{\mathcal{B}}\mathbf{1} & -\mathbf{c}_b^{\mathcal{B}z^\times} \\ \mathbf{c}_b^{\mathcal{B}z^\times} & \mathbf{J}_b^{\mathcal{B}z} \end{bmatrix}, \\
\boldsymbol{\nu}_b &= \begin{bmatrix} \mathbf{v}_b^{zw/a} \\ \boldsymbol{\omega}_b^{ba} \end{bmatrix}, \\
\boldsymbol{\nu}_b^{\otimes} &= \begin{bmatrix} \boldsymbol{\omega}_b^{ba^\times} & \mathbf{0} \\ \mathbf{v}_b^{zw/a^\times} & \boldsymbol{\omega}_b^{ba^\times} \end{bmatrix}, \\
\boldsymbol{f}_b^{\mathcal{B}z} &= \begin{bmatrix} \mathbf{f}_b^{\mathcal{B}} \\ \mathbf{m}_b^{\mathcal{B}z} \end{bmatrix}.
\end{aligned}
$$

The complexity of the coupled translational and rotational equations of motion given in Equations (4.40) and (4.41), or the more concise form given in Equation (4.42), can be simplified collocating the point z with the centre of mass. When z collocated with the centre of mass $\mathbf{c}_b^{\mathcal{B}c} = \mathbf{0}$ and Equations (4.40) and (4.41) become

$$
m_{\mathcal{B}}\left(\dot{\mathbf{v}}_b^{cw/a} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{v}_b^{cw/a}\right) = \mathbf{f}_b^{\mathcal{B}}, \tag{4.43}
$$

$$
\mathbf{J}_b^{\mathcal{B}c}\dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{J}_b^{\mathcal{B}c}\boldsymbol{\omega}_b^{ba} = \mathbf{m}_b^{\mathcal{B}c}. \tag{4.44}
$$

Recalling that

$$
\begin{aligned}
\underrightarrow{v}^{cw/a^{\bullet a}} &= \underrightarrow{v}^{cw/a^{\bullet b}} + \underrightarrow{\omega}^{ba} \times \underrightarrow{v}^{cw/a}, \\
\dot{\mathbf{v}}_a^{cw/a} &= \mathbf{C}_{ba}^{\mathsf{T}}\left(\dot{\mathbf{v}}_b^{cw/a} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{v}_b^{cw/a}\right),
\end{aligned}
$$

Equation (4.43) can be written together with Equation (4.44) as

$$
m_{\mathcal{B}}\dot{\mathbf{v}}_a^{cw/a} = \mathbf{f}_a^{\mathcal{B}}, \tag{4.45}
$$

$$
\mathbf{J}_b^{\mathcal{B}c}\dot{\boldsymbol{\omega}}_b^{ba} + \boldsymbol{\omega}_b^{ba^\times}\mathbf{J}_b^{\mathcal{B}c}\boldsymbol{\omega}_b^{ba} = \mathbf{m}_b^{\mathcal{B}c}. \tag{4.46}
$$

4.4.2.3 The Positive Definite Nature of the Mass Matrix

A natural question to ask is if

$$
\mathbf{M}_b^{\mathcal{B}z} = \begin{bmatrix} m_{\mathcal{B}}\mathbf{1} & -\mathbf{c}_b^{\mathcal{B}z^\times} \\ \mathbf{c}_b^{\mathcal{B}z^\times} & \mathbf{J}_b^{\mathcal{B}z} \end{bmatrix}
$$

is symmetric and positive definite, that is, $\mathbf{M}_b^{\mathcal{B}z} = \mathbf{M}_b^{\mathcal{B}z^{\mathsf{T}}} > 0$, which is equivalent to $\mathbf{x}^{\mathsf{T}}\mathbf{M}_b^{\mathcal{B}z}\mathbf{x} > 0$, $\forall \mathbf{x} \in \mathbb{R}^6 \setminus \mathbf{x} = \mathbf{0}$.

To show that $\mathbf{M}_b^{\mathcal{B}z} = \mathbf{M}_b^{\mathcal{B}z^{\mathsf{T}}}$ simply transpose $\mathbf{M}_b^{\mathcal{B}z}$ and observe that it is equal to $\mathbf{M}_b^{\mathcal{B}z^{\mathsf{T}}}$. To show that $\mathbf{M}_b^{\mathcal{B}z}$ is positive definite, first note that [14, pp. 651]

$$
\mathbf{M}_b^{\mathcal{B}z} > 0 \quad \Leftrightarrow \quad m_{\mathcal{B}} > 0,\ \mathbf{J}_b^{\mathcal{B}z} + \frac{1}{m_{\mathcal{B}}}\mathbf{c}_b^{\mathcal{B}z^\times}\mathbf{c}_b^{\mathcal{B}z^\times} > 0. \tag{4.47}
$$

Using the parallel axis theorem,

$$
\mathbf{J}_b^{\mathcal{B}z} = \mathbf{J}_b^{\mathcal{B}c} - m_{\mathcal{B}}\mathbf{r}_b^{cz^\times}\mathbf{r}_b^{cz^\times},
$$

and the fact that $\mathbf{c}_b^{\mathcal{B}z} = m_{\mathcal{B}} \mathbf{r}_b^{cz}$, it follows that

$$\begin{aligned}
\mathbf{z}^{\mathsf{T}} \left(\mathbf{J}_b^{\mathcal{B}z} + \frac{1}{m_{\mathcal{B}}} \mathbf{c}_b^{\mathcal{B}z^\times} \mathbf{c}_b^{\mathcal{B}z^\times} \right) \mathbf{z} &= \mathbf{z}^{\mathsf{T}} \left(\mathbf{J}_b^{\mathcal{B}c} - m_{\mathcal{B}} \mathbf{r}_b^{cz^\times} \mathbf{r}_b^{cz^\times} + \frac{1}{m_{\mathcal{B}}} \mathbf{c}_b^{\mathcal{B}z^\times} \mathbf{c}_b^{\mathcal{B}z^\times} \right) \mathbf{z} \\
&= \mathbf{z}^{\mathsf{T}} \left(\mathbf{J}_b^{\mathcal{B}c} - m_{\mathcal{B}} \mathbf{r}_b^{cz^\times} \mathbf{r}_b^{cz^\times} + m_{\mathcal{B}} \mathbf{r}_b^{cz^\times} \mathbf{r}_b^{cz^\times} \right) \mathbf{z} \\
&= \mathbf{z}^{\mathsf{T}} \mathbf{J}_b^{\mathcal{B}c} \mathbf{z} > 0,
\end{aligned}$$

owing to the fact that $\mathbf{J}_b^{\mathcal{B}c} = \mathbf{J}_b^{\mathcal{B}c^{\mathsf{T}}} > 0$. As such, from the necessary and sufficient conditions given in Equation (4.47), it follows that $\mathbf{M}_b^{\mathcal{B}z} = \mathbf{M}_b^{\mathcal{B}z^{\mathsf{T}}} > 0$.

4.4.3 Kinetic and Potential Energy of a Continuous-Rigid Body

Definition 4.36 (Kinetic Energy of a Continuous-Rigid Body)**.** Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a continuous-rigid body. The *kinetic energy of body $\mathcal{B}$ relative to w w.r.t. $\mathcal{F}_a$* is defined as

$$T_{\mathcal{B}w/a} = \tfrac{1}{2} \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mw/a} \cdot \underset{\rightarrow}{v}^{\mathrm{d}mw/a} \mathrm{d}m. \tag{4.48}$$

Recalling that

$$\begin{aligned}
\underset{\rightarrow}{r}^{\mathrm{d}mw} &= \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{r}^{zw}, \\
\underset{\rightarrow}{v}^{\mathrm{d}mw/a} &= \underset{\rightarrow}{v}^{\mathrm{d}mz/a} + \underset{\rightarrow}{v}^{zw/a} \\
&= \underbrace{\underset{\rightarrow}{v}^{\mathrm{d}mz/b}}_{\underset{\rightarrow}{0}} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{v}^{zw/a} \\
&= \underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz},
\end{aligned}$$

where z is a point, Equation (4.48) can be written as

$$\begin{aligned}
T_{\mathcal{B}w/a} &= \tfrac{1}{2} \int_{\mathcal{B}} \underset{\rightarrow}{v}^{\mathrm{d}mw/a} \cdot \underset{\rightarrow}{v}^{\mathrm{d}mw/a} \mathrm{d}m \\
&= \tfrac{1}{2} \int_{\mathcal{B}} \left(\underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \cdot \left(\underset{\rightarrow}{v}^{zw/a} + \underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \mathrm{d}m \\
&= \tfrac{1}{2} \int_{\mathcal{B}} \left(\underset{\rightarrow}{v}^{zw/a} \cdot \underset{\rightarrow}{v}^{zw/a} + 2 \underset{\rightarrow}{v}^{zw/a} \cdot \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) + \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \cdot \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz} \right) \right) \mathrm{d}m \\
&= \tfrac{1}{2} m_{\mathcal{B}} \underset{\rightarrow}{v}^{zw/a} \cdot \underset{\rightarrow}{v}^{zw/a} - \underset{\rightarrow}{v}^{zw/a} \cdot \left(\underset{\rightarrow}{c}^{\mathcal{B}z} \times \underset{\rightarrow}{\omega}^{ba} \right) + \tfrac{1}{2} \underset{\rightarrow}{\omega}^{ba} \cdot \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba},
\end{aligned}$$

where

$$
\begin{aligned}
& \tfrac{1}{2}\int_{\mathcal{B}} \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz}\right) \cdot \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz}\right) \mathrm{d}m \\
= \; & -\tfrac{1}{2}\int_{\mathcal{B}} \left(\underset{\rightarrow}{\omega}^{ba} \times \underset{\rightarrow}{r}^{\mathrm{d}mz}\right) \cdot \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} \times \underset{\rightarrow}{\omega}^{ba}\right) \mathrm{d}m \\
= \; & \tfrac{1}{2}\int_{\mathcal{B}} \left((\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz})(\underset{\rightarrow}{\omega}^{ba} \cdot \underset{\rightarrow}{\omega}^{ba}) - (\underset{\rightarrow}{\omega}^{ba} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz})(\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{\omega}^{ba})\right) \mathrm{d}m \\
= \; & \tfrac{1}{2}\int_{\mathcal{B}} \underset{\rightarrow}{\omega}^{ba} \cdot \left((\underset{\rightarrow}{r}^{\mathrm{d}mz} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz})\underset{\rightarrow}{1} - \underset{\rightarrow}{r}^{\mathrm{d}mz}\underset{\rightarrow}{r}^{\mathrm{d}mz}\right) \cdot \underset{\rightarrow}{\omega}^{ba} \mathrm{d}m \\
= \; & \tfrac{1}{2}\underset{\rightarrow}{\omega}^{ba} \cdot \underset{\rightarrow}{J}^{\mathcal{B}z} \cdot \underset{\rightarrow}{\omega}^{ba}
\end{aligned}
$$

Definition 4.37 (Gravitational Potential Energy of a Continuous-Rigid Body)**.** Let $\mathcal{F}_a$ be frame, let w be a point, and let $\mathcal{B}$ be a continuous-rigid body. The *gravitational potential energy of body $\mathcal{B}$ relative to w* is defined as [3, Sec. 9.4]

$$
V_{\mathcal{B}w} = \int_{\mathcal{B}} \mathrm{d}V(\underset{\rightarrow}{r}^{\mathrm{d}mw}) \tag{4.49}
$$

where

$$
\mathrm{d}V(\underset{\rightarrow}{r}^{\mathrm{d}mw}) = -\underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mw}\mathrm{d}m.
$$

Letting $\underset{\rightarrow}{r}^{\mathrm{d}mw} = \underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{r}^{zw}$ where z is a point, from Equation (4.49) notice that

$$
\begin{aligned}
V_{\mathcal{B}w} &= \int_{\mathcal{B}} \mathrm{d}V(\underset{\rightarrow}{r}^{\mathrm{d}mw}) \\
&= \int_{\mathcal{B}} \left(-\underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mw}\mathrm{d}m\right) \\
&= -\int_{\mathcal{B}} \underset{\rightarrow}{g} \cdot \left(\underset{\rightarrow}{r}^{\mathrm{d}mz} + \underset{\rightarrow}{r}^{zw}\right) \mathrm{d}m \\
&= -\int_{\mathcal{B}} \underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{\mathrm{d}mz}\mathrm{d}m - \int_{\mathcal{B}} \underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{zw}\mathrm{d}m \\
&= -\underset{\rightarrow}{g} \cdot \underset{\rightarrow}{c}^{\mathcal{B}z} - m_{\mathcal{B}}\underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{zw},
\end{aligned}
$$

where $m_{\mathcal{B}} = \int_{\mathcal{B}} \mathrm{d}m$ and $\underset{\rightarrow}{c}^{\mathcal{B}z} = \int_{\mathcal{B}} \underset{\rightarrow}{r}^{\mathrm{d}mw}\mathrm{d}m$. When point z is taken to be the centre of mass of $\mathcal{B}$ then

$$
V_{\mathcal{B}w} = -m_{\mathcal{B}}\underset{\rightarrow}{g} \cdot \underset{\rightarrow}{r}^{zw}.
$$

Chapter 5

Lagrangian Approach to the Dynamics of Particles and Rigid Bodies

Unless clarity is required, the time-dependence of a physical vector of scalar will not be written explicitly.

5.1 Generalized Coordinates, Degrees of Freedom, and Constraints

5.1.1 Generalized Coordinates and Degrees of Freedom

Definition 5.1. Let w be a point, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, and let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w. The *generalized coordinates* $q_1, q_2, \ldots, q_n$ are the $n \leq 3\ell$ coordinates used to parameterize $\underset{\rightarrow}{r}^{y_1 w}, \underset{\rightarrow}{r}^{y_2 w}, \ldots, \underset{\rightarrow}{r}^{y_\ell w}$ [2] [39, pp. 34-35] [40, pp. 216].

Note that the generalized coordinates are a function of time.

To solidify ideas, consider a frame $\mathcal{F}_a$, a point w, and the particle y. The position of y relative to w is given by $\underset{\rightarrow}{r}^{yw}$. Resolving $\underset{\rightarrow}{r}^{yw}$ in $\mathcal{F}_a$ results in

$$\underset{\rightarrow}{r}^{yw} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \mathbf{r}_a^{yw} = \underset{\rightarrow}{\boldsymbol{\mathcal{F}}}_a^{\mathsf{T}} \begin{bmatrix} r_{a1}^{yw} \\ r_{a1}^{yw} \\ r_{a2}^{yw} \end{bmatrix}.$$

If the components $r_{a1}^{yw}, r_{a2}^{yw}, r_{a3}^{yw}$ are parameterized using cartesian coordinates, that is $r_{a1}^{yw} = x_a$, $r_{a2}^{yw} = y_a$, $r_{a3}^{yw} = z_a$, then the generalized coordinates are $q_1 = x_a$, $q_2 = y_a$, $q_3 = z_a$. If $r_{a1}^{yw}, r_{a2}^{yw}, r_{a3}^{yw}$ are parameterized using spherical coordinates, that is $r_{a1}^{yw} = r\cos\phi\sin\theta$, $r_{a2}^{yw} = r\sin\phi\sin\theta$, $r_{a3}^{yw} = r\cos\theta$, then the generalized coordinates are $q_1 = \phi$, $q_2 = \theta$, $q_3 = r$.

Again, consider a frame $\mathcal{F}_a$, a point w, and the particle y. The position of y relative to w is given by $\underset{\rightarrow}{r}^{yw}$. Now consider the constraint $\underset{\rightarrow}{r}^{yw} \cdot \underset{\rightarrow}{r}^{yw} = \ell^2$ where ℓ is a constant. If $\underset{\rightarrow}{r}^{yw}$ is parameterized using cartesian coordinates and the generalize coordinates are chosen to be $q_1 = x_a$, $q_2 = y_a$, $q_3 = z_a$ then the generalized coordinates must satisfy $x_a^2 + y_a^2 + z_a^2 = \ell^2$ at all times. In this case, although there are three generalized coordinates, owing to the presence of the constraint, the generalized coordinates are *dependent*. Alternatively, parameterizing $\underset{\rightarrow}{r}^{yw}$ using spherical coordinate and choosing $q_1 = \phi$ and $q_3 = \theta$ as the generalized coordinates, the position of the particle is fully described by $q_1 = \phi$ and $q_3 = \theta$ via $r_{a1}^{yw} = \ell\cos\phi\sin\theta$, $r_{a2}^{yw} = \ell\sin\phi\sin\theta$, $r_{a3}^{yw} = \ell\cos\theta$ but the generalized coordinates are not constrained in anyway. These generalized coordinates are *independent*. Note that it is not always possible to choose a set of generalized coordinates that are independent.

5.1.2 Constraints

Different types of geometric constraints will now be discussed.

Definition 5.2 (Holonomic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Constraints of the form [2] [39, pp. 35-36]

$$\phi_j(\mathbf{q}(t), t) = 0, \quad j = 1, 2, \ldots, m,$$

are *holonomic constraints*.

Definition 5.3 (Nonholonomic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Constraints of the form [2] [39, pp. 36-37]

$$\psi_j(\mathbf{q}(t), \dot{\mathbf{q}}, t) = 0, \quad j = 1, 2, \ldots, m,$$

are *nonholonomic constraints*.

Definition 5.4 (Scleronomic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Constraints that are independent of time are *scleronomic constraints* [2] [39, pp. 37].

Definition 5.5 (Rheonomic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Constraints that are explicitly time dependent are *rheonomic constraints* [2] [39, pp. 37].

A constraint of the form $\phi(\mathbf{q}(t)) = 0$ is both holonomic and scleronomic, while a constraint of the form $\phi(\mathbf{q}(t), t) = 0$ is both holonomic and rheonomic [2].

Definition 5.6 (Catastatic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Nonholonomic constraints that are independent of time are *catastatic constraints* [2] [39, pp. 37].

Definition 5.7 (Acatastatic Constraints)**.** Let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. Nonholonomic constraints that are dependent on time are *acatastatic constraints* [2] [39, pp. 37].

Examples

Consider a frame $\mathcal{F}_a$, a point w, and a particle y. Let the generalized coordinates describing the particle be x_a, y_a, and z_a. In addition, suppose that $x_a^2 + y_a^2 + z_a^2 - \ell^2 = 0$ where ℓ is constant; this is a holonomic constraint (and, in addition, scleronomic) Note, by differentiating the constraint, in rate form the constraint is

$$\begin{aligned} 0 &= 2x_a\dot{x}_a + 2y_a\dot{y}_a + 2z_a\dot{z}_a \\ &= \begin{bmatrix} 2x_a & 2y_a & 2z_a \end{bmatrix} \begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{z}_a \end{bmatrix}. \end{aligned}$$

Again consider a frame $\mathcal{F}_a$, a point w, and a particle y. Suppose a frame $\mathcal{F}_b$ is attached to the particle enabling a discussion of attitude. (Because a particle has no "size", but just mass, a frame cannot really be attached to a particle. But, for the purposes of discussion, assume the particle is actually a sphere of finite volume, so that a frame can be affixed to it.) Let the generalized coordinates describing the particle be x_a, y_a, and θ where $\mathbf{C}_{ba} = \mathbf{C}_3(\theta)$ (and, furthermore, let $z_a = 0$). Assume the particle is constrained to move like a unicycle, that is, it is constrained such that

$$\dot{x}_a(t) = v_b(t)\cos(\theta(t)), \quad \dot{y}_a(t) = v_b(t)\sin(\theta(t)),$$

where $v_b \neq 0$ is the forward velocity along the $\underset{\rightarrow}{b}^1$ physical vector. Eliminating v_b gives

$$\frac{\dot{x}_a}{\dot{y}_a} = \frac{\cos(\theta)}{\sin(\theta)},$$

or

$$0 = \sin(\theta)\dot{x}_a - \cos(\theta)\dot{y}_a = \begin{bmatrix} \sin(\theta) & -\cos(\theta) & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_a \\ \dot{y}_a \\ \dot{\theta} \end{bmatrix}.$$

This is a nonholonomic constraint (but also a catastatic constraint). Note that in this example the generalized coordinates are not subject to any holonomic constraint. However, the time-rate-of-change of the generalized coordinates are subject to a nonholnomic constraint; the nonholonomic constraint does not reduce the degree of freedom of the generalized coordinates nor make them dependent, but the nonholonomic constraint does reduce the degrees of freedom of the time-rate-of-change of the generalized coordinates meaning that there are $n - m = 3 - 1 = 2$ degrees of freedom in the time-rate-of-change of the generalized coordinates [40, pp. 226].

Holonomic systems can sometimes be simplified by eliminating the surplus of generalized coordinates from the formulation. As an example, consider a point w, a frame $\mathcal{F}_a$, a particle y, and let the generalized coordinates describing the particle be x_a, y_a, and z_a subject to $x_a^2 + y_a^2 + z_a^2 - \ell^2 = 0$. Fundamentally there is $n - m = 3 - 1 = 2$ degrees of freedom. In particular, the configuration of the system is described by x_a and y_a because if one assumed z_a to be either positive of negative. As such, it is possible, although not always easy, to eliminate a holonomic constraint. However, the same cannot be said of a nonholonomic constraint.

5.1.3 The Pfaffian Form

Consider a holonomic constraint $\phi_j(\mathbf{q}(t), t) = 0$ where $\mathbf{q}(t) = [q_1(t) \quad q_2(t) \quad \cdots \quad q_n(t)]^\mathsf{T}$. The exact differential of the constraint $\phi_j(\mathbf{q}(t), t) = 0$ is

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t), t)\mathrm{d}q_k + \Xi_{jt}(\mathbf{q}(t), t)\mathrm{d}t = 0, \tag{5.1}$$

where

$$\Xi_{jk}(\mathbf{q}(t), t) = \frac{\partial \phi_j(\mathbf{q}(t), t)}{\partial q_k}, \quad \Xi_{jt}(\mathbf{q}(t), t) = \frac{\partial \phi_j(\mathbf{q}(t), t)}{\partial t}.$$

Equation (5.1) is referred to to as Pfaffian form. In rate form the constraint is

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t), t)\dot{q}_k(t) + \Xi_{jt}(\mathbf{q}(t), t) = 0. \tag{5.2}$$

The matrix form Equation (5.2) is

$$\boldsymbol{\Xi}_j(\mathbf{q}(t), t)\dot{\mathbf{q}}(t) + \Xi_{jt}(\mathbf{q}(t), t) = 0,$$

where

$$\boldsymbol{\Xi}_j(\mathbf{q}(t), t) = \begin{bmatrix} \Xi_{j1}(\mathbf{q}(t), t) & \Xi_{j2}(\mathbf{q}(t), t) & \cdots & \Xi_{jn}(\mathbf{q}(t), t) \end{bmatrix}.$$

Combining all $j = 1, 2, \ldots, m$ constraints together

$$\Xi(\mathbf{q}(t), t)\dot{\mathbf{q}}(t) + \Xi_t(\mathbf{q}(t), t) = \mathbf{0},$$

where

$$\Xi(\mathbf{q}(t), t) = \underset{\substack{j=1,2,\ldots,m \\ k=1,2,\ldots,n}}{\text{matrix}} \{\Xi_{jk}(\mathbf{q}(t), t)\}, \quad \Xi_t(\mathbf{q}(t), t) = \underset{j=1,2,\ldots,m}{\text{col}} \{\Xi_{jt}(\mathbf{q}(t), t)\}.$$

Notice that the definitions of $\Xi(\cdot, \cdot)$ and $\Xi_t(\cdot, \cdot)$ are consistent with the way a Jacobian is defined.

In forthcoming developments only nonholonomic constraints that can be cast in Pfaffian form will be considered.

5.2 Work, Virtual Work, Generalized Forces and Moments, and D'Alembert's Principle

5.2.1 Work

Definition 5.8. Let w be an unforced particle, let y be a particle whose mass is m, and let $\underset{\rightarrow}{f}^{\,y}$ be the force applied to particle y as y moves along the path $\mathcal{C}_y$. Then *the work done on y relative to w by $\underset{\rightarrow}{f}^{\,y}$ as particle y moves along the path $\mathcal{C}_y$* is [3, 9.2] [39, pp. 42]

$$\begin{aligned} W_{yw}(\underset{\rightarrow}{f}^{\,y}, \mathcal{C}_y) &= \int_{C_y} \underset{\rightarrow}{f}^{\,y} \cdot \mathrm{d}\,\underset{\rightarrow}{r}^{\,yw} \\ &= \int_{C_y} \mathrm{d}W_{yw}(\underset{\rightarrow}{f}^{\,y}, \mathrm{d}\,\underset{\rightarrow}{r}^{\,yw}). \end{aligned} \tag{5.3}$$

Recall from N2L that $\underset{\rightarrow}{p}^{\,yw/a^{\bullet a}} = \underset{\rightarrow}{f}^{\,y}$, where w is an unforced particle. The reason Definition 5.8 involves an unforced particle rather than a point is due to the fact that the force $\underset{\rightarrow}{f}^{\,y}$ cause a change in the momentum of y relative to the unforced particle w w.r.t. $\mathcal{F}_a$.

Notice from Equation (5.3) that work is not dependent on a particular coordinate system.

Definition 5.9. Let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and let $\underset{\rightarrow}{f}^{\,y_i}$ be the force applied to particle y_i as y_i moves along the math $\mathcal{C}_{y_i}$. Then *the work done on $\mathcal{B}$ relative to w by $\underset{\rightarrow}{f}^{\,y_i}$ as particle y_i moves along the path $\mathcal{C}_{y_i}$* is [3, 9.2]

$$W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{\,y_1}, \underset{\rightarrow}{f}^{\,y_2}, \ldots, \underset{\rightarrow}{f}^{\,y_\ell}, \mathcal{C}_{y_1}, \mathcal{C}_{y_2}, \ldots, \mathcal{C}_{y_\ell}) = \sum_{i=1}^{\ell} W_{y_i w}(\underset{\rightarrow}{f}^{\,y_i}, \mathcal{C}_{y_i}). \tag{5.4}$$

5.2.2 Virtual Work

Virtual displacements and virtual work will now be discussed.

Definition 5.10. Let w be point, and let y be a particle whose mass is m. Then *the virtual displacement of y relative to w* is [2] [39, pp. 40-46] [41, pp. 246-247]

$$\delta\,\underset{\rightarrow}{r}^{\,yw} \tag{5.5}$$

A virtual displacement is an imagined arbitrarily small displacement that is consistent with geometric constraints of the particle. Virtual displacements are assumed to occur without the passage of time, as if time were "frozen". Virtual displacements are denoted $\delta \underset{\rightarrow}{r}^{yw}$ as opposed to $\mathrm{d}\underset{\rightarrow}{r}^{yw}$, which is a real arbitrarily small displacement that occurs over $\mathrm{d}t$.

When it is said that $\delta \underset{\rightarrow}{r}^{yw}$ is consistent with geometric constraints, or alternatively that $\delta \underset{\rightarrow}{r}^{yw}$ is geometrically admissible, it means that that the virtual displacement does not violate any geometric constraints. For example, if the particle y is constrained to move in a plane then the virtual displacement $\delta \underset{\rightarrow}{r}^{yw}$ is also restricted to be in the plane.

Definition 5.11. Let w be an unforced particle, let y be a particle whose mass is m, and let $\underset{\rightarrow}{f}^{y}$ be the force applied to particle y. Then *the virtual work done on y relative to w by $\underset{\rightarrow}{f}^{y}$ through $\delta \underset{\rightarrow}{r}^{yw}$* is [2] [39, pp. 40-46] [41, pp. 246-247]

$$\delta W_{yw}(\underset{\rightarrow}{f}^{y}, \delta \underset{\rightarrow}{r}^{yw}) = \underset{\rightarrow}{f}^{y} \cdot \delta \underset{\rightarrow}{r}^{yw}. \tag{5.6}$$

Definition 5.12. Let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, and let $\underset{\rightarrow}{f}^{y_i}$ be the force applied to particle y_i. Then *the virtual work done on $\mathcal{B}$ relative to w by $\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}$ through $\delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}$* is [2] [41, pp. 246-247]

$$\delta W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}, \delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}) = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i} \cdot \delta \underset{\rightarrow}{r}^{y_i w}. \tag{5.7}$$

5.2.3 A Comment on δ and the Presence of Constraints

The notation δ does not just represent a virtual quantity, but rather δ represents a linear operator. In particular, δ is called the *variational operator*.

The variational operator δ acts much like the differential operator d [2]. Consider $\phi(u(\cdot), v(\cdot), \cdot)$ and recall that the differential operator applied to $\phi(u(\cdot), v(\cdot), \cdot)$ yields

$$\mathrm{d}\phi(u(t), v(t), t) = \frac{\partial \phi(u(t), v(t), t)}{\partial u(t)}\mathrm{d}u + \frac{\partial \phi(u(t), v(t), t)}{\partial v(t)}\mathrm{d}v + \frac{\partial \phi(u(t), v(t), t)}{\partial t}\mathrm{d}t.$$

Like the differential operator, the variational operator applied to $\phi(u(\cdot), v(\cdot), \cdot)$ yields

$$\delta\phi(u(t), v(t), t) = \frac{\partial \phi(u(t), v(t), t)}{\partial u(t)}\delta u + \frac{\partial \phi(u(t), v(t), t)}{\partial v(t)}\delta v. \tag{5.8}$$

However, while d recognizes changes in time, δ does not, which is why a term of the form $\frac{\partial \phi(u(t), v(t), t)}{\partial t}$ is not included in Equation (5.8). Quantities such as δu and δv do not depend on time. Moreover, because the variational operator does not recognize changes in time, time differentiation and the variational operator can be interchanged, that is [40, pp. 231]

$$\delta \dot{q}_k(t) = \delta \frac{\mathrm{d}q_k(t)}{\mathrm{d}t} = \frac{\mathrm{d}(\delta q_k)}{\mathrm{d}t}, \quad k = 1, 2, \ldots, n,$$

where $\dot{q}_k$ is a generalized coordinate.

Now virtual displacements in the presence of geometric constraints will be discussed. Let w be a point, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, and let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^\mathsf{T}$ be the generalized coordinates. In the absence of any sort of geometric constraint the virtual displacement $\delta \underset{\rightarrow}{r}^{y_i w}$ can be written [2] [39, pp. 41]

$$\delta \underset{\rightarrow}{r}^{y_i w} = \sum_{k=1}^{n} \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k} \delta q_k$$

In the presence of m geometric constraints, either holonomic or nonholonomic, of the form

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t), t)\mathrm{d}q_k + \Xi_{jt}(\mathbf{q}(t), t)\mathrm{d}t = 0,$$

the δq_k's must satisfy the constraint equations instantaneously, that is [2] [39, pp. 41]

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t), t)\delta q_k = 0.$$

In the case of holonomic constraints of the form $\phi_j(\mathbf{q}(t), t) = 0$ this can be derived directly:

$$\begin{aligned} \delta\phi_j(\mathbf{q}(t), t) &= \sum_{k=1}^{n} \underbrace{\frac{\partial \phi_j(\mathbf{q}(t), t)}{\partial q_k}}_{\Xi_{jk}(\mathbf{q}(t),t)} \delta q_k \\ &= \sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t), t)\delta q_k \\ &= 0 \end{aligned}$$

Note that there's no "Ξ_{jt}" term because the variational operator ignores time. In matrix form

$$\boldsymbol{\Xi}_j(\mathbf{q}(t), t)\delta\mathbf{q} = 0.$$

This means $\delta\mathbf{q}$ lies in the null space of $\boldsymbol{\Xi}_j(\cdot, \cdot)$. For all m constraints it follows that

$$\boldsymbol{\Xi}(\mathbf{q}(t), t)\delta\mathbf{q} = \mathbf{0},$$

where $\delta\mathbf{q}$ lies in the null space of $\boldsymbol{\Xi}(\cdot, \cdot)$. Given that $m < n$, rank $\boldsymbol{\Xi}(\cdot, \cdot) = m$.

5.2.4 Workless Constraints

Theorem 5.1. Let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, let $\underset{\rightarrow}{f}^{y_i}$ be the force applied to particle y_i, and let $\phi_i(\underset{\rightarrow}{r}^{y_i w}) = 0$ be a holonomic constraint associated with y_i. Write $\underset{\rightarrow}{f}^{y_i}$ as $\underset{\rightarrow}{f}^{y_i} = \underset{\rightarrow}{f}^{y_i\,\mathrm{ex}} + \underset{\rightarrow}{f}^{y_i\,\mathrm{hcon}} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_i y_j}$ where $\underset{\rightarrow}{f}^{y_i\,\mathrm{ex}}$ is the external force applied to particle y_i and $\underset{\rightarrow}{f}^{y_i\,\mathrm{hcon}}$ is the constraint force applied to particle y_i due to the presence of holonomic constraints, and $\underset{\rightarrow}{f}^{y_i y_j}$ is the force on y_i due to y_j (which could stem from holonomic constraints also). Then [40, pp. 233] [20, pp.211-213]

$$\delta W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{y_1\,\mathrm{ex}}, \underset{\rightarrow}{f}^{y_2\,\mathrm{ex}}, \ldots, \underset{\rightarrow}{f}^{y_\ell\,\mathrm{ex}}, \delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}) = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \mathrm{ex}} \cdot \delta \underset{\rightarrow}{r}^{y_i w}, \tag{5.9}$$

where the virtual displacements $\delta \underset{\rightarrow}{r}^{y_i w}$ are compatible with the geometric constraints.

Proof. First consider the holonomic constraint $\phi_i(\underrightarrow{r}^{y_i w}) = 0$. The virtual displacement $\delta \underrightarrow{r}^{y_i w}$ must satisfy

$$\underrightarrow{\nabla}\phi_i(\underrightarrow{r}^{y_i w}) \cdot \delta \underrightarrow{r}^{y_i w} = 0.$$

In addition, any constraint force $\underrightarrow{f}^{y_i\,\text{hcon}}$ are proportional to, and thus parallel to, $\underrightarrow{\nabla}\phi_i(\underrightarrow{r}^{y_i w})$, that is

$$\underrightarrow{f}^{y_i\,\text{hcon}} = \lambda_i \underrightarrow{\nabla}\phi_i(\underrightarrow{r}^{y_i w}), \quad \lambda_i \in \mathbb{R}.$$

As such, $\underrightarrow{f}^{y_i\,\text{hcon}} \cdot \delta \underrightarrow{r}^{y_i w} = 0$.

Next, using Equation (5.6), the virtual work done on y_i relative to w by $\underrightarrow{f}^{y_i}$ through $\delta \underrightarrow{r}^{y_i w}$ is

$$\begin{aligned} \delta W_{y_i w}(\underrightarrow{f}^{y_i}, \delta \underrightarrow{r}^{y_i w}) &= \underrightarrow{f}^{y_i} \cdot \delta \underrightarrow{r}^{y_i w} \\ &= \underrightarrow{f}^{y_i\,\text{ex}} \cdot \delta \underrightarrow{r}^{y_i w} + \underrightarrow{f}^{y_i\,\text{hcon}} \cdot \delta \underrightarrow{r}^{y_i w} + \sum_{j=1}^{\ell} \underrightarrow{f}^{y_i y_j} \cdot \delta \underrightarrow{r}^{y_i w} \\ &= \underrightarrow{f}^{y_i\,\text{ex}} \cdot \delta \underrightarrow{r}^{y_i w} + \sum_{j=1}^{\ell} \underrightarrow{f}^{y_i y_j} \cdot \delta \underrightarrow{r}^{y_i w}, \end{aligned}$$

where $\underrightarrow{f}^{y_i\,\text{hcon}} \cdot \delta \underrightarrow{r}^{y_i w} = 0$ has been used to simplify. Now, the virtual work done on $\mathcal{B}$ relative to w by $\underrightarrow{f}^{y_1}, \underrightarrow{f}^{y_2}, \ldots, \underrightarrow{f}^{y_\ell}$ through $\delta \underrightarrow{r}^{y_1 w}, \delta \underrightarrow{r}^{y_2 w}, \ldots, \delta \underrightarrow{r}^{y_\ell w}$ is

$$\begin{aligned} &\delta W_{\mathcal{B}w}(\underrightarrow{f}^{y_1}, \underrightarrow{f}^{y_2}, \ldots, \underrightarrow{f}^{y_\ell}, \delta \underrightarrow{r}^{y_1 w}, \delta \underrightarrow{r}^{y_2 w}, \ldots, \delta \underrightarrow{r}^{y_\ell w}) \\ = &\sum_{i=1}^{\ell} \delta W_{y_i w}(\underrightarrow{f}^{y_i}, \delta \underrightarrow{r}^{y_i w}) \\ = &\sum_{i=1}^{\ell} \underrightarrow{f}^{y_i\,\text{ex}} \cdot \delta \underrightarrow{r}^{y_i w} + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \underrightarrow{f}^{y_i y_j} \cdot \delta \underrightarrow{r}^{y_i w} \\ = &\sum_{i=1}^{\ell} \underrightarrow{f}^{y_i\,\text{ex}} \cdot \delta \underrightarrow{r}^{y_i w} + \sum_{i=1}^{\ell-1} \sum_{j=i+1}^{\ell} \underrightarrow{f}^{y_i y_j} \cdot \left(\delta \underrightarrow{r}^{y_i w} - \delta \underrightarrow{r}^{y_j w}\right), \end{aligned}$$

Note that the force $\underrightarrow{f}^{y_i y_j}$ is parallel to $\underrightarrow{r}^{y_i w} - \underrightarrow{r}^{y_j w} = \underrightarrow{r}^{y_i y_j}$, and as such $\underrightarrow{f}^{y_i y_j} \cdot \delta \underrightarrow{r}^{y_i y_j} = 0$ where $\delta \underrightarrow{r}^{y_i w} - \delta \underrightarrow{r}^{y_j w} = \delta \underrightarrow{r}^{y_i y_j}$. Therefore,

$$\delta W_{\mathcal{B}w}(\underrightarrow{f}^{y_1\,\text{ex}}, \underrightarrow{f}^{y_2\,\text{ex}}, \ldots, \underrightarrow{f}^{y_\ell\,\text{ex}}, \delta \underrightarrow{r}^{y_1 w}, \delta \underrightarrow{r}^{y_2 w}, \ldots, \delta \underrightarrow{r}^{y_\ell w}) = \sum_{i=1}^{\ell} \underrightarrow{f}^{y_i\,\text{ex}} \cdot \delta \underrightarrow{r}^{y_i w}.$$

□

The significance of Theorem 5.1 is that the constraint forces stemming from holonomic constraints, $\underrightarrow{f}^{y_i\,\text{hcon}}$, along with the internal forces, $\underrightarrow{f}^{y_i y_j}$, (which may stem from holonomic constraints also) do not contribute to the virtual work on $\mathcal{B}$ relative to w by $\underrightarrow{f}^{y_1}, \underrightarrow{f}^{y_2}, \ldots, \underrightarrow{f}^{y_\ell}$ through $\delta \underrightarrow{r}^{y_1 w}, \delta \underrightarrow{r}^{y_2 w}, \ldots, \delta \underrightarrow{r}^{y_\ell w}$. Such constraint forces are called workless constraint forces, and the constraints overall are called workless constraints.

5.2.5 Static Equilibrium

Theorem 5.2. Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, let $\underset{\rightarrow}{f}^{y_i}$ be the force applied to particle y_i, and let $\phi_i(\underset{\rightarrow}{r}^{y_i w}) = 0$ be a holonomic constraint associated with y_i. Write $\underset{\rightarrow}{f}^{y_i}$ as $\underset{\rightarrow}{f}^{y_i} = \underset{\rightarrow}{f}^{y_i \, \text{ex}} + \underset{\rightarrow}{f}^{y_i \, \text{hcon}} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_i y_j}$ where $\underset{\rightarrow}{f}^{y_i \, \text{ex}}$ is the external force applied to particle y_i and $\underset{\rightarrow}{f}^{y_i \, \text{hcon}}$ is the constraint force applied to particle y_i due to the presence of holonomic constraints, and $\underset{\rightarrow}{f}^{y_i y_j}$ is the force on y_i due to y_j (which could stem from holonomic constraints also). Then [2] [41, pp. 246-247]

$$\delta W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}, \delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}) = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \, \text{ex}} \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0,$$

where the virtual displacements $\delta \underset{\rightarrow}{r}^{y_i w}$ are compatible with the geometric constraints if

$$\underset{\rightarrow}{v}^{y_i w / a^{\bullet a}} = \underset{\rightarrow}{0}.$$

Proof. Using N2L, if $\underset{\rightarrow}{v}^{y_i w / a^{\bullet a}} = \underset{\rightarrow}{0}$ then

$$\underset{\rightarrow}{f}^{y_i} = \underset{\rightarrow}{0}, \quad i = 1, 2, \ldots, \ell.$$

Taking the dot product of the above with $\delta \underset{\rightarrow}{r}^{y_i w}$ and summing yields

$$\sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i} \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0.$$

From Theorem 5.1 forces $\underset{\rightarrow}{f}^{y_i \, \text{hcon}}$ and $\sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_i y_j}$ do not contribute to the virtual work, and therefore

$$\sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \, \text{ex}} \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0.$$

$\square$

5.2.6 Le Principe de D'Alembert

Theorem 5.3 (Le Principe de D'Alembert)**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, let $\underset{\rightarrow}{f}^{y_i}$ be the force applied to particle y_i, and let $\phi_i(\underset{\rightarrow}{r}^{y_i w}) = 0$ be a holonomic constraint associated with y_i. Write $\underset{\rightarrow}{f}^{y_i}$ as $\underset{\rightarrow}{f}^{y_i} = \underset{\rightarrow}{f}^{y_i \, \text{ex}} + \underset{\rightarrow}{f}^{y_i \, \text{hcon}} + \sum_{j=1}^{\ell} \underset{\rightarrow}{f}^{y_i y_j}$ where $\underset{\rightarrow}{f}^{y_i \, \text{ex}}$ is the external force applied to particle y_i and $\underset{\rightarrow}{f}^{y_i \, \text{hcon}}$ is the constraint force applied to particle y_i due to the presence of holonomic constraints, and $\underset{\rightarrow}{f}^{y_i y_j}$ is the force on y_i due to y_j (which could stem from holonomic constraints also). Then

$$\sum_{i=1}^{\ell} \left(\underset{\rightarrow}{f}^{y_i \,\text{ex}} - m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \right) \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0, \tag{5.10}$$

where the virtual displacements $\delta \underset{\rightarrow}{r}^{y_i w}$ are compatible with the geometric constraints.

Proof. From N2L

$$\underset{\rightarrow}{f}^{y_i} = m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}},$$

which can be rewritten as

$$\underset{\rightarrow}{f}^{y_i} - m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} = \underset{\rightarrow}{0},$$

which can be thought of as "dynamic equilibrium". Taking the dot produce of the above with $\delta \underset{\rightarrow}{r}^{y_i w}$ and summing the results yields

$$\sum_{i=1}^{\ell} \left(\underset{\rightarrow}{f}^{y_i} - m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \right) \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0.$$

Employing Theorem 5.1 gives

$$\sum_{i=1}^{\ell} \left(\underset{\rightarrow}{f}^{y_i \,\text{ex}} - m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \right) \cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0.$$

□

5.2.7 The Generalized Forces and Moments

Returning to Equation (5.7),

$$\delta W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}, \delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}) = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i} \cdot \delta \underset{\rightarrow}{r}^{y_i w}.$$

Recall that $\delta \underset{\rightarrow}{r}^{y_i w}$ is a function of (a subset of) the generalized coordinates, $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^{\mathsf{T}}$. Therefore, $\delta \underset{\rightarrow}{r}^{y_i w}$ can be written

$$\delta \underset{\rightarrow}{r}^{y_i w} = \sum_{k=1}^{n} \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q})}{\partial q_k} \delta q_k. \tag{5.11}$$

As such, using Theorem 5.1 it can be written that

$$\begin{aligned}
& \delta W_{\mathcal{B}w}(\underset{\rightarrow}{f}^{y_1}, \underset{\rightarrow}{f}^{y_2}, \ldots, \underset{\rightarrow}{f}^{y_\ell}, \delta \underset{\rightarrow}{r}^{y_1 w}, \delta \underset{\rightarrow}{r}^{y_2 w}, \ldots, \delta \underset{\rightarrow}{r}^{y_\ell w}) \\
= & \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i} \cdot \delta \underset{\rightarrow}{r}^{y_i w} \\
= & \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \,\text{ex}} \cdot \delta \underset{\rightarrow}{r}^{y_i w} \\
= & \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \,\text{ex}} \cdot \left(\sum_{k=1}^{n} \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q})}{\partial q_k} \delta q_k \right) \\
= & \sum_{i=1}^{\ell} \sum_{k=1}^{n} \underset{\rightarrow}{f}^{y_i \,\text{ex}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q})}{\partial q_k} \delta q_k \\
= & \sum_{k=1}^{n} \mathit{f}_k(\mathbf{q}) \delta q_k
\end{aligned}$$

where

$$\mathit{f}_k(\mathbf{q}(t), t) = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \,\text{ex}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k} \tag{5.12}$$

are the generalized forces and moments. The generalized forces and moments can be written as a column matrix,

$$\boldsymbol{f}(\mathbf{q}(t), t) = \underset{k=1,2,\ldots,n}{\text{col}} \left\{ \mathit{f}_k(\mathbf{q}(t), t) \right\}. \tag{5.13}$$

5.3 Lagrange's Equation for a System of Particles

5.3.1 Derivation of Lagrange's Equation for a Unconstrained System of Particles

The derivation of Lagrange's equation where the generalized coordinates are independent will now be considered. This means that the system of particles is either unconstrained or subject only to holonomic constraints and the redundant generalized coordinates have been eliminated.

Theorem 5.4 (Lagrange's Equation for an Unconstrained System of Particles)**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^{\mathsf{T}}$ be the independent generalized coordinates, and let $\underset{\rightarrow}{f}^{y_i}$ be the external force applied to particle y_i. Then the differential equations of motion of $\mathcal{B}$ are given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}} \right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \mathbf{q}} \right)^{\mathsf{T}} = \boldsymbol{f}(\mathbf{q}(t), t), \tag{5.14}$$

where

$$\begin{aligned}
\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}} &= \left[\begin{array}{cccc} \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_1} & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_2} & \cdots & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_n} \end{array}\right], \\
\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}} &= \left[\begin{array}{cccc} \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_1} & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_2} & \cdots & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_n} \end{array}\right], \\
\boldsymbol{f}(\mathbf{q}(t),t) &= \left[\begin{array}{cccc} f_1(\mathbf{q}(t),t) & f_2(\mathbf{q}(t),t) & \cdots & f_n(\mathbf{q}(t),t) \end{array}\right]^{\mathsf{T}},
\end{aligned}$$

and $L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) = T_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) - U_{\mathcal{B}w}(\mathbf{q}(t))$.

Proof. Starting from D'Alembert's Principle, Equation (5.10), which is

$$\sum_{i=1}^{\ell}\left(\underset{\rightarrow}{f}^{y_i\,\text{ex}} - m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}}\right)\cdot \delta \underset{\rightarrow}{r}^{y_i w} = 0.$$

Rearranging and using Equation (5.11) gives

$$\begin{aligned}
\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \cdot \delta \underset{\rightarrow}{r}^{y_i w} &= \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i}\cdot \delta \underset{\rightarrow}{r}^{y_i w}, \\
\sum_{i=1}^{\ell}\sum_{k=1}^{n} m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\delta q_k &= \sum_{i=1}^{\ell}\sum_{k=1}^{n} \underset{\rightarrow}{f}^{y_i}\cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\delta q_k = \sum_{k=1}^{n} f_k \delta q_k,
\end{aligned} \tag{5.15}$$

where the definition of the generalized forces and moments given in Equation (5.12) has been employed. Consider the integration of the term in the double sum on the left-hand-side of Equation (5.15):

$$m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k} = \frac{\mathrm{d}}{\mathrm{d}t}\left(m_i \underset{\rightarrow}{v}^{y_i w/a}\cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\right)\Bigg|_{\mathcal{F}_a} - m_i \underset{\rightarrow}{v}^{y_i w/a}\cdot \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\right)\Bigg|_{\mathcal{F}_a} \tag{5.16}$$

where the $\frac{\mathrm{d}(\cdot)}{\mathrm{d}t}$ is w.r.t. $\mathcal{F}_a$. Note that

$$\underset{\rightarrow}{v}^{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) = \underset{\rightarrow}{r}^{y_i w/a^{\bullet a}}(\mathbf{q}(t),t) = \sum_{k=1}^{n}\frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k}\dot{q}_k(t) + \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial t}. \tag{5.17}$$

Taking the derivative with respect to $\dot{q}_k$ on both the far-left-hand and far-right-hand sides of Equation (5.17) gives

$$\frac{\partial \underset{\rightarrow}{v}^{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_k} = \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k}. \tag{5.18}$$

Furthermore, using the far-left-hand and far-right-hand sides of Equation (5.17) once more,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k}\right)\Bigg|_{\mathcal{F}_a} &= \sum_{s=1}^{n} \frac{\partial}{\partial q_s}\left(\frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k}\right)\dot{q}_s(t) + \frac{\partial}{\partial t}\left(\frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k}\right)\\
&= \sum_{s=1}^{n} \frac{\partial^2 \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_s \partial q_k}\dot{q}_s(t) + \frac{\partial^2 \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial t \partial q_k}\\
&= \sum_{s=1}^{n} \frac{\partial^2 \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k \partial q_s}\dot{q}_s(t) + \frac{\partial^2 \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_k \partial t}\\
&= \frac{\partial}{\partial q_k}\underbrace{\left(\sum_{s=1}^{n} \frac{\partial \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial q_s}\dot{q}_s(t) + \frac{\partial^2 \underset{\rightarrow}{r}^{y_i w}(\mathbf{q}(t))}{\partial t}\right)}_{=\underset{\rightarrow}{v}^{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) \text{ via Eq. (5.17)}}\\
&= \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_k},
\end{aligned}
\tag{5.19}
$$

where the fact that differentiation order can be interchanged has been exploited. Substituting Equations (5.18) and (5.19) into Equation (5.16) results in

$$
\begin{aligned}
m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\right)\Bigg|_{\mathcal{F}_a} - m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k}\right)\Bigg|_{\mathcal{F}_a}\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial \dot{q}_k}\right)\Bigg|_{\mathcal{F}_a} - m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial q_k}.
\end{aligned}
\tag{5.20}
$$

Recall from Definition 4.25 that the kinetic energy of particle y_i relative to point w w.r.t. $\mathcal{F}_a$ is

$$
T_{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) = \tfrac{1}{2} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a},
$$

and the kinetic energy of body $\mathcal{B}$ relative to w w.r.t. $\mathcal{F}_a$ is defined as

$$
T_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t)) = \tfrac{1}{2}\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a}.
$$

Using the definition of the kinetic energy of particle y_i relative to point w w.r.t. $\mathcal{F}_a$,

$$
m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial \dot{q}_k} = \frac{\partial}{\partial \dot{q}_k}\left(\tfrac{1}{2} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a}\right), \qquad m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial q_k} = \frac{\partial}{\partial q_k}\left(\tfrac{1}{2} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a}\right),
$$

and therefore Equation (5.20) can be written

$$
\begin{aligned}
m_i \underset{\rightarrow}{v}^{y_i w/a^{\bullet a}} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial \dot{q}_k}\right)\Bigg|_{\mathcal{F}_a} - m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \frac{\partial \underset{\rightarrow}{v}^{y_i w/a}}{\partial q_k}\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{q}_k}\left(\tfrac{1}{2} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a}\right)\right) - \frac{\partial}{\partial q_k}\left(\tfrac{1}{2} m_i \underset{\rightarrow}{v}^{y_i w/a} \cdot \underset{\rightarrow}{v}^{y_i w/a}\right)\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T_{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_k}\right) - \frac{\partial T_{y_i w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_k}.
\end{aligned}
$$

Summing over the number of particles and using the fact that $T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) = \sum_{i=1}^{\ell} T_{y_i w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ results in

$$\begin{aligned}
\sum_{i=1}^{\ell} m_i \underset{\rightarrow}{v}^{y_i w/a \bullet a} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_i w}}{\partial q_k} &= \sum_{i=1}^{\ell} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{y_i w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial T_{y_i w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q_k} \right] \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q_k}.
\end{aligned}$$

Using this result in conjunction with Equation (5.15) yields

$$\begin{aligned}
\sum_{k=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q_k} \right) \delta q_k &= \sum_{k=1}^{n} f_k(\mathbf{q}(t), t) \delta q_k, \\
\sum_{k=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q_k} - f_k(\mathbf{q}(t), t) \right) \delta q_k &= 0. \qquad (5.21)
\end{aligned}$$

Next, write $\underset{\rightarrow}{f}^{y_i}$ as $\underset{\rightarrow}{f}^{y_i} = -\underset{\rightarrow}{\nabla} U_{y_i w} + \underset{\rightarrow}{f}^{y_i \triangle}$ where $-\underset{\rightarrow}{\nabla} U_{y_i w}$ are conservative forces derivable from a potential and $\underset{\rightarrow}{f}^{y_i \triangle}$ are nonconservative forces. Noting that

$$-\underset{\rightarrow}{\nabla} U_{y_i w} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_1 w}}{\partial q_k} = \frac{\partial U_{y_i w}}{\partial q_k},$$

the generalized forces and moments can be written

$$\begin{aligned}
f_k &= -\frac{\partial}{\partial q_k} \sum_{i=1}^{\ell} U_{y_i w} + \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \triangle} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_1 w}}{\partial q_k} \\
&= -\frac{\partial}{\partial q_k} \sum_{i=1}^{\ell} U_{y_i w} + f_k^{\triangle} \\
&= -\frac{\partial U_{\mathcal{B}w}}{\partial q_k} + f_k^{\triangle} \qquad (5.22)
\end{aligned}$$

where

$$f_k^{\triangle} = \sum_{i=1}^{\ell} \underset{\rightarrow}{f}^{y_i \triangle} \cdot \frac{\partial \underset{\rightarrow}{r}^{y_1 w}}{\partial q_k}$$

are the generalized forces and moments stemming from nonconservative forces and $U_{\mathcal{B}w}(\mathbf{q}(t)) = \sum_{i=1}^{\ell} U_{y_i w}(\mathbf{q}(t))$. Substitution of Equation (5.22) into Equation (5.21) yields

$$\begin{aligned}
\sum_{k=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial q_k} + \frac{\partial U_{\mathcal{B}w}(\mathbf{q}(t))}{\partial q_k} - f_k^{\triangle}(\mathbf{q}(t), t) \right) \delta q_k &= 0, \\
\sum_{k=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))}{\partial \dot{q}_k} \right) - \frac{\partial (T_{\mathcal{B}w/a}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) - U_{\mathcal{B}w}(\mathbf{q}(t)))}{\partial q_k} - f_k^{\triangle}(\mathbf{q}(t), t) \right) \delta q_k &= 0.
\end{aligned}$$

The potential energy is generally a function of the generalized coordinates but not the time-rate-of-change of the generalized coordinates. Thus, by defining $L_{\mathcal{B}w/a} = T_{\mathcal{B}w/a} - U_{\mathcal{B}w}$, it follows that

$$\begin{aligned}\sum_{k=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_k}\right) - \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_k} - f_k^{\triangle}(\mathbf{q}(t),t)\right)\delta q_k &= 0,\\ \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}\right) - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}\right) - \boldsymbol{f}^{\mathsf{T}}(\mathbf{q}(t),t)\right)\delta\mathbf{q} = 0,\end{aligned} \tag{5.23}$$

where, for simplicity, the $f_k^{\triangle}$'s are denoted f_k and $\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}$, $\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}$, and $\boldsymbol{f}(\mathbf{q}(t),t)$ are defined in the theorem statment of Theorem 5.4. The generalized coordinate are independent. Therefore, Equation (5.23) must hold for arbitrary δq_k's leading to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}\right)^{\mathsf{T}} = \boldsymbol{f}(\mathbf{q}(t),t).$$

□

Note that the generalized forces and moments are a function of external forces and moments only.

5.3.2 Derivation of Lagrange's Equation for a Constrained System of Particles

The derivation of Lagrange's equation where the generalized coordinates are dependent and subject to both holonomic and nonholonomic constraints will now be considered.

Theorem 5.5 (Lagrange's Equation for a Constrained System of Particles)**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}$ be a body composed of particles $y_1, y_2, \ldots, y_\ell$ whose masses are $m_1, m_2, \ldots, m_\ell$, respectfully, let $\underset{\rightarrow}{r}^{y_i w}$ be the position of y_i relative to w, let $\mathbf{q}(t) = [q_1(t)\ \ q_2(t)\ \ \cdots\ \ q_n(t)]^{\mathsf{T}}$ be the dependent generalized coordinates, let $\underset{\rightarrow}{f}^{y_i}$ be the external force applied to particle y_i, and let

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t),t)\mathrm{d}q_k + \Xi_{jt}(\mathbf{q}(t),t)\mathrm{d}t = 0, \quad j = 1,2,\ldots,m,$$

be m linearly independent constraints in Pfaffian form. Then the differential equations of motion of $\mathcal{B}$ are given by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}\right)^{\mathsf{T}} &= \boldsymbol{f}(\mathbf{q}(t),t) + \boldsymbol{\Xi}^{\mathsf{T}}(\mathbf{q}(t),t)\boldsymbol{\lambda}(t), \\ \boldsymbol{\Xi}(\mathbf{q}(t),t)\dot{\mathbf{q}} + \boldsymbol{\Xi}_t(\mathbf{q}(t),t) &= \mathbf{0},\end{aligned} \tag{5.24}$$

where $L_{\mathcal{B}w/a} = T_{\mathcal{B}w/a} - U_{\mathcal{B}w}$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the Lagrange multipliers associated with the m constraints, $\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}$, $\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}$, and $\boldsymbol{f}(\mathbf{q})$ are defined in the theorem statement of Theorem 5.4, and

$$\boldsymbol{\Xi}(\mathbf{q}(t),t) = \underset{\substack{j=1,2,\ldots,m\\ k=1,2,\ldots,n}}{\mathrm{matrix}}\left\{\Xi_{jk}(\mathbf{q}(t),t)\right\}, \quad \boldsymbol{\Xi}_t(\mathbf{q}(t),t) = \underset{j=1,2,\ldots,m}{\mathrm{col}}\left\{\Xi_{jt}(\mathbf{q}(t),t)\right\}.$$

Proof. The m constraints are linearly independent. By a suitable ordering of the generalized coordinates the rank m matrix $\boldsymbol{\Xi}$ can be written

$$\boldsymbol{\Xi} = \left[\begin{array}{cc}\boldsymbol{\Xi}_1 & \boldsymbol{\Xi}_2\end{array}\right] \tag{5.25}$$

where $\boldsymbol{\Xi}_1 \in \mathbb{R}^{m\times m}$ is square and full rank and $\boldsymbol{\Xi}_2 \in \mathbb{R}^{m\times(n-m)}$. Owing to the presence of constraints the virtual displacements, $\delta\mathbf{q}$, are no longer independent, and together must satisfy

$$\boldsymbol{\Xi}\delta\mathbf{q} = \mathbf{0}. \tag{5.26}$$

Write $\delta\mathbf{q}$ as

$$\delta\mathbf{q} = \begin{bmatrix} \delta\mathbf{q}_1 \\ \delta\mathbf{q}_2 \end{bmatrix}.$$

Using the partitioning in Equation (5.25), the constraint given in Equation (5.26) can be written as

$$\begin{aligned} \begin{bmatrix} \boldsymbol{\Xi}_1 & \boldsymbol{\Xi}_2 \end{bmatrix} \begin{bmatrix} \delta\mathbf{q}_1 \\ \delta\mathbf{q}_2 \end{bmatrix} &= \mathbf{0}, \\ \delta\mathbf{q}_1 &= -\boldsymbol{\Xi}_1^{-1}\boldsymbol{\Xi}_2\delta\mathbf{q}_2, \end{aligned}$$

where the m virtual displacements $\delta\mathbf{q}_1$ clearly depend on the $n-m$ virtual displacements $\delta\mathbf{q}_2$. Next, using the same ordering of the generalized coordinates as before, Equation (5.23) can be written

$$\delta\mathbf{q}_1^{\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_1}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_1}\right)^{\mathsf{T}} - \boldsymbol{f}_1 \right) + \delta\mathbf{q}_2^{\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_2}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_2}\right)^{\mathsf{T}} - \boldsymbol{f}_2 \right) = 0 \tag{5.27}$$

If Equation (5.26) is true, then

$$\delta\mathbf{q}^{\mathsf{T}}\boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\lambda} = 0 \tag{5.28}$$

is also true for any $\boldsymbol{\lambda} \in \mathbb{R}^m$. The right-hand-side of Equation (5.28) is zero, thus, adding or subtracting a partitioned version of Equation (5.28) to Equation (5.27) results in

$$\delta\mathbf{q}_1^{\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_1}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_1}\right)^{\mathsf{T}} - \boldsymbol{f}_1 - \boldsymbol{\Xi}_1^{\mathsf{T}}\boldsymbol{\lambda} \right) + \delta\mathbf{q}_2^{\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_2}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_2}\right)^{\mathsf{T}} - \boldsymbol{f}_2 - \boldsymbol{\Xi}_2^{\mathsf{T}}\boldsymbol{\lambda} \right) = 0 \tag{5.29}$$

Equation (5.29) must hold for all $\delta\mathbf{q}$'s satisfying Equation (5.26). Given that the $\boldsymbol{\lambda}$'s are arbitrary, let

$$\boldsymbol{\lambda} = \boldsymbol{\Xi}_1^{-\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}}\right)^{\mathsf{T}} - \boldsymbol{f}_1 \right).$$

For this choice of $\boldsymbol{\lambda}$ the first term in Equation (5.29) equals zero,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_1}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_1}\right)^{\mathsf{T}} - \boldsymbol{f}_1 - \boldsymbol{\Xi}_1^{\mathsf{T}}\boldsymbol{\lambda} = \mathbf{0}, \tag{5.30}$$

and from Equation (5.29)

$$\delta\mathbf{q}_2^{\mathsf{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_2}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_2}\right)^{\mathsf{T}} - \boldsymbol{f}_2 - \boldsymbol{\Xi}_2^{\mathsf{T}}\boldsymbol{\lambda} \right) = 0.$$

Owing to the fact that the $\delta\mathbf{q}_2$'s are independent it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \dot{\mathbf{q}}_2}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}}{\partial \mathbf{q}_2}\right)^{\mathsf{T}} - \boldsymbol{f}_2 - \boldsymbol{\Xi}_2^{\mathsf{T}}\boldsymbol{\lambda} = \mathbf{0}. \tag{5.31}$$

Combining Equations (5.30) and (5.31) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}\right)^{\mathsf{T}} = \boldsymbol{f}(\mathbf{q}(t),t) + \boldsymbol{\Xi}^{\mathsf{T}}(\mathbf{q}(t),t)\boldsymbol{\lambda}$$

which, together with $\boldsymbol{\Xi}(\mathbf{q}(t),t)\dot{\mathbf{q}} + \boldsymbol{\Xi}_t(\mathbf{q}(t),t) = \mathbf{0}$, describe the motion of $\mathcal{B}$. □

The above derivation is partially inspired by a discussion with Dr. Anton de Ruiter.

5.4 Lagrange's Equation for Rigid Bodies

Axiom 5.1 (Lagrange's Equation for a System of Rigid Bodies)**.** Let $\mathcal{F}_a$ be an inertial frame, let w be an unforced particle, let $\mathcal{B}_i$ be a continuous rigid-body, let $\mathbf{q}(t) = [q_1(t) \;\; q_2(t) \;\; \cdots \;\; q_n(t)]^{\mathsf{T}}$ be the dependent generalized coordinates, let $\underset{\rightarrow}{f}^{\,i}$ be the external force applied to $\mathcal{B}_i$, and let

$$\sum_{k=1}^{n} \Xi_{jk}(\mathbf{q}(t),t)\mathrm{d}q_k + \Xi_{jt}(\mathbf{q}(t),t)\mathrm{d}t = 0, \quad j = 1,2,\ldots,m,$$

be m linearly independent constraints in Pfaffian form. Then the differential equations of motion of the system of rigid bodies, denoted $\mathcal{B}$, are given by

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}}\right)^{\mathsf{T}} - \left(\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}}\right)^{\mathsf{T}} &= \boldsymbol{f}(\mathbf{q}(t),t) + \boldsymbol{\Xi}^{\mathsf{T}}(\mathbf{q}(t),t)\boldsymbol{\lambda}(t), \\
\boldsymbol{\Xi}(\mathbf{q}(t),t)\dot{\mathbf{q}} + \boldsymbol{\Xi}_t(\mathbf{q}(t),t) &= \mathbf{0},
\end{aligned}$$

where $L_{\mathcal{B}w/a} = T_{\mathcal{B}w/a} - U_{\mathcal{B}w}$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the Lagrange multipliers associated with the m constraints,

$$\begin{aligned}
\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{\mathbf{q}}} &= \begin{bmatrix} \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_1} & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_2} & \cdots & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \dot{q}_n} \end{bmatrix}, \\
\frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial \mathbf{q}} &= \begin{bmatrix} \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_1} & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_2} & \cdots & \frac{\partial L_{\mathcal{B}w/a}(\mathbf{q}(t),\dot{\mathbf{q}}(t))}{\partial q_n} \end{bmatrix}, \\
\boldsymbol{f}(\mathbf{q}(t),t) &= \begin{bmatrix} f_1(\mathbf{q}(t),t) & f_2(\mathbf{q}(t),t) & \cdots & f_n(\mathbf{q}(t),t) \end{bmatrix}^{\mathsf{T}}, \\
\boldsymbol{\Xi}(\mathbf{q}(t),t) &= \underset{\substack{j=1,2,\ldots,m \\ k=1,2,\ldots,n}}{\text{matrix}} \{\Xi_{jk}(\mathbf{q}(t),t)\}, \qquad \boldsymbol{\Xi}_t(\mathbf{q}(t),t) = \underset{j=1,2,\ldots,m}{\text{col}} \{\Xi_{jt}(\mathbf{q}(t),t)\}.
\end{aligned}$$

As discussed in [42, pp. 4-6] the Newton-Euler approach to deriving the differential equation of motion describing a rigid body or a system composed of many rigid bodies can be shown to be equivalent to Axiom 5.1 in many situations. However, it has not been proved, nor does [42] prove, that equivalence holds in all situations. However, to quote [42, pp. 207-214], "There is no doubt that the correct equations of motion for nonholonomic mechanical systems are given by Axiom 5.1" [1].

[1] In [42, pp. 208] Axiom 5.1 is called the Lagrange - D'Alembert principle.

5.5 The General Form of the Differential Equations of Motion

In the unconstrained case, the motion equations of a system of particles or rigid bodies will generally be of the form

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{f}, \tag{5.32}$$

where $\mathbf{M}(\mathbf{q}) = \mathbf{M}^{\mathsf{T}}(\mathbf{q}) > 0, \forall \mathbf{q} \in \mathbb{R}^n$ is the mass matrix, $\mathbf{D} = \mathbf{D}^{\mathsf{T}} \geq 0$ is the damping matrix, $\mathbf{K} = \mathbf{K}^{\mathsf{T}} \geq 0$ is the stiffness matrix, $\mathbf{f}_{\text{non}}(\cdot, \cdot)$ is a column matrix composed of nonlinear coupling terms, $\hat{\mathbf{B}}$ is the input matrix (also called the input-mapping matrix), and $\mathbf{f}$ are the forces and moments applied to the system. The mass, damping, and stiffness matrices are each $n \times n$ where n is the number of (independent) generalized coordinates. The damping and stiffness matrices are generally constant, while the mass matrix is either constant or a function of the generalized coordinates.

5.6 Numerically Solving Differential Equations With and Without the Lagrange Multipliers

In the constrained case, the motion equations of a system of particles or rigid bodies will generally be of the form

$$\begin{aligned} \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} &= \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\lambda} + \mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{f}, \\ \boldsymbol{\Xi}\dot{\mathbf{q}} + \boldsymbol{\Xi}_t &= \mathbf{0}. \end{aligned} \tag{5.33}$$

The mass, damping, and stiffness matrices are each $n \times n$ where n is the number of (dependent) generalized coordinates. There are various ways to solve this system of equations numerically, as discussed next.

5.6.1 With Lagrange Multipliers

First, take the time derivative of the constraint equation, $\boldsymbol{\Xi}\dot{\mathbf{q}} + \boldsymbol{\Xi}_t = \mathbf{0}$, to get

$$\boldsymbol{\Xi}\ddot{\mathbf{q}} + \dot{\boldsymbol{\Xi}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Xi}}_t = \mathbf{0}.$$

Combining this with Equation (5.33) in matrix form yields

$$\begin{bmatrix} \mathbf{M}(\mathbf{q}) & -\boldsymbol{\Xi}^{\mathsf{T}} \\ -\boldsymbol{\Xi} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} -\mathbf{D}\dot{\mathbf{q}} - \mathbf{K}\mathbf{q} + \mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{f} \\ \dot{\boldsymbol{\Xi}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Xi}}_t \end{bmatrix}$$

Recall that $\boldsymbol{\Xi}$ is a full rank matrix (i.e., it has full row rank because the constraints are linearly independent) and that the mass matrix is positive definite. As such, the matrix, which is called the Karush-Kuhn-Tucker (KKT) matrix,

$$\begin{bmatrix} \mathbf{M}(\mathbf{q}) & -\boldsymbol{\Xi}^{\mathsf{T}} \\ -\boldsymbol{\Xi} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)} \tag{5.34}$$

is nonsingular, and hence

$$\begin{bmatrix} \ddot{\mathbf{q}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{M}(\mathbf{q}) & -\boldsymbol{\Xi}^{\mathsf{T}} \\ -\boldsymbol{\Xi} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{D}\dot{\mathbf{q}} - \mathbf{K}\mathbf{q} + \mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{f} \\ \dot{\boldsymbol{\Xi}}\dot{\mathbf{q}} + \dot{\boldsymbol{\Xi}}_t \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\ddot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}) \\ \mathbf{f}_{\boldsymbol{\lambda}}(\mathbf{q}, \dot{\mathbf{q}}) \end{bmatrix}.$$

Numerically integrating

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} \dot{\mathbf{q}} \\ \mathbf{f}_{\ddot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}) \end{bmatrix}$$

yields $\mathbf{q}$ and $\dot{\mathbf{q}}$ versus time.

Often the Lagrange multipliers are related to constraint forces and moments. The Lagrange multipliers can be solved for directly via

$$\boldsymbol{\lambda} = \mathbf{f}_{\boldsymbol{\lambda}}(\mathbf{q}, \dot{\mathbf{q}}).$$

This is usually done in a post-processing step outside of the numerical integration process.

5.6.2 Without Lagrange Multipliers: The Null Space Method

Recall that often a holonomic constraint can be eliminated before employing Lagrange's equation to derive the motion equations of a system. However, it can sometimes be easier to derive the motion equations of a system subject to the constraint $\boldsymbol{\Xi}\dot{\mathbf{q}} = \mathbf{0}$. (Note, the constraints are now assumed to be scleronomic.) Doing so introduces Lagrange multipliers and the need to invert the matrix given in Equation (5.34). If solving for the Lagrange multipliers is unnecessary, and inverting Equation (5.34) is computationally intensive, an alternative approach is to find a independent set of generalized coordinate rates and remove the Lagrange multiplier. This methods also works for systems subject to nonholonomic constraints. This procedure explores the fact that $\dot{\mathbf{q}}$ lies in the null space of $\boldsymbol{\Xi}$, and is called the null space method.

The approach is as follows. Let $\dot{\mathbf{q}}$ be the dependent generalized coordinate rates that must satisfy

$$\boldsymbol{\Xi}\dot{\mathbf{q}} = \mathbf{0} \tag{5.35}$$

at all times, and consider

$$\dot{\mathbf{q}} = \boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}} \tag{5.36}$$

where $\boldsymbol{\Upsilon} \in \mathbb{R}^{n \times r}$ is a full (column) rank mapping matrix between $\dot{\hat{\mathbf{q}}}$, a set of independent generalized coordinate rates. Momentarily the time derivative of Equation (5.36),

$$\ddot{\mathbf{q}} = \boldsymbol{\Upsilon}\ddot{\hat{\mathbf{q}}} + \dot{\boldsymbol{\Upsilon}}\dot{\hat{\mathbf{q}}}, \tag{5.37}$$

will be needed. Substitution of Equation (5.36) into Equation (5.35) gives

$$\boldsymbol{\Xi}\boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}} = \mathbf{0}$$

which must hold for all $\dot{\hat{\mathbf{q}}}$. It follows that

$$\boldsymbol{\Xi}\boldsymbol{\Upsilon} = \mathbf{0}, \quad \boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{\Xi}^{\mathsf{T}} = \mathbf{0}. \tag{5.38}$$

The matrices $\boldsymbol{\Xi}$ and $\boldsymbol{\Upsilon}$ are called orthogonal complements. Next, returning to Equation (5.33),

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\lambda} + \mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{f}.$$

Premultiplying Equation (5.33) by $\boldsymbol{\Upsilon}^{\mathsf{T}}$ yields

$$\boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{D}\dot{\mathbf{q}} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{K}\mathbf{q} = \boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\lambda} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}}) + \boldsymbol{\Upsilon}^{\mathsf{T}}\hat{\mathbf{B}}\mathbf{f}$$

Via Equation (5.38) the term $\boldsymbol{\Upsilon}^{\mathsf{T}}\boldsymbol{\Xi}^{\mathsf{T}}\boldsymbol{\lambda}$ drops out. Using Equations (5.36) and (5.37) results in

$$\boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{M}(\mathbf{q})\boldsymbol{\Upsilon}\ddot{\hat{\mathbf{q}}} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{M}(\mathbf{q})\dot{\boldsymbol{\Upsilon}}\dot{\hat{\mathbf{q}}} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{D}\boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}} + \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{K}\mathbf{q} = \boldsymbol{\Upsilon}^{\mathsf{T}}\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\hat{\mathbf{q}}}) + \boldsymbol{\Upsilon}^{\mathsf{T}}\hat{\mathbf{B}}\mathbf{f}.$$

This can be written as

$$\hat{\mathbf{M}}(\mathbf{q})\ddot{\hat{\mathbf{q}}} + \hat{\mathbf{D}}\dot{\hat{\mathbf{q}}} + \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{K}\mathbf{q} = \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\hat{\mathbf{q}}}) + \boldsymbol{\Upsilon}^\mathsf{T}\hat{\mathbf{B}}\mathbf{f}, \tag{5.39}$$

where $\hat{\mathbf{M}} = \hat{\mathbf{M}}^\mathsf{T} > 0$ and $\hat{\mathbf{D}} = \hat{\mathbf{D}}^\mathsf{T} \geq 0$ are the mass and damping matrices associated with $\dot{\hat{\mathbf{q}}}$. Premultiplying Equation (5.39) by the inverse of $\hat{\mathbf{M}}$ results in

$$\ddot{\hat{\mathbf{q}}} = \hat{\mathbf{M}}^{-1}(\mathbf{q})\left(-\hat{\mathbf{D}}\dot{\hat{\mathbf{q}}} - \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{K}\mathbf{q} + \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\hat{\mathbf{q}}}) + \boldsymbol{\Upsilon}^\mathsf{T}\hat{\mathbf{B}}\mathbf{f}\right).$$

Numerically integrating

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \dot{\hat{\mathbf{q}}} \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{bmatrix} \boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}} \\ \hat{\mathbf{M}}^{-1}(\mathbf{q})\left(-\hat{\mathbf{D}}\dot{\hat{\mathbf{q}}} - \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{K}\mathbf{q} + \boldsymbol{\Upsilon}^\mathsf{T}\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\hat{\mathbf{q}}}) + \boldsymbol{\Upsilon}^\mathsf{T}\hat{\mathbf{B}}\mathbf{f}\right) \end{bmatrix}$$

yields $\mathbf{q}$ and $\dot{\hat{\mathbf{q}}}$ versus time. Should it be needed $\dot{\mathbf{q}}$ can be computed at each numerical integration step via Equation (5.36). For instance, often $\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}})$ is already computed; rather than manually computing $\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\hat{\mathbf{q}}})$, one simply computes $\dot{\mathbf{q}}$ via $\dot{\mathbf{q}} = \boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}}$ and then computes $\mathbf{f}_{\text{non}}(\mathbf{q}, \dot{\mathbf{q}})$.

The above formulation hinges on two relationships: the first is Equation (5.35), $\boldsymbol{\Xi}\dot{\mathbf{q}} = \mathbf{0}$, and Equation (5.36), $\dot{\mathbf{q}} = \boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}}$. Also, this formulation works equally well for holonomic as well as nonholnomic systems. The reason it works for nonholonomic systems is that the formulation is based on the relationship $\dot{\mathbf{q}} = \boldsymbol{\Upsilon}\dot{\hat{\mathbf{q}}}$, a relationship between dependent generalized coordinate rates and independent generalized coordinate rates.

This procedure of eliminating the Lagrange multipliers and finding an alternative expression of the motion equations is known as the null space method, the natural orthogonal complement method, and as Maggie's formulation.

Appendices

Appendix A

Linear Algebra, Differential Equations, Stability, and Numerical Methods

In this chapter linear algebra, differential equations, stability of differential equations, and numerical methods for solving differential equations will be reviewed.

A.1 Linear Algebra

Linear algebra is an invaluable tool, not just when studying kinematics and dynamics, but in all branches of engineering and science. As such, a brief review of the essentials of linear algebra will be reviewed.

A.1.1 Vector Spaces and Subspaces

Roughly speaking (and for the precise definition of a vector space, see Definition A.1) a vector space is composed of four things: two sets, $\boldsymbol{\mathcal{V}}$ and $\mathbb{F}$, and two algebraic operations called *vector addition* and *scalar multiplication*. The set $\boldsymbol{\mathcal{V}}$ is a nonempty set of objects called vectors that will be denoted $\mathbf{v}$. The set $\mathbb{F}$ is a scalar field, usually $\mathbb{R}$, the set of real numbers, or $\mathbb{C}$, the set of complex numbers. Vector addition, $\mathbf{v}^1 + \mathbf{v}^2$ where $\mathbf{v}^1, \mathbf{v}^2 \in \boldsymbol{\mathcal{V}}$, is an operation between the elements of $\boldsymbol{\mathcal{V}}$, while scalar multiplication, $\alpha\mathbf{v}$ where $\alpha \in \mathbb{F}$ and $\mathbf{v} \in \boldsymbol{\mathcal{V}}$, is an operation between elements of both $\boldsymbol{\mathcal{V}}$ and $\mathbb{F}$. A vector space is a set of rules stipulating how these four things work together.

Definition A.1. [10, pp. 292] [9, pp. 160] A *vector space* $\boldsymbol{\mathcal{V}}$ over a field of scalars $\mathbb{F}$ (e.g., $\mathbb{R}$ or $\mathbb{C}$) is a set of objects, called vectors, that together with two operations, addition and scalar multiplication, satisfy the following ten axioms.

1. $\mathbf{v}^1 + \mathbf{v}^2 \in \boldsymbol{\mathcal{V}}, \quad \forall \mathbf{v}^1, \mathbf{v}^2 \in \boldsymbol{\mathcal{V}}$ (closure under addition).
2. $(\mathbf{v}^1 + \mathbf{v}^2) + \mathbf{v}^3 = \mathbf{v}^1 + (\mathbf{v}^2 + \mathbf{v}^3), \quad \forall \mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3 \in \boldsymbol{\mathcal{V}}$ (associative law of vector addition).
3. $\mathbf{v}^1 + \mathbf{v}^2 = \mathbf{v}^2 + \mathbf{v}^1, \quad \forall \mathbf{v}^1, \mathbf{v}^2 \in \boldsymbol{\mathcal{V}}$ (commutative law of vector addition).
4. $\exists \mathbf{0} \in \boldsymbol{\mathcal{V}}$ s.t. $\mathbf{v} + \mathbf{0} = \mathbf{v}, \quad \forall \mathbf{v} \in \boldsymbol{\mathcal{V}}$ (the existence of a zero vector).
5. If $\mathbf{v}^1 \in \boldsymbol{\mathcal{V}}, \quad \exists - \mathbf{v}^1 \in \boldsymbol{\mathcal{V}}$ s.t. $\mathbf{v}^1 + (-\mathbf{v}^1) = \mathbf{0}$ ($-\mathbf{v}^1$ is the additive inverse of $\mathbf{v}^1$).
6. $\alpha\mathbf{v} \in \boldsymbol{\mathcal{V}}, \quad \forall \mathbf{v} \in \boldsymbol{\mathcal{V}}, \forall \alpha \in \mathbb{F}$ (closure under scalar multiplication).
7. $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v}), \quad \forall \mathbf{v} \in \boldsymbol{\mathcal{V}}, \forall \alpha, \beta \in \mathbb{F}$ (associative law of scalar multiplication).

8. $\alpha(\mathbf{v}^1 + \mathbf{v}^2) = \alpha\mathbf{v}^2 + \alpha\mathbf{v}^1, \quad \forall \mathbf{v}^1, \mathbf{v}^2 \in \boldsymbol{\mathcal{V}}, \forall \alpha \in \mathbb{F}$ (distributive law).
9. $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}, \quad \forall \mathbf{v} \in \boldsymbol{\mathcal{V}}, \forall \alpha, \beta \in \mathbb{F}$ (distributive law with α and β).
10. $1\mathbf{v} = \mathbf{v}, \quad \forall \mathbf{v} \in \boldsymbol{\mathcal{V}}$.

Example A.1. [9, pp. 161] The set of real $m \times n$ matrices is a (real) vector space. Also, the set of complex $m \times n$ matrices is a (complex) vector space.

Example A.2. [10, pp. 293] Lines in $\mathbb{R}^2$ passing through the origin constitute a (real) vector space.

Proof. Let $\boldsymbol{\mathcal{V}} = \{x, y \in \mathbb{R} \mid y = mx, \ \forall m \in \mathbb{R}\}$. Consider two elements of $\boldsymbol{\mathcal{V}}$

$$\mathbf{v}^1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ mx_1 \end{bmatrix}, \quad \mathbf{v}^2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ mx_2 \end{bmatrix}.$$

Then,

$$\mathbf{v}^1 + \mathbf{v}^2 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ mx_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ mx_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ m(x_1 + x_2) \end{bmatrix} \in \boldsymbol{\mathcal{V}}.$$

Hence, Axiom 1 in Definition A.1 is satisfied.

Next, consider $\mathbf{v}^1$ defined above and $\alpha \in \mathbb{R}$. Then,

$$\alpha\mathbf{v}^1 = \begin{bmatrix} \alpha x_1 \\ m(\alpha x_1) \end{bmatrix} \in \boldsymbol{\mathcal{V}}.$$

Hence, Axiom 6 in Definition A.1 is satisfied.

The remaining eight axioms can be shown to hold as well. Hence, $\boldsymbol{\mathcal{V}}$ is a real vector space. □

Example A.3. [10, pp. 295] Let $\mathcal{P}^n$ denote the set of polynomials with real coefficients of degree less than or equal to n, that is, if $p \in \mathcal{P}^n$ then

$$p(x, a_0, a_1, \ldots, a_{n-1}, a_n) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = \sum_{i=0}^{n} a_i x^i, \quad \forall a_i \in \mathbb{R}, \quad i = 1, 2, \ldots, n.$$

$\mathcal{P}^n$ constitutes a real vector space.

Proof. Let $p, q \in \mathcal{P}^n$ where $p(x, a_0, a_1, \ldots, a_{n-1}, a_n) = \sum_{i=0}^{n} a_i x^i$ and $q(x, b_0, b_1, \ldots, b_{n-1}, b_n) = \sum_{i=0}^{n} b_i x^i$. Then

$$\begin{aligned} & p(x, a_0, a_1, \ldots, a_{n-1}, a_n) + q(x, b_0, b_1, \ldots, b_{n-1}, b_n) \\ = \ & (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0) = \sum_{i=0}^{n}(a_i + b_i)x^i \in \mathcal{P}^n. \end{aligned}$$

Hence, Axiom 1 in Definition A.1 is satisfied.

By defining the zero polynomial to be $p(x, 0, 0, \ldots, 0, 0) = 0x^n + 0x^{n-1} + \cdots 0x + 0$, the remaining nine axioms can be shown to hold as well. Hence, $\mathcal{P}^n$ is a real vector space. □

Subspaces are discussed next.

Definition A.2. [10, pp. 300] [9, pp. 162] A nonempty subset $\boldsymbol{\mathcal{S}}$ of a vector space $\boldsymbol{\mathcal{V}}$ is a *subspace* of $\boldsymbol{\mathcal{V}}$ iff

$$\forall \mathbf{v}^1, \mathbf{v}^2 \in \boldsymbol{\mathcal{S}} \Rightarrow \mathbf{v}^1 + \mathbf{v}^2 \in \boldsymbol{\mathcal{S}},$$
$$\forall \mathbf{v}^1 \in \boldsymbol{\mathcal{S}} \Rightarrow \alpha\mathbf{v}^1 \in \boldsymbol{\mathcal{S}} \quad \forall \alpha \in \mathbb{F}.$$

The subset $\{\mathbf{0}\}$ of $\boldsymbol{\mathcal{V}}$ is a subspace because $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $\alpha\mathbf{0} = \mathbf{0}\ \forall\alpha \in \mathbb{F}$, and is called the *trivial subspace*. Also, the vector space $\boldsymbol{\mathcal{V}}$ is a subspace of itself. Subspaces other than the trivial subspace and $\boldsymbol{\mathcal{V}}$ are called *proper subspaces* [10, pp. 300].

Example A.4. [10, pp. 301] Let

$$\boldsymbol{\mathcal{S}} = \left\{ \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \middle| \ v_1 = at, v_2 = bt, v_3 = ct,\ t \in \mathbb{R} \right\},$$

which consists of vectors in $\mathbb{R}^3$ lying on a straight line passing through the origin. Show that $\boldsymbol{\mathcal{S}}$ is a subspace of $\mathbb{R}^3$.

Proof. Let

$$\mathbf{v}^1 = \begin{bmatrix} at_1 \\ bt_1 \\ ct_1 \end{bmatrix} \in \boldsymbol{\mathcal{S}}, \quad \mathbf{v}^2 = \begin{bmatrix} at_2 \\ bt_2 \\ ct_2 \end{bmatrix} \in \boldsymbol{\mathcal{S}}.$$

Then,

$$\mathbf{v}^1 + \mathbf{v}^2 = \begin{bmatrix} at_1 \\ bt_1 \\ ct_1 \end{bmatrix} + \begin{bmatrix} at_2 \\ bt_2 \\ ct_2 \end{bmatrix} = \begin{bmatrix} a(t_1 + t_2) \\ b(t_1 + t_2) \\ c(t_1 + t_2) \end{bmatrix} \in \boldsymbol{\mathcal{S}},$$

and

$$\alpha\mathbf{v}^1 = \begin{bmatrix} a(\alpha t_1) \\ b(\alpha t_1) \\ c(\alpha t_1) \end{bmatrix} \in \boldsymbol{\mathcal{S}}.$$

Thus, $\boldsymbol{\mathcal{S}}$ is a subspace of $\mathbb{R}^3$. □

In fact, any line through the origin and any plane through the origin are subspaces of $\mathbb{R}^3$.

A.1.2 Linear Combinations and Span

Linear combinations of vectors and the notion of span are discussed next. These are needed to understand linear independence as well as basis and dimension, which are discussed in the next section, Section A.1.3.

Definition A.3. [10, pp. 306] Let $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n \in \boldsymbol{\mathcal{V}}$. Then any vector of the form

$$x_1\mathbf{v}^1 + x_2\mathbf{v}^2 + \ldots + x_n\mathbf{v}^n,$$

where $x_1, \ldots, x_n \in \mathbb{F}$ is called a *linear combination* of $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n$.

Example A.5. [10, pp. 306] Every polynomial in $\mathcal{P}^n$ can be written as a linear combination of $1, x, x^2, \ldots, x^n$.

Definition A.4. [10, pp. 306] The vectors $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n \in \boldsymbol{\mathcal{V}}$ are said to *span* $\boldsymbol{\mathcal{V}}$ if every vector in $\boldsymbol{\mathcal{V}}$ can be written as a linear combination of $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n$, that is, for every $\mathbf{v} \in \boldsymbol{\mathcal{V}}$, there exists $x_1, x_2, \ldots, x_n \in \mathbb{F}$ such that

$$\mathbf{v} = x_1\mathbf{v}^1 + x_2\mathbf{v}^2 + \cdots + x_n\mathbf{v}^n.$$

Example A.6. [9, pp. 165] The column matrices

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

span $\mathbb{R}^3$.

Example A.7. The column matrices

$$\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

also span $\mathbb{R}^3$.

Example A.8. [9, pp. 165] Let $\boldsymbol{\mathcal{S}} = \left\{\mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^n\right\}$ be a subspace of $\mathbb{R}^m$ (i.e., $\mathbf{a}^i \in \mathbb{R}^m$, $i = 1 \ldots n$), and let $\mathbf{A} = \left[\begin{array}{cccc} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{array}\right] \in \mathbb{R}^{m \times n}$. The following statements are equivalent.

(1) $\boldsymbol{\mathcal{S}}$ spans $\mathbb{R}^m$.

(2) $\forall \mathbf{b} \in \mathbb{R}^m$, $\exists \mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Proof. (2) $\Rightarrow$ (1) will be proven first. This is the "only if" part, or "necessity". To begin, note that

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \left[\begin{array}{cccc} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n.$$

Thus, $\mathbf{b}$ is a linear combination of the columns of $\mathbf{A}$. Because $\mathbf{b}$ is an arbitrary element of $\mathbb{R}^m$, the columns of $\mathbf{A}$ must span $\mathbb{R}^m$, that is, $\boldsymbol{\mathcal{S}}$ spans $\mathbb{R}^m$. Hence, (2) $\Rightarrow$ (1), as required.

Next, (1) $\Rightarrow$ (2) will be proven. This is the "if" part, or "sufficiency". Recall that, by definition, $\boldsymbol{\mathcal{S}}$ spans $\mathbb{R}^m$ means any $\mathbf{b} \in \mathbb{R}^m$ can be written as a linear combination of the elements of $\boldsymbol{\mathcal{S}}$,

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n, \quad x_i \in \mathbb{R}, i = 1, \ldots, n.$$

Rearranging,

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n = \left[\begin{array}{cccc} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{A}\mathbf{x}.$$

Thus, $\boldsymbol{\mathcal{S}}$ spans $\mathbb{R}^m$ implies that $\forall \mathbf{b} \in \mathbb{R}^m$, $\exists \mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$. Hence, (1) $\Rightarrow$ (2), as required. $\square$

A.1.3 Linear Independence, Basis, and Dimension

Now linear independence will be defined.

Definition A.5. [6, pp. 419] The set of vectors $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n \in \boldsymbol{\mathcal{V}}$ are *linear independent* if

$$x_1\mathbf{v}^1 + x_2\mathbf{v}^2 + \ldots + x_n\mathbf{v}^n = \mathbf{0}$$

implies $x_i = 0$, $i = 1, \ldots, n$. If the vectors are not linearly independent, they're *linearly dependent*.

Lemma A.1. [6, pp. 419] If $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n \in \boldsymbol{\mathcal{V}}$ is a linearly independent set of vectors and $\mathbf{v} \in \operatorname{span}\left\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\right\}$ then the relation

$$\mathbf{v} = x_1\mathbf{v}^1 + x_2\mathbf{v}^2 + \ldots + x_n\mathbf{v}^n$$

is unique.

Proof. Suppose not, that is, suppose

$$\mathbf{v} = y_1\mathbf{v}^1 + y_2\mathbf{v}^2 + \ldots + y_n\mathbf{v}^n$$

is another relation where $x_i \neq y_i$, $i = 1, \ldots, n$. Then,

$$\begin{aligned} \mathbf{0} &= \mathbf{v} - \mathbf{v} \\ &= x_1\mathbf{v}^1 + x_2\mathbf{v}^2 + \ldots + x_n\mathbf{v}^n - \left(y_1\mathbf{v}^1 + y_2\mathbf{v}^2 + \ldots + y_n\mathbf{v}^n\right) \\ &= (x_1 - y_1)\mathbf{v}^1 + (x_2 - y_2)\mathbf{v}^2 + \ldots + (x_n - y_n)\mathbf{v}^n, \end{aligned}$$

where $(x_i - y_i) \neq 0$, $i = 1, \ldots, n$. This contradicts the assumption of linear independence. □

Example A.9. Are

(1)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

(2)

$$\begin{bmatrix} 5 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 5 \end{bmatrix},$$

(3)

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 6 \end{bmatrix},$$

linearly independent?

Proof. (1) Yes. (2) Yes. (3) No, because the second column matrix is two times the first column matrix. □

The notion of a basis is defined next.

Definition A.6. [10, pp. 337] A finite set of vectors $\left\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\right\}$ is a *basis* for the vector space $\boldsymbol{\mathcal{V}}$ if

1. $\left\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\right\}$ is linearly independent, and
2. $\left\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\right\}$ spans $\boldsymbol{\mathcal{V}}$.

Example A.10. Are

(1)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

(2)

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

(3)

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix},$$

bases for $\mathbb{R}^2$, $\mathbb{R}^2$, and $\mathbb{R}^{2\times 2}$, respectively?

Proof. (1) Yes. (2) Yes. (3) Yes. In (1) and (2) the pairs of column matrices are linearly independent and span $\mathbb{R}^2$. Similarly, in (3) the 2×2 matrices are linearly independent and span $\mathbb{R}^{2\times 2}$. $\square$

Example A.11. [10, pp. 338] Find a basis for the plane

$$\boldsymbol{\pi} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \middle| \; 2x - y + 3z = 0 \right\}.$$

Proof. First, notice that $\boldsymbol{\pi}$ is a (real) vector space. Vectors in $\boldsymbol{\pi}$ are of the form

$$\mathbf{v} = \begin{bmatrix} x \\ 2x + 3z \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix},$$

and are thus 3×1 column matrices. The two column matrices

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$$

form a basis for $\boldsymbol{\pi}$ as they are linearly independent and span $\boldsymbol{\pi}$. $\square$

Definition A.7. [10, pp. 337] The *dimension* of the vector space $\boldsymbol{\mathcal{V}}$ is equal to the number of vectors in any basis of $\boldsymbol{\mathcal{V}}$.

Example A.12. [10, pp. 341] The polynomials $\left\{1, x, x^2, x^3, \ldots, x^n\right\}$ constitute a basis for $\mathcal{P}^n$. Thus, $\dim \mathcal{P}^n = n + 1$.

Example A.13. The dimension of $\boldsymbol{\pi}$ in Example A.11 is two, that is, $\dim \boldsymbol{\pi} = 2$.

A.1.4 The Null Space and Range of a Matrix

Definition A.8. [10, pp. 348] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. The *null space of* $\mathbf{A}$, or the *kernel of* $\mathbf{A}$, is

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}.$$

The *nullity of* $\mathbf{A}$ is $\nu(\mathbf{A}) = \dim \mathcal{N}(\mathbf{A})$.

Example A.14. Find the $\mathcal{N}(\mathbf{A})$ and $\nu(\mathbf{A})$ of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Proof. Writing out $\mathbf{A}\mathbf{x} = \mathbf{0}$ gives

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which leads to the two equations

$$x_1 + 2x_2 - x_3 = 0, \quad x_1 = -x_3.$$

Substitution of the second equation into the first gives $2x_2 - 2x_3 = 0$ or $x_2 = x_3$. Thus,

$$\mathcal{N}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

It is said that "the null space of $\mathbf{A}$ is spanned by the column matrix $[-1 \ \ 1 \ \ 1]^\mathsf{T}$", or "$[-1 \ \ 1 \ \ 1]^\mathsf{T}$ forms a basis for the null space of $\mathbf{A}$". The dimension of $\mathcal{N}(\mathbf{A})$ is 1, that is $\nu(\mathbf{A}) = 1$.

□

Definition A.9. [6, pp. 430] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. The *range of* $\mathbf{A}$, or the *image of* $\mathbf{A}$, is

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m \mid \exists \mathbf{x} \in \mathbb{R}^n \text{ such that } \mathbf{y} = \mathbf{A}\mathbf{x}\}.$$

The *rank of* $\mathbf{A}$ is $\rho(\mathbf{A}) = \dim\mathcal{R}(\mathbf{A})$.

The range of $\mathbf{A}$ is often called the column space of $\mathbf{A}$ because it is the space spanned by the columns of $\mathbf{A}$. Similarly, $\mathcal{R}(\mathbf{A}^\mathsf{T})$ is often called the row space of $\mathbf{A}$ because it is the space spanned by the columns of $\mathbf{A}^\mathsf{T}$ which are the same as the rows of $\mathbf{A}$.

The rank of a matrix $\mathbf{A}$ is also equal to *a*) the number of linearly independent columns of $\mathbf{A}$, and *b*) the number of linearly independent rows of $\mathbf{A}$. Additionally, $\rho(\mathbf{A}) \le \min(m, n)$.

Example A.15. Find a basis for both $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A})$ as well as $\rho(\mathbf{A})$ and $\nu(\mathbf{A})$ where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Proof. The fourth column $\mathbf{A}$ is a linear combination of the first and third columns. Therefore, a basis for $\mathcal{R}(\mathbf{A})$ is

$$\mathcal{R}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^3.$$

It is said that "the range of $\mathbf{A}$ is spanned by the column matrices $[1 \ \ 0 \ \ 0]^\mathsf{T}$, $[0 \ \ 1 \ \ 0]^\mathsf{T}$, $[0 \ \ 0 \ \ 1]^\mathsf{T}$", or "$[1 \ \ 0 \ \ 0]^\mathsf{T}$, $[0 \ \ 1 \ \ 0]^\mathsf{T}$, $[0 \ \ 0 \ \ 1]^\mathsf{T}$ is a basis for the range of $\mathbf{A}$". The rank of $\mathbf{A}$ is the number of linearly independent columns of $\mathbf{A}$ which is 3, thus, $\rho(\mathbf{A}) = 3$.

Next, the null space of $\mathbf{A}$ is the subspace of all vectors satisfying $\mathbf{A}\mathbf{x} = \mathbf{0}$. Thus,

$$\mathcal{N}(\mathbf{A}) = \text{span}\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

It is said that "the null space of $\mathbf{A}$ is spanned by the column matrix $\frac{1}{\sqrt{3}}[-1 \ \ 0 \ \ -1 \ \ 1]^\mathsf{T}$". The nullity is the dimension of $\mathcal{N}(\mathbf{A})$ which is 1, that is, $\nu(\mathbf{A}) = 1$. Note that $[-1 \ \ 0 \ \ -1 \ \ 1]^\mathsf{T}$ has been normalized; any scalar multiple of $[-1 \ \ 0 \ \ -1 \ \ 1]^\mathsf{T}$, such as $\frac{1}{\sqrt{3}}[-1 \ \ 0 \ \ -1 \ \ 1]^\mathsf{T}$, is said to span $\mathcal{N}(\mathbf{A})$. □

Theorem A.1. [10, pp. 356] [6, pp. 431] Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then $\rho(\mathbf{A}) + \nu(\mathbf{A}) = n$.

Theorem A.1 is called the Rank-Nullity Theorem or Sylvester's Law of Nullity.

Consider the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. The system $\mathbf{Ax} = \mathbf{b}$ may have [9, pp. 53-54]

(1) one solution (a unique solution),

(2) infinitely many solutions, or

(3) no solution.

Definition A.10. [9, pp. 53-54] The system $\mathbf{Ax} = \mathbf{b}$ is *consistent* if it has at least one solution. The system $\mathbf{Ax} = \mathbf{b}$ is *inconsistent* if it does not have a solution.

Theorem A.2. [10, pp. 358] Consider the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. The following statements are equivalent.

(1) $\mathbf{Ax} = \mathbf{b}$ has at least one solution.

(2) $\mathbf{b} \in \mathcal{R}(\mathbf{A})$.

(3) $\rho(\mathbf{A}) = \rho([\mathbf{A}\ \mathbf{b}])$.

Proof. (1) $\Rightarrow$ (2) will be proven first. If $\mathbf{Ax} = \mathbf{b}$ has one solution then

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n, \quad x_i \in \mathbb{R}, i = 1, \ldots, n,$$

where $\mathbf{b}$ is written as a linear combination of the columns of $\mathbf{A}$. Thus, $\mathbf{b} \in \mathcal{R}(\mathbf{A})$, and (1) $\Rightarrow$ (2).

Next (2) $\Rightarrow$ (1) will be shown to hold. If $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ then $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$, that is

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n, \quad x_i \in \mathbb{R}, i = 1, \ldots, n.$$

This can be rewritten as

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n = \begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{Ax}.$$

Hence, (2) $\Rightarrow$ (1).

Now (3) $\Rightarrow$ (2) will be shown to hold. If $\mathbf{A}$ and $[\mathbf{A}\ \mathbf{b}]$ have the same rank then $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$, meaning

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n, \quad x_i \in \mathbb{R}, i = 1, \ldots, n.$$

Thus, $\mathbf{b} \in \mathcal{R}(\mathbf{A})$, and (3) $\Rightarrow$ (2).

Last (2) $\Rightarrow$ (3) will be shown to hold. If $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ then

$$\mathbf{b} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + \ldots + x_n\mathbf{a}^n, \quad x_i \in \mathbb{R}, i = 1, \ldots, n.$$

The rank of a matrix is equal to the number of linearly independent columns. Because $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$ the matrix $[\mathbf{A}\ \mathbf{b}]$ and $\mathbf{A}$ have the same rank. Thus, (2) $\Rightarrow$ (3).

From (1) $\Leftrightarrow$ (2) and (2) $\Leftrightarrow$ (3) it follows that (1) $\Leftrightarrow$ (3).

$\square$

Theorem A.3. Consider the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. The following statements are equivalent.

(1) The solution to $\mathbf{Ax} = \mathbf{b}$ is unique (i.e., $\mathbf{x}$ is unique).

(2) $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$ (i.e., $\nu(\mathbf{A}) = 0$).

(3) $\rho(\mathbf{A}) = n \le m$.

Proof. The proof of (1) $\Rightarrow$ (2) will be shown first. To prove (1) $\Rightarrow$ (2), the contrapositive will be proven, that is, $\neg$(2) $\Rightarrow$ $\neg$(1) which reads "not (2) implies not (1)". To begin, $\neg$(2) means $\nu(\mathbf{A}) \neq 0$, which in turn means $\exists \mathbf{n} \in \mathcal{N}(\mathbf{A})$ (i.e., $\exists \mathbf{n} \in \mathbb{R}^n$ such that $\mathbf{An} = \mathbf{0}$). Next, consider $\mathbf{Ap} = \mathbf{b}$ where $\mathbf{p}$ is a particular solution. Another solution is

$$\mathbf{x} = \mathbf{p} + \alpha\mathbf{n},$$

where $\alpha \in \mathbb{R}$ is arbitrary. Given that α is arbitrary and

$$\mathbf{Ax} = \mathbf{A}(\mathbf{p} + \alpha\mathbf{n}) = \mathbf{Ap} + \alpha\mathbf{0} = \mathbf{b},$$

there are an infinite number of solutions $\mathbf{x}$ to $\mathbf{Ax} = \mathbf{b}$ of the form $\mathbf{x} = \mathbf{p} + \alpha\mathbf{n}$. Hence $\neg$(2) $\Rightarrow$ $\neg$(1). This is equivalent to (1) $\Rightarrow$ (2).

Next, (2) $\Rightarrow$ (1) will be proven. Again, the contrapositive will be proven, that is, $\neg$(1) $\Rightarrow$ $\neg$(2). The statement $\neg$(1) means the solution to $\mathbf{Ax} = \mathbf{b}$ is not unique. Consider two different solutions, $\mathbf{x}^1 \in \mathbb{R}^n$ and $\mathbf{x}^2 \in \mathbb{R}^n$ such that

$$\mathbf{Ax}^1 = \mathbf{b} \quad \text{and} \quad \mathbf{Ax}^2 = \mathbf{b}.$$

Subtracting the first solution from the second solution yields

$$\mathbf{A}\underbrace{\left(\mathbf{x}^2 - \mathbf{x}^1\right)}_{\mathbf{n}} = \mathbf{0}.$$

Therefore, $\mathbf{n} \in \mathcal{N}(\mathbf{A})$, and thus $\nu(\mathbf{A}) \neq 0$. Therefore, $\neg$(1) $\Rightarrow$ $\neg$(2), or equivalently, (2) $\Rightarrow$ (1).

Next, (2) $\Leftrightarrow$ (3) will be shown to hold. First, recall that $\rho(\mathbf{A}) \le \min(m, n)$. If $\rho(\mathbf{A}) = n$ then it must be that $\rho(\mathbf{A}) = n \le m$.

Now (2) $\Rightarrow$ (3) will be shown to hold. If (2) holds (i.e., $\nu(\mathbf{A}) = 0$) then, using the Rank Nullity Theorem (Theorem A.1),

$$\rho(\mathbf{A}) + 0 = n.$$

implying (3).

Last (3) $\Rightarrow$ (2) will be proven. The Rank Nullity Theorem (Theorem A.1) states that

$$\rho(\mathbf{A}) + \nu(\mathbf{A}) = n.$$

Thus, $\nu(\mathbf{A}) = 0$ because $\rho(\mathbf{A}) = n$. Thus, (3) $\Rightarrow$ (2).

From (1) $\Leftrightarrow$ (2) and (2) $\Leftrightarrow$ (3) it follows that (1) $\Leftrightarrow$ (3).

$\square$

Example A.16. Consider

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Does $\mathbf{Ax} = \mathbf{b}$ have a solution? Is it unique?

Proof. The matrix $\mathbf{A}$ is 3×2, thus $m = 3$ and $n = 2$. The rank of $\mathbf{A}$ is 2, that is $\rho(\mathbf{A}) = 2$, because the columns of $\mathbf{A}$ are linearly independent. The rank of $[\mathbf{A}\ \mathbf{b}]$ is also 2, that is $\rho([\mathbf{A}\ \mathbf{b}]) = 2$, because $\mathbf{b}$ is a linear combination of the columns of $\mathbf{A}$. Specifically, $\mathbf{b}$ is equal to $\mathbf{A}$'s first column plus $\mathbf{A}$'s second column, and as such the solution $\mathbf{x}$ to $\mathbf{Ax} = \mathbf{b}$ is $\mathbf{x} = [1\ \ 1]^\mathsf{T}$. Because $\rho(\mathbf{A}) = \rho([\mathbf{A}\ \mathbf{b}]) = 2$ and $\rho(\mathbf{A}) = 2 \leq 3$ the solution is unique. $\square$

Example A.17. Consider

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}.$$

Does $\mathbf{Ax} = \mathbf{b}$ have a solution? How many solutions? Does $\mathbf{Ax} = \mathbf{c}$ have a solution?

Proof. The matrix $\mathbf{A}$ is 3×4, thus $m = 3$ and $n = 4$. Denote the four columns of $\mathbf{A}$ as $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$, $\mathbf{a}^4$. Note that $\mathbf{a}^1 + \mathbf{a}^2 = \mathbf{a}^3$ and $\mathbf{a}^1 - \mathbf{a}^2 = \mathbf{a}^4$. Thus, the number of linearly independent columns is two. Hence, $\rho(\mathbf{A}) = 2$. A basis for the range of $\mathbf{A}$ is

$$\mathcal{R}(\mathbf{A}) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

The rank of $[\mathbf{A}\ \mathbf{b}]$ is two. Notice that $2\mathbf{a}^1 + \mathbf{a}^2 = \mathbf{b}$. Thus, a solution exists because $\rho(\mathbf{A}) = \rho([\mathbf{A}\ \mathbf{b}]) = 2$. However, the solution is not unique because $\rho(\mathbf{A}) = 2 \leq 4$. There are infinitely many solutions.

Concerning $\mathbf{Ax} = \mathbf{c}$, the rank of $[\mathbf{A}\ \mathbf{c}]$ is three. The column matrix $\mathbf{c}$ is not a linear combination of the columns of $\mathbf{A}$. Because $\rho(\mathbf{A}) \neq \rho([\mathbf{A}\ \mathbf{b}])$, no solution exists.

$\square$

Theorem A.4. [10, pp. 360] Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. The following statements are equivalent.

1. $\mathbf{A}$ is invertible.
2. The only solution to $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.
3. The solution to $\mathbf{Ax} = \mathbf{b}$ is unique (i.e., $\mathbf{x}$ is unique).
4. The rows and columns of $\mathbf{A}$ are linearly independent.
5. $\det \mathbf{A} \neq 0$.
6. $\rho(\mathbf{A}) = n$.
7. $\nu(\mathbf{A}) = 0$.

If one of the above fails to hold then $\mathbf{Ax} = \mathbf{b}$ has either no solution of an infinite number of solutions; it has an infinite number of solutions iff $\rho(\mathbf{A}) = \rho([\mathbf{A}\ \mathbf{b}])$.

A.1.5 The Inverse, Determinant, and Trace of a Matrix

Definition A.11. [6, pp. 414] Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$. The matrix $\mathbf{A}$ is *invertible* or *nonsingular* if $\exists \mathbf{J} \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{AJ} = \mathbf{JA} = \mathbf{1}.$$

The matrix $\mathbf{J}$ is called the *inverse of* $\mathbf{A}$ and is often written as $\mathbf{A}^{-1}$.

Theorem A.5. Consider $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Then

1. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$,
2. $(\mathbf{A}^{-1})^{\mathsf{T}} = (\mathbf{A}^{\mathsf{T}})^{-1} = \mathbf{A}^{-\mathsf{T}}$.

Definition A.12. [10, pp. 172] Consider $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

The determinant of $\mathbf{A}$ is

$$\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21},$$

Definition A.13. [10, pp. 173] Consider $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

The determinant of $\mathbf{A}$ is

$$\det \mathbf{A} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

Definition A.14. [10, pp. 176] Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$ where

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

The $1-k$ *minor* of $\mathbf{A}$, denoted $\mathbf{M}_{1k} \in \mathbb{R}^{(n-1)\times(n-1)}$, is constructed by deleting the first row and k^{th} column of $\mathbf{A}$. The *determinant* of $\mathbf{A}$ is

$$\det \mathbf{A} = \sum_{k=1}^{n} a_{1k}\mathbf{A}_{1k},$$

where $\mathbf{A}_{1k} = (-1)^{1+k} \det \mathbf{M}_{1k}$ is the $1-k$ cofactor.

Theorem A.6. [10, pp. 358, 190, 192, 195, 211] Consider $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Then

1. $\det(\mathbf{AB}) = \det \mathbf{A} \det \mathbf{B}$,
2. $\det \mathbf{A} = \det \mathbf{A}^{\mathsf{T}}$,
3. $\det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}}$ for any nonsingular $\mathbf{A}$,
4. if any row or column of $\mathbf{A}$ is zero then $\det \mathbf{A} = 0$,

5. if $\mathbf{A}$ is rank deficient (i.e., $\rho(\mathbf{A}) < n$) then $\det \mathbf{A} = 0$,
6. if $\det \mathbf{A} \neq 0$ then $\rho(\mathbf{A}) = n$ and $\mathbf{A}^{-1}$ exists.

Definition A.15. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$. The *trace* of $\mathbf{A}$ is

$$\mathrm{tr}\mathbf{A} = \sum_{i=1}^{n} a_{ii}.$$

Theorem A.7. Consider $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$, and $\mathbf{C} \in \mathbb{R}^{p\times m}$. Then

$$\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB}) = \mathrm{tr}(\mathbf{BCA}).$$

Theorem A.8. Trace is linear, that is

$$\begin{aligned} \mathrm{tr}(\mathbf{A}+\mathbf{B}) &= \mathrm{tr}\mathbf{A} + \mathrm{tr}\mathbf{B}, \quad \forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}, \\ \mathrm{tr}(\alpha\mathbf{A}) &= \alpha\mathrm{tr}\mathbf{A}, \quad \forall \mathbf{A} \in \mathbb{R}^{n\times n},\ \alpha \in \mathbb{R}. \end{aligned}$$

Proof. First, superposition will be shown to hold. Consider

$$\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \sum_{i=1}^{n}(a_{ii}+b_{ii}),$$

and

$$\mathrm{tr}\mathbf{A} + \mathrm{tr}\mathbf{B} = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \sum_{i=1}^{n}(a_{ii}+b_{ii}).$$

Thus, trace satisfies superposition.

Next, scaling will be shown to hold. Consider

$$\mathrm{tr}(\alpha\mathbf{A}) = \sum_{i=1}^{n} \alpha a_{ii} = \alpha \sum_{i=1}^{n} a_{ii} = \alpha\mathrm{tr}\mathbf{A}.$$

Thus, trace satisfies scaling.

□

A.1.6 Orthongonality and Orthogonal Complements

Definition A.16. [6, pp. 424] [9, pp. 286]

1. [9, pp. 286] A *inner product* on a vector space is a function $\langle \cdot, \cdot \rangle : \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{S}} \to \mathbb{R}$ satisfying the following four properties.

   - Let $\mathbf{x} \in \boldsymbol{\mathcal{S}}$. $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$ with $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
   - Let $\mathbf{x}^1, \mathbf{x}^2 \in \boldsymbol{\mathcal{S}}$ and $\alpha \in \mathbb{R}$. $\left\langle \mathbf{x}^1, \alpha\mathbf{x}^2 \right\rangle = \left\langle \alpha\mathbf{x}^1, \mathbf{x}^2 \right\rangle = \alpha \left\langle \mathbf{x}^1, \mathbf{x}^2 \right\rangle$.
   - Let $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3 \in \boldsymbol{\mathcal{S}}$. $\left\langle \mathbf{x}^1, \mathbf{x}^2 + \mathbf{x}^3 \right\rangle = \left\langle \mathbf{x}^1, \mathbf{x}^2 \right\rangle + \left\langle \mathbf{x}^1, \mathbf{x}^3 \right\rangle$.
   - Let $\mathbf{x}^1, \mathbf{x}^2 \in \boldsymbol{\mathcal{S}}$. $\left\langle \mathbf{x}^1, \alpha\mathbf{x}^2 \right\rangle = \left\langle \mathbf{x}^2, \alpha\mathbf{x}^1 \right\rangle^{\mathsf{H}}$ where $(\cdot)^{\mathsf{H}}$ denotes the complex conjugate or Hermitian. (When $\boldsymbol{\mathcal{S}}$ is a real vector space, $\left\langle \mathbf{x}^1, \alpha\mathbf{x}^2 \right\rangle = \left\langle \mathbf{x}^2, \alpha\mathbf{x}^1 \right\rangle$.)

   Any real or complex vector space that is equipped with an inner product is called an inner product space.

2. [9, pp. 286] The *standard inner product* associated with $\mathbb{R}^n$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\mathsf{T}\mathbf{y} = \sum_{i=1}^{n} x_i y_i,$$

where $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Note that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\mathsf{T}\mathbf{y} = \mathbf{y}^\mathsf{T}\mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$.

3. [9, pp. 286] The *standard inner product for matrices* associated with $\mathbb{R}^{m\times n}$ is

$$\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}^\mathsf{T}\mathbf{B})$$

where $\langle \cdot, \cdot \rangle : \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \to \mathbb{R}$. Note that $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}^\mathsf{T}\mathbf{B}) = \mathrm{tr}(\mathbf{B}^\mathsf{T}\mathbf{A}) = \langle \mathbf{B}, \mathbf{A} \rangle$ because $\mathrm{tr}\mathbf{Q} = \mathrm{tr}\mathbf{Q}^\mathsf{T}$. See Definition A.15.

4. [9, pp. 288] Let $\boldsymbol{\mathcal{S}}$ be an inner product space with inner product $\langle \cdot, \cdot \rangle : \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{S}} \to \mathbb{R}$. Then

$$\|\star\| = \sqrt{\langle \star, \star \rangle}$$

defines a norm on $\boldsymbol{\mathcal{S}}$.

5. A set of vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k \in \mathbb{R}^n$ is an *orthogonal set* if $\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle = 0, \forall i \neq j$.

6. A set of vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k \in \mathbb{R}^n$ is an *orthonormal set* if $\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle = 0$, $\forall i \neq j$, and $\left\|\mathbf{x}^i\right\| = \sqrt{\langle \mathbf{x}^i, \mathbf{x}^i \rangle} = 1$, $i = 1, 2, \ldots, k$.

7. Let $\boldsymbol{\mathcal{S}}$ be a subspace of $\mathbb{R}^n$. A set of vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k \in \mathbb{R}^n$ is an *orthogonal basis* for $\boldsymbol{\mathcal{S}}$ if it is an orthonormal set and a basis for $\boldsymbol{\mathcal{S}}$.

8. Let $\boldsymbol{\mathcal{S}}$ be a subspace of $\mathbb{R}^n$. The *orthogonal complement* of $\boldsymbol{\mathcal{S}}$ is defined as

$$\boldsymbol{\mathcal{S}}^\perp = \left\{ \mathbf{y} \in \mathbb{R}^n | \langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad \forall \mathbf{x} \in \boldsymbol{\mathcal{S}} \right\}.$$

9. [9, pp. 299] If $\boldsymbol{\mathcal{O}} = \left\{ \mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^n \right\}$ is an orthonormal basis for an inner product space $\boldsymbol{\mathcal{V}}$, then each $\mathbf{x} \in \boldsymbol{\mathcal{V}}$ can be written as

$$\mathbf{x} = \left\langle \mathbf{u}^1, \mathbf{x} \right\rangle \mathbf{u}^1 + \left\langle \mathbf{u}^2, \mathbf{x} \right\rangle \mathbf{u}^2 + \cdots + \left\langle \mathbf{u}^n, \mathbf{x} \right\rangle \mathbf{u}^n = \sum_{i=1}^{n} \left\langle \mathbf{u}^i, \mathbf{x} \right\rangle \mathbf{u}^i.$$

This is called a *Fourier expansion* of $\mathbf{x}$.

Theorem A.9. [6, pp. 434] For $\mathbf{A} \in \mathbb{R}^{m\times n}$, $[\mathcal{R}(\mathbf{A})]^\perp = \mathcal{N}(\mathbf{A}^\mathsf{T})$ and $[\mathcal{N}(\mathbf{A})]^\perp = \mathcal{R}(\mathbf{A}^\mathsf{T})$.

Proof. To show $[\mathcal{R}(\mathbf{A})]^\perp = \mathcal{N}(\mathbf{A}^\mathsf{T})$ it must be shown that $[\mathcal{R}(\mathbf{A})]^\perp \subset \mathcal{N}(\mathbf{A}^\mathsf{T})$ and $\mathcal{N}(\mathbf{A}^\mathsf{T}) \subset [\mathcal{R}(\mathbf{A})]^\perp$.

First, it must be shown that $[\mathcal{R}(\mathbf{A})]^\perp$ is a subset of $\mathcal{N}(\mathbf{A}^\mathsf{T})$, that is $[\mathcal{R}(\mathbf{A})]^\perp \subset \mathcal{N}(\mathbf{A}^\mathsf{T})$. Let $\mathbf{y} \in [\mathcal{R}(\mathbf{A})]^\perp$ and $\mathbf{z} \in \mathcal{R}(\mathbf{A})$ be two arbitrary vectors. Then, $\langle \mathbf{y}, \mathbf{z} \rangle = 0$ because $\mathbf{y}$ and $\mathbf{z}$ are orthogonal. Write $\mathbf{z}$ as $\mathbf{z} = \mathbf{A}\mathbf{x}$ where $\mathbf{x} \in \mathbb{R}^n$ is arbitrary; clearly $\mathbf{z} \in \mathcal{R}(\mathbf{A})$. Then, $\langle \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{A}\mathbf{x} \rangle = 0$ which can be written $\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{x} = (\mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}\mathbf{x} = 0$. Because $\mathbf{x}$ is arbitrary, $\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0}$, or $\mathbf{y} \in \mathcal{N}(\mathbf{A}^\mathsf{T})$. Thus, $[\mathcal{R}(\mathbf{A})]^\perp \subset \mathcal{N}(\mathbf{A}^\mathsf{T})$.

Next, it must be shown that $\mathcal{N}(\mathbf{A}^\mathsf{T})$ is a subset of $[\mathcal{R}(\mathbf{A})]^\perp$, that is $\mathcal{N}(\mathbf{A}^\mathsf{T}) \subset [\mathcal{R}(\mathbf{A})]^\perp$. Let $\mathbf{y} \in \mathcal{N}(\mathbf{A}^\mathsf{T})$ and $\mathbf{x} \in \mathbb{R}^n$ be both arbitrary. Then, $\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{0}$ and $\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{y} = 0$. Defining $\mathbf{z} = \mathbf{A}\mathbf{x}$, write $\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{z}, \mathbf{y} \rangle = 0$. Given that $\mathbf{z} \in \mathcal{R}(\mathbf{A})$ it follows that $\mathbf{y} \in [\mathcal{R}(\mathbf{A})]^\perp$. Hence, $\mathcal{N}(\mathbf{A}^\mathsf{T}) \subset [\mathcal{R}(\mathbf{A})]^\perp$.

Proving $[\mathcal{N}(\mathbf{A})]^\perp = \mathcal{R}(\mathbf{A}^\mathsf{T})$ follows in a similar manner. $\square$

A.1.7 Gram-Schmidt Orthogonalization

This section is based on [9, pp. 307].

Given an arbitrary basis $\boldsymbol{\mathcal{B}} = \left\{\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^n\right\}$ for the n-dimensional inner product space $\boldsymbol{\mathcal{S}}$ with inner product $\langle \star, \star \rangle$ and norm $\|\star\| = \sqrt{\langle \star, \star \rangle}$, how can an orthonormal basis $\boldsymbol{\mathcal{O}} = \left\{\mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^n\right\}$ be constructed for $\boldsymbol{\mathcal{S}}$? To be clear, given $\mathbf{v} \in \boldsymbol{\mathcal{S}}$, $\mathbf{v}$ can be written as a linear combination of $\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^n$, that is, $\mathbf{v} = \alpha_1 \mathbf{x}^1 + \alpha_2 \mathbf{x}^2 + \cdots + \alpha_2 \mathbf{x}^n$. However, it would be nice if $\mathbf{v}$ could be written as a linear combination of some orthonormal basis of $\boldsymbol{\mathcal{S}}$, that is, write $\mathbf{v}$ as $\mathbf{v} = \beta_1 \mathbf{u}^1 + \beta_2 \mathbf{u}^2 + \cdots + \beta_n \mathbf{u}^n$, where $\mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^n$ are linearly independent, span $\boldsymbol{\mathcal{S}}$, and are orthonormal.

For simplicity, assume $\boldsymbol{\mathcal{S}}$ is a real inner product space. The orthogonal basis $\boldsymbol{\mathcal{O}}$ will be sequentially constructed using $\boldsymbol{\mathcal{B}}$ because $\boldsymbol{\mathcal{B}}$ is an basis for $\boldsymbol{\mathcal{S}}$ already. Let $\boldsymbol{\mathcal{O}}^k = \left\{\mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^k\right\}$ be an orthonormal basis for $\boldsymbol{\mathcal{S}}^k$, and $\boldsymbol{\mathcal{B}}^k = \left\{\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^k\right\}$ be a basis for $\boldsymbol{\mathcal{S}}^k$. It is desired to construct $\boldsymbol{\mathcal{O}}^{k+1} = \left\{\mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^{k+1}\right\}$, an orthonormal basis for $\boldsymbol{\mathcal{S}}^{k+1}$, using $\boldsymbol{\mathcal{O}}^k$ and $\boldsymbol{\mathcal{B}}^{k+1} = \left\{\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^{k+1}\right\}$, which is already a basis for $\boldsymbol{\mathcal{S}}^{k+1}$. Also, for $k = 1$, $\mathbf{u}^1 = \mathbf{x}^1 / \left\|\mathbf{x}^1\right\|$. Note that the Fourier expansion of $\mathbf{x}^{k+1}$ in terms of $\boldsymbol{\mathcal{O}}^{k+1}$ is

$$\mathbf{x}^{k+1} = \sum_{i=1}^{k+1} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i.$$

This can be written as

$$\begin{aligned}
\mathbf{x}^{k+1} &= \left\langle \mathbf{u}^{k+1}, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^{k+1} + \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i, \\
\mathbf{u}^{k+1} &= \frac{\mathbf{x}^{k+1} - \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i}{\left\langle \mathbf{u}^{k+1}, \mathbf{x}^{k+1} \right\rangle}. \qquad \text{(A.1)}
\end{aligned}$$

The norm of $\mathbf{u}^{k+1}$ is equal to 1 because it is an element of the orthonormal basis $\boldsymbol{\mathcal{O}}^{k+1}$, that is, $\left\|\mathbf{u}^{k+1}\right\| = 1$. As such,

$$\begin{aligned}
1 &= \left\| \frac{\mathbf{x}^{k+1} - \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i}{\left\langle \mathbf{u}^{k+1}, \mathbf{x}^{k+1} \right\rangle} \right\|, \\
\left\langle \mathbf{u}^{k+1}, \mathbf{x}^{k+1} \right\rangle &= \left\| \mathbf{x}^{k+1} - \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i \right\|. \qquad \text{(A.2)}
\end{aligned}$$

Note that the expression for $\left\langle \mathbf{u}^{k+1}, \mathbf{x}^{k+1} \right\rangle$ in Equation (A.2) depends only on $\mathbf{u}^i$, $i = 1, 2, \ldots k$, and not $\mathbf{u}^{k+1}$. Let

$$\nu_{k+1} = \left\| \mathbf{x}^{k+1} - \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i \right\|,$$

so that Equation (A.1) can be written as

$$\mathbf{u}^{k+1} = \frac{\mathbf{x}^{k+1} - \sum_{i=1}^{k} \left\langle \mathbf{u}^i, \mathbf{x}^{k+1} \right\rangle \mathbf{u}^i}{\nu_{k+1}}.$$

Computing $\mathbf{u}^1, \mathbf{u}^2, \cdots, \mathbf{u}^n$ in this way, where for $k = 1$, $\mathbf{u}^1 = \mathbf{x}^1 / \left\|\mathbf{x}^1\right\|$, is referred to as the Gram-Schmidt Orthogonalization Procedure.

A.1.8 The QR Factorization and Finding a Solution to $\mathbf{Ax} = \mathbf{b}$

This section is based on [9, pp. 310-314].

Gram-Schmidt orthogonalization can be used to find a solution of the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. In the special case where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\rho(\mathbf{A}) = n$ (i.e., $\mathbf{A}$ is full rank), it is tempting to solve for $\mathbf{x}$ by inverting $\mathbf{A}$, that is, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Not only is inverting $\mathbf{A}$ computationally demanding and completely impractical in most real-world situations, a solution to $\mathbf{Ax} = \mathbf{b}$ cannot be found via matrix inversion if $\mathbf{A}$ is not square.

The QR Factorization

It is possible to efficiently find a solution to $\mathbf{Ax} = \mathbf{b}$ when $\mathbf{A} \in \mathbb{R}^{m\times n}$ has $\rho(\mathbf{A}) = n \leq m$ (i.e., $\mathbf{A}$ has full column rank meaning the columns of $\mathbf{A}$ are linearly independent) using the so-called QR factorization. Let $\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{bmatrix} \in \mathbb{R}^{m\times n}$. The QR factorization boils down to using the columns of $\mathbf{A}$, $\left\{\mathbf{a}^1, \mathbf{a}^2, \cdots, \mathbf{a}^n\right\}$, which are linearly independent and span $\mathcal{R}(\mathbf{A})$, to form an orthonormal basis $\left\{\mathbf{q}^1, \mathbf{q}^2, \cdots, \mathbf{q}^n\right\}$ for $\mathcal{R}(\mathbf{A})$. Each $\mathbf{q}^i$, $i = 1, 2, \ldots, n$, is computed using the Gram-Schmidt procedure. Specifically,

$$\mathbf{q}^1 = \frac{\mathbf{a}^1}{\nu_1}, \quad \mathbf{q}^k = \frac{\mathbf{a}^k - \sum_{i=1}^{k-1} \left\langle \mathbf{q}^i, \mathbf{a}^k \right\rangle \mathbf{q}^i}{\nu_k}, \quad k = 2, 3, \ldots, n,$$

where $\nu_1 = \left\|\mathbf{a}^1\right\|$ and $\nu_k = \left\|\mathbf{a}^k - \sum_{i=1}^{k-1} \left\langle \mathbf{q}^i, \mathbf{a}^k \right\rangle \mathbf{q}^i\right\|$, $k = 2, 3, \ldots, n$. Notice that it is possible to write each $\mathbf{a}^i$ in terms of $\mathbf{q}^i$ via

$$\begin{aligned}
\mathbf{a}^1 &= \nu_1 \mathbf{q}^1, \\
\mathbf{a}^k &= \sum_{i=1}^{k-1} \left\langle \mathbf{q}^i, \mathbf{a}^k \right\rangle \mathbf{q}^i + \mathbf{q}^k \nu_k \\
&= \left\langle \mathbf{q}^1, \mathbf{a}^k \right\rangle \mathbf{q}^1 + \left\langle \mathbf{q}^2, \mathbf{a}^k \right\rangle \mathbf{q}^2 + \cdots + \left\langle \mathbf{q}^{k-1}, \mathbf{a}^k \right\rangle \mathbf{q}^{k-1} + \mathbf{q}^k \nu_k, \quad k = 2, 3, \ldots, n.
\end{aligned} \tag{A.3}$$

Equation (A.3) can be written in matrix form as

$$\underbrace{\begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \cdots & \mathbf{a}^n \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} \mathbf{q}^1 & \mathbf{q}^2 & \cdots & \mathbf{q}^n \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} \nu_1 & \left\langle \mathbf{q}^1, \mathbf{a}^2 \right\rangle & \left\langle \mathbf{q}^1, \mathbf{a}^3 \right\rangle & \left\langle \mathbf{q}^1, \mathbf{a}^4 \right\rangle & \cdots & \left\langle \mathbf{q}^1, \mathbf{a}^n \right\rangle \\ 0 & \nu_2 & \left\langle \mathbf{q}^2, \mathbf{a}^3 \right\rangle & \left\langle \mathbf{q}^2, \mathbf{a}^4 \right\rangle & \cdots & \left\langle \mathbf{q}^2, \mathbf{a}^n \right\rangle \\ 0 & 0 & \nu_3 & \left\langle \mathbf{q}^3, \mathbf{a}^3 \right\rangle & \cdots & \left\langle \mathbf{q}^3, \mathbf{a}^n \right\rangle \\ 0 & 0 & 0 & \nu_4 & \cdots & \left\langle \mathbf{q}^4, \mathbf{a}^n \right\rangle \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \nu_n \end{bmatrix}}_{\mathbf{R}} \tag{A.4}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{Q} \in \mathbb{R}^{m\times n}$, $\mathbf{R} \in \mathbb{R}^{n\times n}$. Equation (A.4) is the QR factorization of $\mathbf{A}$. The columns of $\mathbf{Q}$ are an orthonormal basis for $\mathcal{R}(\mathbf{A})$, and as such

$$\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{1},$$

which is equivalent to

$$\mathbf{q}^{ij\mathsf{T}}\mathbf{q}^{ij} = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{when } i \neq j, \end{cases}, \quad i, j = 1, 2, \ldots, n.$$

The matrix $\mathbf{R}$ is an upper-triangular matrix.

How is a QR factorization used to solve $\mathbf{Ax} = \mathbf{b}$?

This section is based on [9, pp. 312-314].

Consider the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, and $\rho(\mathbf{A}) = n \leq m$ meaning that $\mathbf{A}$ has linearly independent columns. Using a QR factorization of $\mathbf{A}$,

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{QRx} = \mathbf{b}.$$

Premultiplying by $\mathbf{Q}^\mathsf{T}$ and using the fact that $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{1}$, it follows that

$$\mathbf{Q}^\mathsf{T}\mathbf{QRx} = \mathbf{Q}^\mathsf{T}\mathbf{b} \quad \Leftrightarrow \quad \mathbf{Rx} = \mathbf{Q}^\mathsf{T}\mathbf{b}.$$

From Theorem A.3, $\mathbf{A} \in \mathbb{R}^{m\times n}$ has a unique solutions when $\rho(\mathbf{A}) = n \leq m$, and as such, the solution $\mathbf{x}$ is easily found via *backward substitution*. First, let $\mathbf{s} = \mathbf{Q}^\mathsf{T}\mathbf{b}$, which is easy to compute. Next, notice that $\mathbf{Rx} = \mathbf{Q}^\mathsf{T}\mathbf{b} = \mathbf{s}$ is of the form

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n-1} & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n-1} & r_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & r_{n-1n-1} & r_{n-1n} \\ 0 & 0 & \cdots & 0 & r_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n-1} \\ s_n \end{bmatrix}$$

Back substitution means first solving for $x_n = s_n/r_{nn}$, then solving for $x_{n-1} = (s_{n-1} - r_{n-1n}x_n)/r_{n-1n-1}$, and so on. In general, and with $x_n = s_n/r_{nn}$,

$$x_i = \frac{1}{r_{ii}}\left(s_i - \sum_{k=i+1}^{n} r_{ik}x_k\right), \quad i = n-1, n-2, \cdots, 1.$$

Classic Least Squares

This section is based on [9, pp. 312-314].

Consider the objective function

$$\begin{aligned} J(\mathbf{x}) &= \tfrac{1}{2}(\mathbf{Ax} - \mathbf{b})^\mathsf{T}(\mathbf{Ax} - \mathbf{b}) \\ &= \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{Ax} - \mathbf{b}^\mathsf{T}\mathbf{Ax} + \tfrac{1}{2}\mathbf{b}^\mathsf{T}\mathbf{b} \end{aligned}$$

to be minimized where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$. Defining the *error* to be $\boldsymbol{\rho} = \mathbf{Ax} - \mathbf{b}$ the objective function can be written as

$$J(\mathbf{x}) = \tfrac{1}{2}\boldsymbol{\rho}^\mathsf{T}\boldsymbol{\rho} = \sum_{i=1}^{m} \rho_i^2.$$

The problem at hand is to find an optimal value of $\mathbf{x}$, call it $\mathbf{x}^\star$, such that the error is minimized in a least-squares sense. The minimizing solution,

$$\mathbf{x}^\star = \underset{\mathbf{x}\in\mathbb{R}^n}{\operatorname{argmin}}\, J(\mathbf{x}),$$

can be found by differentiating J with respect to $\mathbf{x}$,

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A} - \mathbf{b}^\mathsf{T}\mathbf{A}$$

and setting the result to zero,

$$\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A} - \mathbf{b}^\mathsf{T}\mathbf{A} = \mathbf{0}.$$

Rearranging leads to

$$\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b},$$

which are called the *associated system of normal equations*, or simple the *normal equations*. The name normal equations comes from observing that, when the solution is written as $\mathbf{A}^\mathsf{T}(\mathbf{A}\mathbf{x} - \mathbf{b})$, $\mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0}$ is normal (i.e., orthogonal) to $\mathbf{A}^\mathsf{T}$. Said another way, $\mathbf{A}\mathbf{x} - \mathbf{b}$ lies in the null space of $\mathbf{A}^\mathsf{T}$. Notice that a solution to the normal equations always exists because, even if $\rho(\mathbf{A}) < n$, both $\mathbf{A}\mathbf{x}$ and $\mathbf{b}$ are multiplied by $\mathbf{A}^\mathsf{T}$, and as such, both the left-hand and right-hand sides of the normal equations lie in $\mathcal{R}(\mathbf{A}^\mathsf{T})$.

In the special case where $\rho(\mathbf{A}) = n \leq m$, the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ is of rank n and symmetric, and as such the least square solution is unique. As such, the inverse of $\mathbf{A}^\mathsf{T}\mathbf{A}$ exists, but actually computing this inverse may be impractical. In order to avoid computing the inverse of $\mathbf{A}^\mathsf{T}\mathbf{A}$ when finding $\mathbf{x}^\star$, the minimizing solution of the objective function, the QR decomposition can be employed. Consider a QR decomposition of $\mathbf{A}$, that is, $\mathbf{A} = \mathbf{Q}\mathbf{R}$. Then, the normal equations are

$$\begin{aligned}
\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{A}^\mathsf{T}\mathbf{b} \quad &\Leftrightarrow \quad (\mathbf{Q}\mathbf{R})^\mathsf{T}\mathbf{Q}\mathbf{R}\mathbf{x} = (\mathbf{Q}\mathbf{R})^\mathsf{T}\mathbf{b} \\
&\Leftrightarrow \quad \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{Q}\mathbf{R}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b} \\
&\Leftrightarrow \quad \mathbf{R}^\mathsf{T}\mathbf{R}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b},
\end{aligned}$$

where $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{1}$. The system $\mathbf{R}^\mathsf{T}\mathbf{R}\mathbf{x} = \mathbf{R}^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{b}$ can be solved by *forward substitution* followed by *backward substitution*. However, the matrix $\mathbf{R}$ is full rank and hence invertible, leading to

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^\mathsf{T}\mathbf{b},$$

which is easily solved via backward substitution.

A.1.9 Eigenvalues, Eigenvectors, and Diagonalization

Definition A.17. [10, pp. 533] Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$. The number λ, which could be real or complex, is an *eigenvalue of* $\mathbf{A}$ if there exists a nonzero vector $\mathbf{v} \in \mathbb{C}^n$ such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \tag{A.5}$$

The vector $\mathbf{v} \neq \mathbf{0}$ is the *eigenvector of* $\mathbf{A}$ corresponding to λ.

Rearrange (A.5) and write

$$(\mathbf{A} - \lambda\mathbf{1})\mathbf{v} = \mathbf{0}.$$

Given that $\mathbf{v} \neq \mathbf{0}$, the above equation can be interpreted as follows: $\mathbf{v}$ lies in the null space of the matrix $(\mathbf{A} - \lambda\mathbf{1})$, that is, $\mathbf{v} \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{1}) = E_\lambda$ where E_λ is the eigenspace of $\mathbf{A}$ corresponding to the eigenvalue λ. If $(\mathbf{A} - \lambda\mathbf{1})$ has a null space then $(\mathbf{A} - \lambda\mathbf{1})$ is singular. If $(\mathbf{A} - \lambda\mathbf{1})$ is singular then $\det(\mathbf{A} - \lambda\mathbf{1}) = 0$.

Theorem A.10. [10, pp. 535] Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$. Then λ is an eigenvalue of $\mathbf{A}$ iff $p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{1}) = 0$. $p(\lambda)$ is the *characteristic polynomial of* $\mathbf{A}$.

Because the characteristic polynomial $p(\lambda)$ has n roots there are n λ's (not necessarily distinct).

Theorem A.11. [10, pp. 536] Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$ and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the distinct eigenvalues of $\mathbf{A}$ (i.e., $\lambda_i \neq \lambda_j$ if $i \neq j$) and $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n$ be the corresponding eigenvectors. Then $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n$ are linearly independent.

Proof. The statement above will be proven using mathematical induction. Starting with $n = 2$, suppose that $c_1\mathbf{v}^1 + c_2\mathbf{v}^2 = \mathbf{0}$. Left multiply both sides by $\mathbf{A}$ to get

$$\begin{aligned} c_1\mathbf{A}\mathbf{v}^1 + c_2\mathbf{A}\mathbf{v}^2 &= \mathbf{0}, \\ c_1\lambda_1\mathbf{v}^1 + c_2\lambda_2\mathbf{v}^2 &= \mathbf{0}. \end{aligned}$$

Multiply $c_1\mathbf{v}^1 + c_2\mathbf{v}^2 = \mathbf{0}$ by λ_1 and subtract the result from the above to get

$$\begin{aligned} c_1\lambda_1\mathbf{v}^1 + c_2\lambda_2\mathbf{v}^2 - c_1\lambda_1\mathbf{v}^1 - c_2\lambda_1\mathbf{v}^2 &= \mathbf{0}, \\ c_2(\lambda_2 - \lambda_1)\mathbf{v}^2 &= \mathbf{0}. \end{aligned}$$

Because $\mathbf{v}^2 \neq \mathbf{0}$ and $\lambda_1 \neq \lambda_2$ it follows that $c_2 = 0$. Substitution of $c_2 = 0$ into $c_1\mathbf{v}^1 + c_2\mathbf{v}^2 = \mathbf{0}$ and recalling that $\mathbf{v}^1 \neq \mathbf{0}$ it follows that $c_1 = 0$. Thus, $\mathbf{v}^1$ and $\mathbf{v}^2$ are linearly independent.

A similar argument can be take for $n = 3, 4, \ldots, n$. □

Example A.18. Let

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}.$$

Find the eigenvalues and eigenvectors of $\mathbf{A}$.

Proof. Start with computing $\det(\mathbf{A} - \lambda\mathbf{1})$ and setting the result to zero gives

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{1}) &= \det\begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix} \\ &= (4-\lambda)(3-\lambda) - 6 \\ &= \lambda^2 - 7\lambda + 6 \\ &= (\lambda - 1)(\lambda - 6). \end{aligned}$$

Thus, the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = 6$. To solve for the eigenvector corresponding to λ_1, denoted $\mathbf{v}^1$, solve

$$\begin{aligned} (\mathbf{A} - \lambda\mathbf{1})\mathbf{v}^1 &= \mathbf{0}, \\ \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

From this it follows that $3v_1^1 + 2v_2^1 = 0$ and thus

$$\mathbf{v}^1 = \begin{bmatrix} -2/3 \\ 1 \end{bmatrix}.$$

Similarly, to find $\mathbf{v}^2$, solve

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{1})\mathbf{v}^2 &= \mathbf{0},\\ \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.\end{aligned}$$

From this it follows that $-2v_1^2 + 2v_2^2 = 0$, and thus

$$\mathbf{v}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

□

Theorem A.12. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$ where $\mathbf{A} = \mathbf{A}^\mathsf{T}$ (i.e., $\mathbf{A}$ is symmetric), λ_i are the n eigenvalues and $\mathbf{v}^i$ are the n eigenvectors. The eigenvalues are real, $\lambda_i \in \mathbb{R}$, as are the eigenvectors, $\mathbf{v}^i \in \mathbb{R}^n$.

Theorem A.13. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$ where $\mathbf{A} = \mathbf{A}^\mathsf{T}$ (i.e., $\mathbf{A}$ is symmetric) and assume that the n eigenvalues of $\mathbf{A}$ are distinct and the eigenvectors n have been normalized (i.e., are unit-length). Then the eigenvectors are orthogonal, that is

$${\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Proof. From Theorem A.12 both λ_i and $\mathbf{v}^i$ are real. Consider

$$\lambda_i\mathbf{v}^i = \mathbf{A}\mathbf{v}^i.$$

Transposing the above and right multiplying both sides of the equation by $\mathbf{v}^j$ where $i \neq j$ results in

$$\lambda_i{\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j = {\mathbf{v}^i}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{v}^j = {\mathbf{v}^i}^\mathsf{T}\mathbf{A}\mathbf{v}^j = \lambda_j{\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j,$$

that is,

$$(\lambda_i - \lambda_j){\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j = 0.$$

The assumption of $\lambda_i \neq \lambda_j$ results in the conclusion that ${\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j = 0$, that is, the eigenvectors satisfy

$${\mathbf{v}^i}^\mathsf{T}\mathbf{v}^j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

□

Theorem A.14. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$ and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of and $\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n$ be the corresponding eigenvectors. Assume that the n eigenvectors are linearly independent and let

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}^1 & \mathbf{v}^2 & \cdots & \mathbf{v}^n \end{bmatrix}, \quad \mathbf{\Lambda} = \text{diag}\left\{\lambda_1, \lambda_2, \cdots, \lambda_n\right\}.$$

Under the assumption that the eigenvectors are linearly independent the matrix $\mathbf{V}$ is nonsingular. The matrix $\mathbf{A}$ can be *diagonalized* or written in terms of its *eigendecomposition*,

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}.$$

Proof. This can be shown by noting that

$$\mathbf{A}\underbrace{\begin{bmatrix} \mathbf{v}^1 & \mathbf{v}^2 & \cdots & \mathbf{v}^n \end{bmatrix}}_{\mathbf{V}} = \begin{bmatrix} \mathbf{v}^1 & \mathbf{v}^2 & \cdots & \mathbf{v}^n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Right multiplying both sides by $\mathbf{V}^{-1}$ yields

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}.$$

Note that, from Theorem A.11, when the eigenvalues are distinct the eigenvectors are linearly independent and thus $\mathbf{V}$ is full rank and hence invertible. □

Theorem A.15. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$ where $\mathbf{A} = \mathbf{A}^\mathsf{T}$ (i.e., $\mathbf{A}$ is symmetric) and assume that the n eigenvalues of $\mathbf{A}$ are distinct and the eigenvectors n have been normalized (i.e., are unit-length). Then

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\mathsf{T}.$$

Proof. From Theorem A.13 it is known that the eigenvectors are orthogonal and therefore

$$\mathbf{V}^\mathsf{T}\mathbf{V} = \begin{bmatrix} {\mathbf{v}^1}^\mathsf{T} \\ {\mathbf{v}^2}^\mathsf{T} \\ \vdots \\ {\mathbf{v}^n}^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{v}^1 & \mathbf{v}^2 & \cdots & \mathbf{v}^n \end{bmatrix} = \begin{bmatrix} {\mathbf{v}^1}^\mathsf{T}\mathbf{v}^1 & {\mathbf{v}^1}^\mathsf{T}\mathbf{v}^2 & \cdots & {\mathbf{v}^1}^\mathsf{T}\mathbf{v}^n \\ {\mathbf{v}^2}^\mathsf{T}\mathbf{v}^1 & {\mathbf{v}^2}^\mathsf{T}\mathbf{v}^2 & \cdots & {\mathbf{v}^2}^\mathsf{T}\mathbf{v}^n \\ \vdots & \vdots & \ddots & \vdots \\ {\mathbf{v}^n}^\mathsf{T}\mathbf{v}^1 & {\mathbf{v}^n}^\mathsf{T}\mathbf{v}^2 & \cdots & {\mathbf{v}^n}^\mathsf{T}\mathbf{v}^n \end{bmatrix} = \mathbf{1}.$$

The matrix $\mathbf{V}$ is an *orthogonal matrix*. Consider right multiplying $\mathbf{V}^\mathsf{T}\mathbf{V} = \mathbf{1}$ by $\mathbf{V}^{-1}$,

$$\begin{aligned} \mathbf{V}^\mathsf{T}\mathbf{V}\mathbf{V}^{-1} &= \mathbf{V}^{-1}, \\ \mathbf{V}^\mathsf{T} &= \mathbf{V}^{-1}. \end{aligned}$$

Therefore,

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\mathsf{T}.$$

□

Theorem A.16. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$, and assume that the n eigenvalues of $\mathbf{A}$ are distinct. Then

$$\det \mathbf{A} = \prod_{i=1}^{n} \lambda_i.$$

Proof. Recall that for matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ that $\det(\mathbf{ABC}) = \det\mathbf{A}\det\mathbf{B}\det\mathbf{C}$. Also recall that $\det\mathbf{A}^{-1} = 1/\det\mathbf{A}$. Therefore, using the eigendecomposition of $\mathbf{A}$

$$\det\mathbf{A} = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}) = \det\mathbf{V}\det\mathbf{\Lambda}\det\mathbf{V}^{-1} = \det\mathbf{V}\det\mathbf{\Lambda}\frac{1}{\det\mathbf{V}} = \det\mathbf{\Lambda} = \prod_{i=1}^{n}\lambda_i.$$

Note that from Theorem A.11 that when the eigenvalues are distinct the eigenvectors are linearly independent and thus $\mathbf{V}$ is invertible □

Theorem A.17. Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$, and assume that the n eigenvalues of $\mathbf{A}$ are distinct. Then

$$\mathrm{tr}\mathbf{A} = \sum_{i=1}^{n} \lambda_i.$$

Proof. Recall that for matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ that $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB})$. Therefore, using the eigendecomposition of $\mathbf{A}$,

$$\mathrm{tr}\mathbf{A} = \mathrm{tr}(\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}) = \mathrm{tr}(\mathbf{V}^{-1}\mathbf{V}\boldsymbol{\Lambda}) = \mathrm{tr}\boldsymbol{\Lambda} = \sum_{i=1}^{n} \lambda_i.$$

□

It can be shown that $\mathrm{tr}\mathbf{A} = \sum_{i=1}^{n} \lambda_i$ without assuming the n eigenvalues of $\mathbf{A}$ are distinct. Doing so requires the notion of the Jordan form of $\mathbf{A}$.

A.1.10 The Singular Value Decomposition

Eigenvalues and eigenvectors are only relevant to square matrices. For rectangular matrices, a useful decomposition is the singular value decomposition.

Definition A.18. [9, pp. 412] Consider $\mathbf{A} \in \mathbb{R}^{m\times n}$ of rank r. There exists a diagonal matrix $\mathbf{D} \in \mathbb{R}^{r\times r}$ and orthonormal matrices $\mathbf{U} \in \mathbb{R}^{m\times m}$ and $\mathbf{V} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{A} = \mathbf{U}\begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^{\mathsf{T}}$$

where $\mathbf{D} = \mathrm{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_r\}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. The σ_i's are the singular values of $\mathbf{A}$, and the columns in $\mathbf{U}$ and $\mathbf{V}$ are the left-hand and right-had singular vectors of $\mathbf{A}$, respectively.

There is an elegant relationship between the singular values of $\mathbf{A}$ and the eigenvalues of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ or $\mathbf{A}\mathbf{A}^{\mathsf{T}}$.

Theorem A.18. Consider $\mathbf{A} \in \mathbb{R}^{m\times n}$ with singular value decomposition $\mathbf{A} = \mathbf{U}\mathrm{diag}(\mathbf{D}, \mathbf{0})\mathbf{V}^{\mathsf{T}}$ where $\mathbf{D} = \mathrm{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_r\}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $\mathbf{V} = [\mathbf{v}^1\ \mathbf{v}^2\ \cdots \mathbf{v}^n]$ and $\mathbf{U} = [\mathbf{v}^1\ \mathbf{v}^2\ \cdots \mathbf{v}^m]$. Then σ_i^2 and $\mathbf{v}^i$ are an eigenpair of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ meaning that the nonzero singular values of $\mathbf{A}$ are the positive square roots of the nonzero eigenvalues of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ and right-hand singular vectors $\mathbf{v}^i$ of $\mathbf{A}$ are (particular) eigenvectors of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$. Similarly, σ_i^2 and $\mathbf{u}^i$ are an eigenpair of $\mathbf{A}\mathbf{A}^{\mathsf{T}}$, meaning that the nonzero singular values of $\mathbf{A}$ are the positive square roots of the nonzero eigenvalues of $\mathbf{A}\mathbf{A}^{\mathsf{T}}$ and left-hand singular vectors $\mathbf{u}^i$ of $\mathbf{A}$ are (particular) eigenvectors of $\mathbf{A}\mathbf{A}^{\mathsf{T}}$.

Proof. By direct computation,

$$\mathbf{A}^{\mathsf{T}}\mathbf{A} = \mathbf{V}\begin{bmatrix} \mathbf{D}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^{\mathsf{T}}.$$

In a similar manner,

$$\mathbf{A}\mathbf{A}^{\mathsf{T}} = \mathbf{U}^{\mathsf{T}}\begin{bmatrix} \mathbf{D}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{U}.$$

□

A.1.11 Symmetric and Skewsymmetric Projection Operators

Consider $\mathbf{P} \in \mathbb{R}^{n\times n}$. Reader are assumed to be familiar with symmetric matrices, those that satisfy $\mathbf{P} = \mathbf{P}^{\mathsf{T}}$. A skew symmetric matrix is equal to the negative of its transpose, $\mathbf{P} = -\mathbf{P}^{\mathsf{T}}$. Readers not familiar with skewsymmetric matrices might find it helpful to know that any $n \times n$ matrix, such as $\mathbf{P} \in \mathbb{R}^{n \times n}$ can be decomposed into symmetric and skewsymmetric parts using the symmetric and skewsymmetric projection operators.

Definition A.19. Consider $\mathbf{P} \in \mathbb{R}^{n \times n}$ and assume that $\mathbf{P}$ is unstructured (i.e., it is neither symmetric nor skewsymmetric). $\mathbf{P} \in \mathbb{R}^{n \times n}$ can be decomposed into symmetric and skewsymmetric parts,

$$\mathbf{P} = \tfrac{1}{2}\mathbf{P} + \tfrac{1}{2}\mathbf{P} = \tfrac{1}{2}\mathbf{P} + \tfrac{1}{2}\mathbf{P} + \tfrac{1}{2}\mathbf{P}^{\mathsf{T}} - \tfrac{1}{2}\mathbf{P}^{\mathsf{T}} = \underbrace{\tfrac{1}{2}\left(\mathbf{P} + \mathbf{P}^{\mathsf{T}}\right)}_{\boldsymbol{\mathcal{P}}_s(\mathbf{P})} + \underbrace{\tfrac{1}{2}\left(\mathbf{P} - \mathbf{P}^{\mathsf{T}}\right)}_{\boldsymbol{\mathcal{P}}_a(\mathbf{P})},$$

where $\boldsymbol{\mathcal{P}}_s(\mathbf{P})$ and $\boldsymbol{\mathcal{P}}_a(\mathbf{P})$ are symmetric and skewsymmetric matrices, respectively. The *symmetric* and *skewsymmetric projection operators* are $\boldsymbol{\mathcal{P}}_s(\mathbf{P}) = \frac{1}{2}(\mathbf{P} + \mathbf{P}^{\mathsf{T}})$ and $\boldsymbol{\mathcal{P}}_a(\mathbf{P}) = \frac{1}{2}(\mathbf{P} - \mathbf{P}^{\mathsf{T}})$, respectively.

A useful identity involving the trace of a 3×3 matrix and the skewsymmetric projection operator is

$$\tfrac{1}{2}\mathrm{tr}\left(\mathbf{v}^{\times}\mathbf{P}\right) = -\mathbf{v}^{\mathsf{T}}\boldsymbol{\mathcal{P}}_a(\mathbf{P})^{\vee}, \quad \forall \mathbf{v} \in \mathbb{R}^3, \forall \mathbf{P} \in \mathbb{R}^{3 \times 3},$$

where

$$\mathbf{v}^{\times} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}, \quad \forall \mathbf{v} \in \mathbb{R}^3.$$

A.1.12 Positive (Semi)Definite, Negative (Semi)Definite, and Indefinite Matrices

Definition A.20. [11, pp. 42]. Consider the symmetric matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$. The matrix $\mathbf{P}$

1. is *positive definite* if
$$\mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x} > 0, \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \mathbf{x} = \mathbf{0},$$
2. is *positive semidefinite* if
$$\mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x} \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$
3. is *negative definite* if
$$\mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x} < 0, \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \mathbf{x} = \mathbf{0},$$
4. is *negative semidefinite* if
$$\mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x} \leq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$
5. and is indefinite if $\mathbf{x}^{\mathsf{T}}\mathbf{P}\mathbf{x}$ is neither positive nor negative definite.

Theorem A.19. [11, pp. 43]. Consider the symmetric matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ and let $\underline{\lambda}(\mathbf{P})$ and $\bar{\lambda}(\mathbf{P})$ denote the minimum and maximum eigenvalue of $\mathbf{P}$. The matrix $\mathbf{P}$

1. is *positive definite* iff
$$\underline{\lambda}(\mathbf{P}) > 0,$$
2. is *positive semidefinite* iff
$$\underline{\lambda}(\mathbf{P}) \geq 0,$$
3. is *negative definite* if
$$\bar{\lambda}(\mathbf{P}) < 0$$

4. is *negative semidefinite* if

$$\bar{\lambda}(\mathbf{P}) \leq 0,$$

5. and is neither positive nor negative definite iff $\underline{\lambda}(\mathbf{P}) < 0$ and $\bar{\lambda}(\mathbf{P}) > 0$.

Proof. To see why the positive or negative definiteness of $\mathbf{x}^\mathsf{T}\mathbf{P}\mathbf{x}$ is dictated by the eigenvalues of $\mathbf{P}$, let $\mathbf{P} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ and notice that

$$\begin{aligned}\mathbf{x}^\mathsf{T}\mathbf{P}\mathbf{x} &= \mathbf{x}^\mathsf{T}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{x} \\ &= (\mathbf{V}\mathbf{x})^\mathsf{T}\mathbf{\Lambda}\mathbf{V}^\mathsf{T}\mathbf{x} \\ &= \mathbf{z}^\mathsf{T}\mathbf{\Lambda}\mathbf{z} \\ &= \sum_{i=1}^{n} \lambda_i z_i^2\end{aligned}$$

where $\mathbf{z} = \mathbf{V}^\mathsf{T}\mathbf{x}$ and, because $\mathbf{P}$ is symmetric $\mathbf{V}^{-1} = \mathbf{V}^\mathsf{T}$. □

A.1.13 The Schur Complement and The Matrix Inversion Lemma

This section is based on [12–14]

Theorem A.20. [12] Consider the matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$. Partition $\mathbf{X}$ in block matrix form

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

where $\mathbf{A} \in \mathbb{R}^{p\times p}$, $\mathbf{B} \in \mathbb{R}^{p\times q}$, $\mathbf{C} \in \mathbb{R}^{q\times p}$, and $\mathbf{D} \in \mathbb{R}^{q\times q}$. If both $\mathbf{D}$ and $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ are invertible then

$$\mathbf{X}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix} \tag{A.6}$$

The matrix $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ is the *Schur Complement of* $\mathbf{D}$ *in* $\mathbf{X}$. If both $\mathbf{A}$ and $\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ are invertible then

$$\mathbf{X}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} \tag{A.7}$$

The matrix $\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ is the *Schur Complement of* $\mathbf{A}$ *in* $\mathbf{X}$.

Proof. Equality of the left-hand and right-hand sides of Equation (A.6) will be shown first by showing that $\mathbf{X}^{-1}\mathbf{X} = \mathbf{1}$ when the expression for $\mathbf{X}^{-1}$ given in Equation (A.6) is used. Begin with factoring the right-hand-side of Equation (A.6) as

$$\begin{aligned}\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{1} \end{bmatrix} \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}.\end{aligned}$$

It follows that

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{1} \end{bmatrix}.$$

By direct computation it is straightforward to verify that $\mathbf{X}^{-1}\mathbf{X} = \mathbf{1}$. To this end,

$$\begin{aligned}
&\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{1} \end{bmatrix} \begin{bmatrix} (\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{1} \end{bmatrix} \\
&= \mathbf{1}.
\end{aligned}$$

Equality of Equation (A.7) can be proven in a similar way. Being with factoring the right-hand-side of Equation (A.7) as

$$\begin{aligned}
\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{1} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ -(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{1} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{1} \end{bmatrix}.
\end{aligned}$$

It follows that

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}.$$

By direct computation it is straightforward to verify that $\mathbf{X}^{-1}\mathbf{X} = \mathbf{1}$. It follows that

$$\begin{aligned}
&\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{1} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \\
&= \mathbf{1}.
\end{aligned}$$

$\square$

In the special case where $\mathbf{A}$, $\mathbf{D}$, the Schur complement of $\mathbf{A}$ in $\mathbf{X}$, $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}$, and the Schur complement of $\mathbf{D}$ in $\mathbf{X}$, $\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}$, are all invertible, equating the $(1,1)$ block of each expression for $\mathbf{X}^{-1}$ in Equations (A.6) and (A.7) leads to [12]

$$(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1},$$

which is called the Woodbury matrix identity. When $\mathbf{D} = \mathbf{1}$ and $\mathbf{B}$ is changed to $-\mathbf{B}$ results in the *Matrix Inversion Lemma*, [12]

$$(\mathbf{A}+\mathbf{B}\mathbf{C})^{-1} = \mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{1}+\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}.$$

A.2 Differential Equations

As mentioned Chapter 1, a dynamic analysis (or, more precisely, a kinematic followed by a dynamic analysis) is really about deriving the differential equation describing of motion of a body. As such, linear and nonlinear differential equations will be briefly discussed.

A.2.1 Linear Differential Equations

Consider the following *differential equation*,

$$m\ddot{q}(t) + d\dot{q}(t) + kq(t) = 0, \quad q_0 = q(0),\ \dot{q}_0 = \dot{q}(0), \tag{A.8}$$

where $0 < m < \infty$, $0 < d < \infty$, $0 < k < \infty$, $q_0 = q(0)$ and $\dot{q}_0 = \dot{q}(0)$ are initial conditions (ICs), and $q(\cdot)$ is the dependent variable and t is the independent variable. The differential equation in Equation (A.8) is an ordinary differential equation (ODE) because there is only one independent variable [5, pp. 645-646]. In this text ODEs will be primarily dealt with, and unless clarity is required the "O" in ODE will be dropped, and DE will be written. Moreover, the DE in Equation (A.8) is a second-order DE. The order of a DE is equal to the highest derivative in the differential equation [5, pp. 647].

Before discussing solutions Equation (A.8), writing Equation (A.8) in a first-order form will be discussed. Equation (A.8) can be equivalently written as

$$\begin{bmatrix} \dot{q}(t) \\ \ddot{q}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{d}{m} \end{bmatrix} \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix}, \quad \begin{bmatrix} q_0 \\ \dot{q}_0 \end{bmatrix} = \begin{bmatrix} q(0) \\ \dot{q}(0) \end{bmatrix}. \tag{A.9}$$

By defining

$$\mathbf{x}(t) = \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{d}{m} \end{bmatrix},$$

where $x_1(t) = q(t)$ and $x_2(t) = \dot{q}(t) = \dot{x}_1(t)$ are referred to as the system states, Equation (A.9) can be written

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad \mathbf{x}_0 = \mathbf{x}(0). \tag{A.10}$$

Equation (A.10) is the first-order form of Equation (A.8), also called the first-order-matrix form or first-order state-space form, and it is not unique. When it comes time to find an analytic or numerical solution to Equation (A.10) this first order form is often preferred over the second order scalar form of the DE in Equation (A.8).

A.2.1.1 The Solution of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ and the Matrix Exponential

Next, consider the solution to Equation (A.10). There are various ways to find solutions of Equation (A.10).

Before discussing the solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ given the ICs $\mathbf{x}_0 = \mathbf{x}(0)$, the solution to the first-order DE

$$\dot{x}(t) = ax(t), \quad x_0 = x(0)$$

will be discussed. The solution this first order scalar DE is

$$x(t) = e^{at}x_0, \tag{A.11}$$

where [6, pp. 53]

$$e^{at} = \sum_{k=0}^{\infty} \frac{a^k t^k}{k!}. \tag{A.12}$$

Motivated by the previous scalar example, it seems natural that the solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ is of the form

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0, \quad \mathbf{x}_0 = \mathbf{x}(0) \tag{A.13}$$

where $e^{\mathbf{A}t}$ is the *matrix exponential*. It can be rigorously shown that Equation (A.13) is indeed the (unique) solution to Equation (A.10) [15, pp. 396-397] [16, pp. 78-99]. The matrix exponential is defined as [6, pp. 53] [16, pp. 85]

$$e^{\mathbf{A}t} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!},$$

and has several interesting properties [6, pp. 54]. For any $\mathbf{A} \in \mathbb{R}^{n\times n}$

1. $e^{\mathbf{A}t}$ is the unique matrix satisfying $\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}$, $e^{\mathbf{A}t}|_{t=0} = \mathbf{1}$,
2. $e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{A}t_1}e^{\mathbf{A}t_2}, \forall t_1, t_2 \in \mathbb{R}$, which means that $e^{\mathbf{A}t}e^{-\mathbf{A}t} = e^{\mathbf{A}(t-t)} = e^{\mathbf{A}(0)} = \mathbf{1}$,
3. $e^{\mathbf{A}t}$ is nonsingular with inverse $(e^{\mathbf{A}t})^{-1} = e^{-\mathbf{A}t}$,
4. $\mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A}$ (that is, $\mathbf{A}$ and $e^{\mathbf{A}t}$ commute),
5. $e^{(\mathbf{A}+\mathbf{B})t} = e^{\mathbf{A}t}e^{\mathbf{B}t}$ $\forall t \in \mathbb{R}$ iff $\mathbf{AB} = \mathbf{BA}$ (i.e., $\mathbf{A}$ and $\mathbf{B}$ commute), and
6. $e^{\mathbf{PAP}^{-1}t} = \mathbf{P}e^{\mathbf{A}t}\mathbf{P}^{-1}$ for all nonsingular $\mathbf{P}^{n\times n}$.

Now that the solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$ is known, that being $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$ where $\mathbf{x}_0 = \mathbf{x}(0)$, how to compute a solution can be discussed. Assume the n eigenvalues of $\mathbf{A}$ are distinct and thus, from Theorem A.11, the n eigenvectors are linearly independent. Under the assumption that the n eigenvectors are linearly independent and thus span $\mathbb{R}^n$, $\mathbf{x}_0$ can be written as a linear combination of the eigenvectors of $\mathbf{A}$,

$$\mathbf{x}_0 = \sum_{i=1}^{n} c_i \mathbf{v}^i.$$

Using the matrix exponential it follows that

$$\begin{aligned}
\mathbf{x}(t) &= e^{\mathbf{A}t}\mathbf{x}_0 \\
&= \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{A}^k\mathbf{x}_0 \\
&= \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{A}^k \left[\sum_{i=1}^{n} c_i\mathbf{v}^i\right] \\
&= \sum_{k=0}^{\infty}\sum_{i=1}^{n} c_i\frac{t^k}{k!}\mathbf{A}^k\mathbf{v}^i \\
&= \sum_{i=1}^{n} c_i \sum_{k=0}^{\infty} \frac{t^k}{k!}\lambda_i^k\mathbf{v}^i \\
&= \sum_{i=1}^{n} c_i e^{\lambda_i t}\mathbf{v}^i,
\end{aligned}$$

where the expression for $e^{\lambda_i t}$ comes from Equation (A.12). Therefore, when the n eigenvalues of $\mathbf{A}$ are distinct the solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ is a linear combination of the eigenvectors of $\mathbf{A}$. This solution method is known as *modal decomposition*.

The solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ when the eigenvalues are not distinct is discussed in [15, pp. 401-410] [16, pp. 88-90]. It is worth mentioning what solutions to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ look like when eigenvalues are repeated. Consider the case where $\mathbf{A} \in \mathbb{R}^{2\times 2}$, $\lambda_1 = \lambda_2 = \lambda$; the solution is of the form $\mathbf{x}(t) = c_1 e^{\lambda t}\mathbf{v}^1 + c_2 t e^{\lambda t}\mathbf{v}^2$. Notice the $te^{\lambda t}$ term.

Example A.19. [15, pp. 395] Find the solution of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}.$$

Proof. First, find the eigenvalues of $\mathbf{A}$. To this end,

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{1}) &= (1-\lambda)(1-\lambda) - 4 \\ &= (1 - 2\lambda + \lambda^2) - 4 \\ &= \lambda^2 - 2\lambda - 3 \\ &= (\lambda + 1)(\lambda - 3). \end{aligned}$$

Therefore, $\lambda_1 = -1$ and $\lambda_2 = 3$.

Next, find the eigenvectors of $\mathbf{A}$, starting with $\mathbf{v}^1$. To this end,

$$\begin{aligned} (\mathbf{A} - \lambda_1 \mathbf{1})\mathbf{v}^1 &= \mathbf{0}, \\ \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}\begin{bmatrix} v_1^1 \\ v_2^1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ -2v_1^1 + v_2^1 &= 0, \\ 4v_1^1 - 2v_2^1 &= 0, \\ \mathbf{v}^1 &= \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \end{aligned}$$

Similarly,

$$\begin{aligned} (\mathbf{A} - \lambda_2 \mathbf{1})\mathbf{v}^2 &= \mathbf{0}, \\ \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} v_1^2 \\ v_2^2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ 2v_1^2 + v_2^2 &= 0, \\ 4v_1^2 + 2v_2^2 &= 0, \\ \mathbf{v}^1 &= \begin{bmatrix} 1 \\ -2 \end{bmatrix}. \end{aligned}$$

Therefore, the solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t}$$

where c_1 and c_2 depend on $\mathbf{x}_0$.

□

A.2.1.2 Linearity

Next it will be shown that $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$, $\mathbf{x}_0 = \mathbf{x}(0)$ given in Equation (A.13), which is the solution to $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ given in Equation (A.10), is *linear* [16, pp. 49-50]. For $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$, $\mathbf{x}_0 = \mathbf{x}(0)$ to be linear it must satisfy two properties: *superposition* and *scaling*. Showing superposition and scaling hold will be done simultaneously. Let $\mathbf{x}^1(\cdot)$ be the solution to Equation (A.10) given the IC $\mathbf{x}_0^1$ and $\mathbf{x}^2(\cdot)$ be the solution to Equation (A.10) given the IC $\mathbf{x}_0^2$. Also, let $\mathbf{x}_0 = \alpha_1\mathbf{x}_0^1 + \alpha_2\mathbf{x}_0^2$ for any $\alpha_1, \alpha_2 \in \mathbb{R}$. Then,

$$\begin{aligned}
\mathbf{x}(t) &= e^{\mathbf{A}t}\mathbf{x}_0 \\
&= e^{\mathbf{A}t}(\alpha_1\mathbf{x}_0^1 + \alpha_2\mathbf{x}_0^2) \\
&= \alpha_1 e^{\mathbf{A}t}\mathbf{x}_0^1 + \alpha_2 e^{\mathbf{A}t}\mathbf{x}_0^2 \\
&= \alpha_1\mathbf{x}^1(t) + \alpha_2\mathbf{x}^2(t).
\end{aligned}$$

Thus, $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$, $\mathbf{x}_0 = \mathbf{x}(0)$ is linear.

In reality all systems are, to some degree, nonlinear. So, why then are linear systems important? Because "we know how to solve them[1]". Or, to quote Professor Pierre T. Kabamba,

The world is nonlinear, this is true. But when signals are small over all, $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ *will do!*

A.2.2 Nonlinear Differential Equations

Consider the following differential equation,

$$m\ddot{q}(t) + d\dot{q}(t) + \left(k_1 + k_2 q^2(t)\right) q(t) = 0, \quad q_0 = q(0),\ \dot{q}_0 = \dot{q}(0), \tag{A.14}$$

where $0 < m < \infty$, $0 < d < \infty$, $0 < k_1 < \infty$, $k_2 \in \mathbb{R}$, $q_0 = q(0)$ and $\dot{q}_0 = \dot{q}(0)$ are initial conditions (ICs). The DE in Equation (A.14) is a special case of a Duffing oscillator. Equation (A.14) is a nonlinear DE owing to the presence of the $k_2 q^3(\cdot)$ term.

Like the linear DE in Equation (A.8), the nonlinear DE in Equation (A.14) can be cast into first-order form,

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)), \quad \mathbf{x}_0 = \mathbf{x}(0), \tag{A.15}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix}, \quad \mathbf{f}(\mathbf{x}(t)) = \begin{bmatrix} x_2(t) \\ -\frac{k_1}{m}x_1(t) - \frac{k_2}{m}x_1^3(t) - \frac{d}{m}x_2(t) \end{bmatrix}, \quad \mathbf{x}_0 = \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}, \tag{A.16}$$

and $x_1(t) = q(t)$ and $x_2(t) = \dot{q}(t) = \dot{x}_1(t)$ are the system states.

Note that the nonlinear function $\mathbf{f}(\cdot)$ in Equation (A.15) is not an explicit function of time. As such, the DE in Equation (A.15) is called an *autonomous* system or a *time-invariant* system. DEs that do explicitly depend on time, such as

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), \quad \mathbf{x}_0 = \mathbf{x}(0),$$

are called *nonautonmous* systems or *time-varying* systems.

A.2.2.1 Linearization

Before discussing linearization, both a nominal state trajectory and an equilibrium point must be defined [6, pp. 18] .

[1]This quote is attributed to Mr. Raja Mukherji, previously the manager of the control systems group at MacDonald Detweiler and Associates, Brampton, ON, Canada. Mr. Mukherji unfortunately died of stomach cancer in July 2018.

Definition A.21. The *nominal state trajectory* $\bar{\mathbf{x}}(\cdot)$ satisfies

$$\dot{\bar{\mathbf{x}}}(t) = \mathbf{f}(\bar{\mathbf{x}}(t), t), \quad \forall t \in \mathbb{R}^+.$$

An *equilibrium point*, or *equilibrium state*, is a constant state $\bar{\mathbf{x}}$ that satisfies

$$\mathbf{0} = \mathbf{f}(\bar{\mathbf{x}}, t), \quad \forall t \in \mathbb{R}^+.$$

Often we're interested in the deviation of the system state about a particular state trajectory or equilibrium point [6, pp. 18–20] [16, pp. 21-24]. Consider a small perturbation $\delta\mathbf{x}(\cdot)$ to the nominal state trajectory,

$$\mathbf{x}(t) = \bar{\mathbf{x}}(t) + \delta\mathbf{x}(t).$$

Substitution of $\mathbf{x}(t) = \bar{\mathbf{x}}(t) + \delta\mathbf{x}(t)$ into $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$ given in Equation (A.15) and expanding using a Taylor series expansion results in

$$\begin{aligned} \dot{\bar{\mathbf{x}}}(t) + \delta\dot{\mathbf{x}}(t) &= \mathbf{f}(\bar{\mathbf{x}}(t) + \delta\mathbf{x}(t), t) \\ &= \mathbf{f}(\bar{\mathbf{x}}(t), t) + \left.\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}}\right|_{\bar{\mathbf{x}}(t)} \delta\mathbf{x}(t) + \text{higher-order terms}, \end{aligned}$$

where

$$\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}(t),t)}{\partial x_1} & \frac{\partial f_1(\mathbf{x}(t),t)}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x}(t),t)}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x}(t),t)}{\partial x_1} & \frac{\partial f_2(\mathbf{x}(t),t)}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x}(t),t)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x}(t),t)}{\partial x_1} & \frac{\partial f_n(\mathbf{x}(t),t)}{\partial x_2} & \cdots & \frac{\partial f_n(\mathbf{x}(t),t)}{\partial x_n} \end{bmatrix}$$

is the Jacobain. Subtracting the nominal solution $\dot{\bar{\mathbf{x}}}(t) = \mathbf{f}(\bar{\mathbf{x}}(t), t)$ and, under the assumption of small $\delta\mathbf{x}(\cdot)$, neglecting the higher-order terms yields

$$\delta\dot{\mathbf{x}}(t) = \underbrace{\left.\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}}\right|_{\bar{\mathbf{x}}(t)}}_{\mathbf{A}(t)} \delta\mathbf{x}(t) = \mathbf{A}(t)\delta\mathbf{x}(t),$$

the linearized system. $\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t)$ is the linearized system about the nominal state trajectory $\bar{\mathbf{x}}(\cdot)$. If the system is autonomous and linearization is performed about an equilibrium point rather than a nominal state trajectory then

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t),$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is constant [16, pp. 22]. $\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t)$ is the linearized system about the equilibrium point $\bar{\mathbf{x}}$.

Example A.20. Find the equilibrium point(s) of Equation (A.14) and linearize it about those points to find the linearized DE.

Proof. Using the state-space form in Equation (A.16) the linearization points can be found, that being

$$\begin{aligned} \mathbf{0} &= \mathbf{f}(\bar{\mathbf{x}}) \\ &= \begin{bmatrix} \bar{x}_2 \\ -\frac{k_1}{m}\bar{x}_1 - \frac{k_2}{m}\bar{x}_1^3 - \frac{d}{m}\bar{x}_2 \end{bmatrix}, \end{aligned}$$

which implies

$$\begin{aligned} \bar{x}_2 &= 0, \\ \left(\frac{k_1}{m} + \frac{k_2}{m}\bar{x}_1^2\right)\bar{x}_1 + \frac{d}{m}\bar{x}_2 &= 0. \end{aligned}$$

The first of these two equations yields $\bar{x}_2 = 0$ while the second results in

$$\bar{x}_1 = 0 \quad \text{or} \quad \bar{x}_1 = \pm j\sqrt{\frac{k_1}{k_2}}.$$

The state $x_1(\cdot)$ is the position and cannot take on an imaginary value, thus $\bar{x}_1 = 0$. Therefore, the only valid equilibrium point is

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Next, linearize the system by expanding $\mathbf{x}(\cdot)$ and $\mathbf{f}(\cdot)$ using a Taylor series expansion,

$$\begin{aligned} \dot{\bar{\mathbf{x}}} + \delta\dot{\mathbf{x}}(t) &= \mathbf{f}(\bar{\mathbf{x}} + \delta\mathbf{x}(t)) \\ &= \begin{bmatrix} \bar{x}_2 + \delta\bar{x}_2 \\ -\frac{k_1}{m}(\bar{x}_1 + \delta x_1) - \frac{k_2}{m}(\bar{x}_1 + \delta x_1)^3 - \frac{d}{m}(\bar{x}_2 + \delta x_2) \end{bmatrix} \\ &= \begin{bmatrix} \bar{x}_2 \\ -\frac{k_1}{m}\bar{x}_1 - \frac{k_2}{m}\bar{x}_1^3 - \frac{d}{m}\bar{x}_2 \end{bmatrix} + \begin{bmatrix} \delta\bar{x}_2 \\ -\frac{k_1}{m}\delta x_1 - \frac{k_2}{m}(3\bar{x}_1^2\delta x_1 + 3\delta x_1^2\bar{x}_1 + \delta x_1^3) - \frac{d}{m}\delta x_2 \end{bmatrix}. \end{aligned}$$

Subtracting the nominal solution $\mathbf{0} = \mathbf{f}(\bar{\mathbf{x}})$, neglecting the higher-order terms, and substitution of $\bar{x}_1 = 0$ and $\bar{x}_2 = 0$ yields

$$\begin{bmatrix} \delta\dot{x}_1 \\ \delta\dot{x}_2 \end{bmatrix} = \begin{bmatrix} \delta\bar{x}_2 \\ -\frac{k_1}{m}\delta x_1 - \frac{d}{m}\delta x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m} & -\frac{d}{m} \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix},$$

the linearized DE.

□

A.3 Stability of Linear and Nonlinear Differential Equations

A.3.1 What is Stability?

Consider the autonomous system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)), \quad \mathbf{x}_0 = \mathbf{x}(0) \tag{A.17}$$

and the equilibrium points $\bar{\mathbf{x}}$ which are constant $n \times 1$ column matrices that satisfy $\mathbf{0} = \mathbf{f}(\bar{\mathbf{x}})$. A nonlinear system, such as the one in Equation (A.17), can have multiple isolated equilibrium points. The system response about each of these isolated equilibrium points may be quite different. Therefore, when assessing the stability of a system, what equilibrium point stability is being assessed about must be clearly identified.

Let $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\mathsf{T}\mathbf{x}}$ denote the Euclidian norm. Additionally, when $\bar{\mathbf{x}} \neq \mathbf{0}$, via a change of variables, the equilibrium point can be shifted to the origin, that is, $\bar{\mathbf{x}} = \mathbf{0}$. As such, in the following definitions the equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ will be considered, but it is understood that the stability of any equilibrium point can be assessed via a change of variables.

Definition A.22. [6, pp. 200] The equilibrium point $\bar{\mathbf{x}}$ of Equation (A.17) is:

1. *stable* if given any $0 < \epsilon < \infty$ there corresponds a $0 < \delta < \infty$ such that $\|\mathbf{x}_0\|_2 < \delta$ implies $\|\mathbf{x}(t)\|_2 < \epsilon, \forall t \in \mathbb{R}^+$;
2. *unstable* if it is not stable;
3. *asymptotically stable* if it is stable and there exists a $0 < \delta < \infty$ such that $\|\mathbf{x}_0\|_2 < \delta$ implies $\lim_{t\to\infty} \|\mathbf{x}(t)\|_2 \to 0$;
4. *globally asymptotically stable* if it is stable and $\lim_{t\to\infty} \|\mathbf{x}(t)\|_2 \to 0, \forall \mathbf{x}_0 \in \mathbb{R}^n$;
5. *exponentially stable* if there exists $0 < \delta < \infty$, $0 < k < \infty$, and $0 < \lambda < \infty$ such that $\|\mathbf{x}_0\|_2 < \delta$ implies $\|\mathbf{x}(t)\|_2 \leq ke^{-\lambda t} \|\mathbf{x}_0\|_2, \forall t \in \mathbb{R}^+$;
6. *globally exponentially stable* if there exists $0 < k < \infty$ and $0 < \lambda < \infty$ such that $\|\mathbf{x}(t)\|_2 \leq ke^{-\lambda t} \|\mathbf{x}_0\|_2, \forall t \in \mathbb{R}^+, \forall \mathbf{x}_0 \in \mathbb{R}^n$.

A.3.2 Stability of Linear Systems

When the DE in Equation (A.17) is linear, that is

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad \mathbf{x}_0 = \mathbf{x}(0), \tag{A.18}$$

stability can be assessed by looking at the eigenvalues of $\mathbf{A}$.

Theorem A.21. [17] The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of the Equation (A.18) is:

1. stable if Re $\{\lambda_i\} \leq 0$, $i = 1, 2, \ldots, n$ and there are no repeated eigenvalues on the imaginary axis;
2. unstable if there is at least one eigenvalue with Re $\{\lambda_i\} > 0$;
3. (globally) asymptotically stable if Re $\{\lambda_i\} < 0$, $i = 1, 2, \ldots, n$.

Figure A.1 shows the location of eigenvalues that correspond to stability, instability, and asymptotic stability.

A.3.3 Lyapunov's Indirect and Direct Methods

Stability of nonlinear systems will now be discussed.

A.3.3.1 Lyapunov's Indirect Method

Theorem A.22. [17] Consider the linearization of Equation (A.17), $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))$, $\mathbf{x}_0 = \mathbf{x}(0)$ about $\bar{\mathbf{x}} \in \mathbb{R}^n$ (a constant)

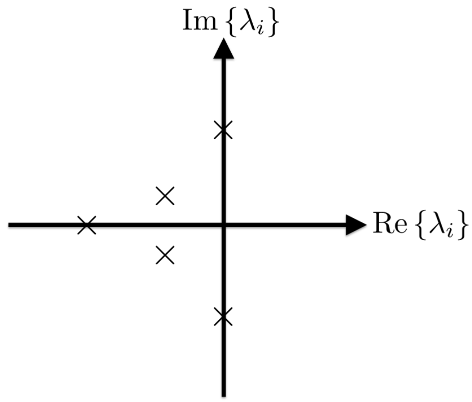
$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}\delta\mathbf{x}(t), \tag{A.19}$$

where

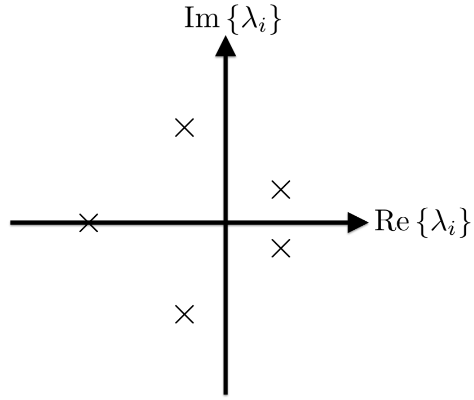
$$\mathbf{A} = \left.\frac{\partial \mathbf{f}(\mathbf{x}(t))}{\partial \mathbf{x}}\right|_{\bar{\mathbf{x}}}.$$

1. If Equation (A.19) is asymptotically stable then $\bar{\mathbf{x}}$ is an asymptotically stable equilibrium point of the nonlinear system.
2. If Equation (A.19) is unstable then $\bar{\mathbf{x}}$ is an unstable equilibrium point of the nonlinear system.
3. If Equation (A.19) is stable then no conclusion can be drawn about the nonlinear system.

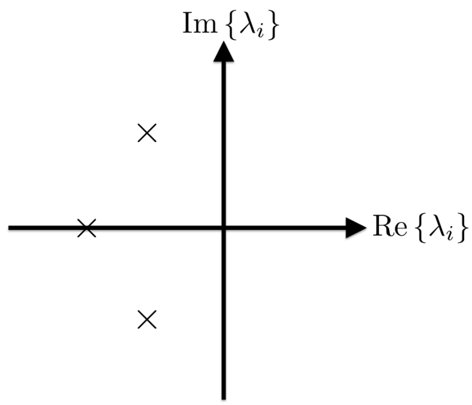
Note that one cannot conclude global asymptotic stability using Lyapunov's indirect method.



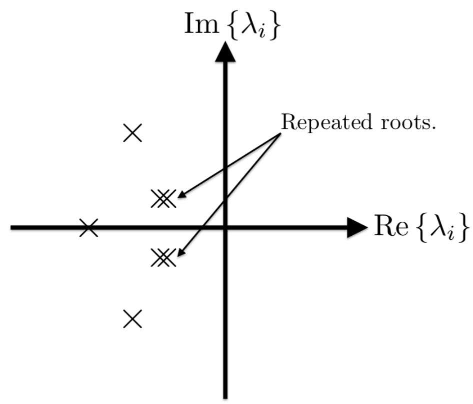
(a) Stable.



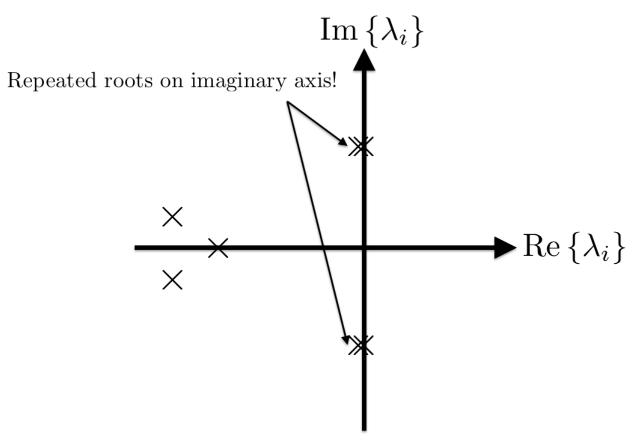
(b) Unstable.



(c) Asymptotically stable.



(d) Repeated eigenvlaues, but still asymptotically stable.



(e) Repeated eigenvlaues on the imaginary axis leading to instability.

Figure A.1: Location of eigenvalues and stability.

A.3.3.2 Lyapunov's Direct Method

Definition A.23. [17] Let *the closed ball of radius r* be defined $\bar{B}_r = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq r\}$.

Definition A.24. [17] The function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ is *locally positive definite* if $V(\mathbf{0}, t) = 0$ and there exists r such that $V(\mathbf{x}(t), t) > 0$, $\forall \mathbf{x}(t) \in \bar{B}_r \setminus \mathbf{x}(t) \neq \mathbf{0}$, $\forall t \in \mathbb{R}^+$. The function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ is *locally negative definite* if $-V(\cdot, \cdot)$ is locally positive definite.

Definition A.25. [17] The function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ is *positive definite* $V(\mathbf{0}, t) = 0$ and $V(\mathbf{x}(t), t) > 0$, $\forall \mathbf{x}(t) \in \mathbb{R}^n \setminus \mathbf{x}(t) \neq \mathbf{0}$, $\forall t \in \mathbb{R}^+$. The function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ is *negative definite* if $-V(\cdot, \cdot)$ is positive definite.

Definition A.26. [17] The function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ is positive definite and *radially unbounded* if it is positive definite and $V(\mathbf{x}(t), t) \to \infty$ as $\|\mathbf{x}(t)\|_2 \to \infty$, $\forall t \in \mathbb{R}^+$.

Theorem A.23. [17] Consider the nonautonomous system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t), \quad \mathbf{x}_0 = \mathbf{x}(0) \tag{A.20}$$

and the equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$.

1. The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of Equation (A.20) is *stable* if there exists a continuously differentiable locally positive definite function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ such that

   $$\dot{V}(\mathbf{x}(t), t) \leq 0, \quad \forall \mathbf{x}(t) \in \bar{B}_r, \forall t \in \mathbb{R}^+$$

   where $\dot{V}(\cdot, \cdot)$ is evaluated along the trajectories of Equation (A.20), that is

   $$\dot{V}(\mathbf{x}(t), t) = \frac{\partial V(\mathbf{x}(t), t)}{\partial t} + \frac{\partial V(\mathbf{x}(t), t)}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}(t), t).$$

2. The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of Equation (A.20) is *asymptotically stable* if there exists a continuously differentiable locally positive definite function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ such that

   $$\dot{V}(\mathbf{x}(t), t) < 0, \quad \forall \mathbf{x}(t) \in \bar{B}_r, \forall t \in \mathbb{R}^+.$$

3. The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of Equation (A.20) is *globally asymptotically stable* if there exists a continuously differentiable positive definite function $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$ such that $V(\cdot, \cdot)$ is radially unbounded and

   $$\dot{V}(\mathbf{x}(t), t) < 0, \quad \forall \mathbf{x}(t) \in \mathbb{R}^n \setminus \mathbf{x}(t) \neq \mathbf{0}, \forall t \in \mathbb{R}^+.$$

The function $V(\cdot, \cdot)$ is called a Lyapunov function. The statements in Theorem A.23 are sufficient conditions for stability, asymptotic stability, and global asymptotic stability, but not necessary conditions.

Example A.21. Consider the system described by Equation (A.14). With respect to the equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$, is the system stable? Asymptotically Stable? Globally asymptotically stable?

Proof. As shown in Example (A.20), $\bar{\mathbf{x}} = \mathbf{0}$ is indeed an equilibrium point. Although stability could be assessed using Lyapunov's indirect method, Lyapunov's direct method will be used to assess stability. Consider the Lyapunov-function candidate

$$V(\mathbf{x}(t)) = \tfrac{1}{2} m x_2^2(t) + \tfrac{1}{2} k_1 x_1^2(t).$$

Notice that $V(\mathbf{0}) = 0$, $V(\mathbf{x}(t)) > 0$, $\forall \mathbf{x}(t) \in \mathbb{R}^n$, $\forall t \in \mathbb{R}^+$, and $V(\mathbf{x}(t), t) \to \infty$ as $\|\mathbf{x}(t)\|_2 \to \infty$, $\forall t \in \mathbb{R}^+$.

The time-rate-of-chafe of the Lyapunov-function candidate is

$$\begin{aligned}
\dot{V}(\mathbf{x}(t)) &= m\dot{x}_2(t)x_2(t) + k_1x_1(t)\dot{x}_1(t) \\
&= m\dot{x}_2(t)x_2(t) + k_1x_1(t)x_2(t) \\
&= \Big(m\dot{x}_2(t) + k_1x_1(t)\Big)\ x_2(t) \\
&= \Big(-dx_2(t) - k_2x_1^3(t)\Big)\ x_2(t) \\
&= -dx_2^2(t) - k_2x_1^3(t)x_2(t)
\end{aligned}$$

where $\dot{x}_1(t) = x_2(t)$ and the motion equation in Equation (A.14) has been used to simplify. Notice that although $-dx_2^2(t) <$ the term $-k_2x_1^3(t)x_2(t)$ is not negative definite. As such, one cannot conclude if the system is stable, asymptotically stable, nor globally asymptotically stable using this particular Lyapunov-function candidate.

Next, try a different Lyapunov-function candidate. In particular, consider the Lyapunov-function candidate

$$V(\mathbf{x}(t)) = \tfrac{1}{2}mx_2^2(t) + \tfrac{1}{2}k_1x_1^2(t) + \tfrac{1}{4}k_2x_1^4(t)$$

which possesses the following properties: $V(\mathbf{0}) = 0$, $V(\mathbf{x}(t)) > 0$, $\forall \mathbf{x}(t) \in \mathbb{R}^n$, $\forall t \in \mathbb{R}^+$, and $V(\mathbf{x}(t), t) \to \infty$ as $\|\mathbf{x}(t)\|_2 \to \infty$, $\forall t \in \mathbb{R}^+$. The time-rate-of-change of the Lyapunov-function candidate is

$$\begin{aligned}
\dot{V}(\mathbf{x}(t)) &= m\dot{x}_2(t)x_2(t) + k_1x_1(t)\dot{x}_1(t) + k_2x_1^3(t)\dot{x}_1(t) \\
&= m\dot{x}_2(t)x_2(t) + k_1x_1(t)x_2(t) + k_2x_1^3(t)x_2(t) \\
&= \Big(m\dot{x}_2(t) + k_1x_1(t) + k_2x_1^3(t)\Big)\ x_2(t) \\
&= \Big(-dx_2(t)\Big)\ x_2(t) \\
&= -dx_2^2(t)
\end{aligned}$$

where the motion equation in Equation (A.14) has been used to simplify. In this case $\dot{V}(\mathbf{x}(t)) \leq 0$, $\forall \mathbf{x}(t) \in \mathbb{R}^n$, $\forall t \in \mathbb{R}^+$. As such, with respect to $\bar{\mathbf{x}} = \mathbf{0}$ the DE given in Equation (A.14) is stable, but not asymptotically stable nor globally asymptotically stable. □

A.4 Numerical Solutions to Differential Equations

Rarely do nonlinear differential equations of the form $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$, $\mathbf{x}_0 = \mathbf{x}(0)$ have analytic solutions. As such, often DEs are solved numerically.

A.4.1 Euler Integration

Euler integration is the simplest (and least accurate) form of numerical integration. Owing to its conceptual simplicity (forward) Euler integration will be reviewed.

Euler integration begins by defining the step size of the numerical integration method. Let t_k denote the current time and define the step size as $h = t_{k+1} - t_k$. Next, let the current state at time t_k be

$$\mathbf{x}_k = \mathbf{x}(t_k).$$

Now approximate $\dot{\mathbf{x}}$ using a simple forward difference, that being

$$\dot{\mathbf{x}}(t) \doteq \frac{\mathbf{x}(t_{k+1}) - \mathbf{x}(t_k)}{t_{k+1} - t_k} = \frac{\mathbf{x}_{k+1} - \mathbf{x}_k}{h}.$$

Use the above approximation in $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$ gives

$$\begin{aligned} \frac{\mathbf{x}_{k+1} - \mathbf{x}_k}{h} &= \mathbf{f}(\mathbf{x}_k, t_k), \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + h\mathbf{f}(\mathbf{x}_k, t_k). \end{aligned}$$

Given $\mathbf{x}_1 = \mathbf{x}(0)$ at $t_1 = 0$, "march forward in time".

A.4.2 Fourth-Order Runge-Kutta

Fourth-order Runge-Kutta integration is based on a higher-order Taylor series approximation of $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t)$, and is much more accurate than Euler integration. Numerical integration is accomplished via

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{1}{6}\left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right)$$

where

$$\begin{aligned} \mathbf{k}_1 &= h\mathbf{f}(\mathbf{x}_k, t_k) \\ \mathbf{k}_2 &= h\mathbf{f}(\mathbf{x}_k + \frac{1}{2}\mathbf{k}_1, t_k + \frac{1}{2}h) \\ \mathbf{k}_3 &= h\mathbf{f}(\mathbf{x}_k + \frac{1}{2}\mathbf{k}_2, t_k + \frac{1}{2}h) \\ \mathbf{k}_4 &= h\mathbf{f}(\mathbf{x}_k + \mathbf{k}_3, t_k + h) \end{aligned}$$

Given $\mathbf{x}_1 = \mathbf{x}(0)$ at $t_1 = 0$, "march forward in time".

A.5 Notes

A.5.1 To Do

- Polar decomposition.
- $\lambda_{min}(\mathbf{P})\mathbf{x}^\mathsf{T}\mathbf{x} \leq \mathbf{x}^\mathsf{T}\mathbf{P}\mathbf{x} \leq \lambda_{max}(\mathbf{P})\mathbf{x}^\mathsf{T}\mathbf{x}$, see [11].
- Use stability and Lyapunov stability definitions presented in [11, pp. 65-85].
- Introduce LaSalle's Invariant Set Theorem; see [11, pp. 85].

Bibliography

[1] P. C. Hughes, *Spacecraft Attitude Dynamics*. Mineola, New York: Dover, 2nd ed., 2004.

[2] G. M. T. D'Eleutario and G. R. Heppler, *Newton's Second Law And All That (Preprint)*. Cambridge University Press, 2011.

[3] D. S. Bernstein, *Geometry, Kinematics, Statics, and Dynamcis (Preprint)*. Princeton University Press, 2013. February 18, 2013.

[4] A. V. Rao, *Dynamics of Particles and Rigid Bodies: A Systematic Approach*. New York, New York: Cambridge University Press, 2006.

[5] N. J. Kasdin and D. A. Paley, *Engineering Dynamics: A Comprehensive Introduction*. Princeton, New Jersey: Princeton University Press, 2011.

[6] R. L. Williams and D. A. Lawrence, *Linear State-Space Control Systems*. Hoboken, NJ: John Wiley & Sones, 2007.

[7] J. H. Ginsberg and J. Genin, *Statics and Dynamics: Combined Version*. Hoboken, New Jersey: John Wiley & Sons, Inc., 1977.

[8] R. C. Hibbeler, *Mechanics of Materials*. Upper Saddle River, New Jersey: Pearson Prentice Hall, 9th ed., 2014.

[9] C. D. Meyer, *Matrix Analysis and Applied linear Algebra*. Philadelphia, PA: SIAM, 2000.

[10] S. I. Grossman, *Elementary Linear Algebra*. Toronto: Harcourt College Publisher, 5th ed., 1994.

[11] H. J. Marquez, *Nonlinear Control Systems*. Hoboken, NJ: John Wiley and Sons, Inc., 2003.

[12] J. Gallier, *The Schur Complement and Symmetric Positive Semidefinite (and Definite) Matrices*. Course Notes, 2010.

[13] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1994.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[15] W. E. Boyce and R. C. Diprima, *Elementary Differential Equations and Boundary Value Problems*. Hoboken, NJ: John Wiley & Sons, 7th ed., 2003.

[16] P. J. Antsaklis and A. N. Michel, *A Linear Systems Primer*. Boston: Birkhauser, 2007.

[17] C. J.Damaren, "AER 1513 H – Spacecraft Dynamics and Control II Course Notes," *University of Toronto Institute For Aerospace Studies*, 2013.

[18] A. H. J. de Ruiter, C. J. Damaren, and J. R. Forbes, *Spacecraft Dynamics and Control: An Introduction*. West Sussex, UK: John Wiley & Sons, Ltd., 2013.

[19] L. Meirovitch, *Methods of Analytical Dynamics*. Toronto: McGraw-Hill, Inc., 1970.

[20] H. Schaub and J. L. Junkins, *Analytical Mechanics of Space Systems*. Reston, VA: American Institute of Aeronautics and Astronautics, Inc., 2nd ed., 2009.

[21] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems*. New York: Springer Science + Business Media, Inc., 2005.

[22] R. N. Murry, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulatrion*. Boca Raton, FL: CRC Press, Inc., 1993.

[23] M. J. Sidi, *Spacecraft Dynamics and Control: A practical Engineering Approach*. New York, NY: Cambridge University Press, 1997.

[24] T. R. Kane, P. W. Likins, and D. A. Levinson, *Spacecraft Dynamics*. Toronto: McGraw-Hill, 1983.

[25] G. E. Mase, *Theory and Problems of Continuum Mechanics (Schaum's Outline Series in Engineering)*. Toronto: McGraw-Hill, 1970.

[26] R. A. Tenenbaum, *Fundamentals of Applied Dynamics*. Springer-Verlag: New York, New York, 2004.

[27] W. T. Thomson, *Introduction to Space Dynamics*. Dover: Mineola, New York, 1986.

[28] N. A. Chaturvedi, A. K. Sanyal, and N. H. McClamroch, "Rigid-Body Attitude Control: Using Rotation Matrices for Continuous Singularity-Free Control Laws," *IEEE Control Systems Magazine*, vol. 31, pp. 30–51, June 2011.

[29] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control*. New York, New York: Springer, 2014.

[30] J. J. McPhee and R. N. Dubey, "Dynamics of Multihody Systems With Known Configuration Changes," *ASME Journal of Applied Mechanics*, vol. 58, pp. 215–221, March 1991.

[31] D. T. Greenwood, *Classical Dynamics*. Mineola, New York: Dover, 1997.

[32] J. L. Meriam and L. G. Kraige, *Engineering Mechanics: Dynamics*. Hoboken, New Jersey: John Wiley & Sons, Inc., 7th ed., 2012.

[33] A. Laulusa and O. A. Bauchau, "Review of Classical Approaches for Constraint Enforcement in Multibody Systems," *ASME Journal of Computational and Nonlinear Dynamics*, vol. 3, pp. 011004–(1–8), January 2008. doi: 10.1115/1.2803257.

[34] M. D. Schuster, "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, vol. 41, no. 4, pp. 439–517, 1993.

[35] P. Tsiotras, "Further Passivity Results for the Attitude Control Problem," *IEEE Transactions on Automatic Control*, vol. 43, no. 11, pp. 1597–1600, 1998.

[36] Y. Su and C. Zheng, “Globally Asymptotic Stabilization of Spacecraft with Simple Saturated Proportional-Derivative Control,” *AIAA Journal of Guidance, Control, and Dynamics*, vol. 34, no. 6, pp. 1932–1935, 2011.

[37] H. D. Curtis, *Orbital Mechanics for Engineering Students*. Oxford, UK: Butterworth-Heinemann / Elsevier, 2009.

[38] D. S. Bernstein, “Newton’s Frames,” *IEEE Control Systems Magazine*, vol. 28, pp. 17–19, February 2008.

[39] D. T. Greenwood, *Advanced Dynamics*. Cambridge University Press: Cambridge, 2003.

[40] H. Baruh, *Analytical Dynamics*. New York: McGraw-Hill, 1999.

[41] D. T. Greenwood, *Principles of Dynamics*. Englewood Cliffs, New Jersey: Prentice-Hall, 2nd ed., 1988.

[42] A. Bloch, *Nonholonomic Mechanics and Control*. New York: Springer-Verlag, 2003.