Fundamentals of Mechanical Systems
(with Application to Robotics)

by Tristan Perez

Lecture Notes MCHA3900–Mechatronic Design II

School of Engineering, Discipline of Mechanics and Mechatronics,

The University of Newcastle, AUSTRALIA

THE UNIVERSITY OF
NEWCASTLE
AUSTRALIA

Preface

These lecture notes revisit some of the fundamental concepts of classical mechanics[1] with the purpose of describing robotic systems—mobile robots (aerospace, marine and land vehicles) and robotic manipulators. The material is organised into four chapters.

Chapter 1 discusses the underlying ideas of the vectorial and analytical approaches to the study of mechanics. Chapter 2 reviews geometrical aspects of motion (kinematics). It defines the mathematical notation for vectors, reference frames and coordinate systems. It then discusses motion magnitudes and their kinematic transformations. The reader is assumed to have been exposed to simplified versions of these topics, which in these notes is presented in both tensor and matrix forms. The tensor form provides an economical way of writing and deriving kinematical relationships independent of the coordinate systems in which the vectors are ultimately expressed for computations. The matrix forms are linked to a coordinate system and are useful for implementation of computer simulation codes as well as robotic system analysis and design. Chapter 3 moves on to the Newton-Euler formulation of the equations of motion for rigid bodies in space. The equations of motion are formulated in body-fixed coordinates as commonly used to describe vehicle dynamics in aerospace, marine and land applications. We present an examples of aircraft and marine vehicle models.

Chapter 4 introduces concepts of analytical mechanics. Whilst vectorial mechanics provides a suitable tool for describing the motion of simple mechanical systems and free bodies in space, kinematic constraints often render it cumbersome. Analytical mechanics deals naturally with constraints and provides a more economical approach for deriving equations of motion. The definition of generalised coordinates and kinematic constraints are reviewed. We relate the Euler-Lagrange equations to both d'Alembert's and Hamilton's principle. This chapter also provides a brief introduction the canonical equations of Hamilton and Port-Hamiltonian Systems. The latter models have received a significant attention within the control community in the past two decades, for they provide a neat framework for energy-based control of non-linear mechanical systems. We provide a short introduction to this topic.

Tristan Perez
Newcastle, Australia, 2014

[1] Classical mechanics excludes relativistic and quantum mechanics. In this case of these lecture notes, the review of concepts also excludes mechanics of continuum.

Contents

Chapter 1

Vectorial and Analytical Mechanics

"If nature were not beautiful, it would not be worth knowing, and if nature were not worth knowing, life would not be worth living."

Henri Poincare (1854-1912)
French Mathematician, Physicist, and Philosopher

Mechanics deals with the dynamics of particles, bodies, and continua (fluids, elastic materials, thermodynamics). Mechanics has a rich history of milestone developments which span for over two thousand years. It has provided a basis for developments in different branches of science and engineering like astronomy, optics, and control theory. It has also motivated the development of calculus, and still today continues to motivate developments in mathematics like differential geometry and topology (Marsden and Ratiu, 2010). The analysis and design of robotic systems requires a solid background on mechanics. The aim of these lecture notes is to revisit some of the fundamental concepts of classical mechanics with the purpose of describing robotic systems—mobile robots (aerospace, marine and land vehicles) and robotic manipulators. The review of concepts has thus a specific focus which naturally shapes the way in which the material is presented and also the examples used to illustrate some of the concepts.

The following summary follows from the delightful work of Lanczos (1970), and it provides an overview of the relation among various topics. It is organised so as to highlight the links among the different developments rather than the actual chronological order in which the developments were made.

Newton (1642-1727), based on previous work of Leonardo (1452-1519) and Galileo (1564-1642), provided the three fundamental laws that govern the motion of particles: the first law is the law of inertia of Galileo, the second law relates the force and rate of change of linear momentum, and the third law deals with action and reaction forces. Euler (1707-1783) extended Newton's laws and described the motion of rigid bodies. He introduced the angular velocity as an auxiliary kinematic variable and the laws of conservation of linear momentum of the centre of mass and conservation of angular momentum. Because the magnitudes used by Newton and Euler are vectors (**momentum** and **force**), the study of mechanics based on their laws is normally referred to as **vectorial mechanics**.

Whilst Newton used momentum and force as fundamental quantities, Leibniz (1646-1716) proposed instead the use of two scalar quantities. One such quantity was *vis viva* (living force), which is twice of what we know today as the kinetic energy and the other quantity the work of the force, which is related to the potential energy. It may seem strange, but these two scalar quantities contain all the information about dynamics of most mechanical systems.

Hamilton's Principle—formulated by Hamilton (1805-1865) and related to the Principle of Least Action established in the early work of Euler (1707-1783) and Lagrange (1736-1813)—states that *the configuration*[1] *of a mechanical system described in terms of a set of generalised coordinates*[2] $q_1, q_2, \ldots, q_n$ *evolves in such a way that the definite integral*

$$I(q_1, q_2, \ldots, q_n) = \int_{t_1}^{t_2} \mathcal{L}(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n)\, dt \tag{1.1}$$

is minimised, where $\mathcal{L}$ *is the difference between the kinetic and potential energy of the system.*

The integral (1.1) is called the **action**, and the principle shows the evolution of mechanical systems in nature is such that the action is minimised.

Euler solved this minimisation problem for conservative systems using standard differential calculus, but Lagrange, who also solved it, invented a new branch of mathematics called **Calculus of Variations**. When we try to find whether a function $f(x)$ has a extremum at a point $\bar{x}$, we use the differential $\tilde{x} = \bar{x} + dx$ to check if the differential of the function $df = f(\tilde{x}) - f(\bar{x})$, caused by dx is zero. If this holds true, we say that the function has a **stationary point** at $\bar{x}$, and we need a further condition on the second differentials to establish the nature of the stationary point (maximum, minimum, or saddle). The minimisation of the definite integral (1.1) is with respect to the functions $q_j(t)$. That is, for every set of functions q_j we may choose, the integral gives us a number (because it is a definite integral), then we seek the functions q_j that give the smallest value of the integral. Lagrange devised the concept of variation δq, and expressed $\tilde{q} = \bar{q} + \delta q$. The variation δq is a function of our choice—a mathematical experiment imposed on q_j at a given time but not triggered by a dt. If the variation $\delta I = I(\tilde{q}_1, \ldots, \tilde{q}_n) - I(\bar{q}_1, \ldots, \bar{q}_n)$ vanishes, then $\bar{q}_j$ makes I stationary. Using this, Hamilton's principle can be written as a **variational principle**:

$$\delta I(q_1, q_2, \ldots, q_n) = 0. \tag{1.2}$$

For conservative systems in which the generalised coordinates are not subject to constraints, (1.2) leads to the **Euler-Lagrange equations**:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \qquad j = 1, 2, \ldots, n. \tag{1.3}$$

[1] The configuration of a mechanical system is known when the location of every particle of the system is known.

[2] The configuration can usually be described by a number of coordinates smaller than the actual number of particles. These are called generalised coordinates. The number of generalised coordinates is equal to the number of degrees of freedom of the mechanical system. The term generalised arises because these coordinates may not all have the same units, e.g., the coordinates can be displacements and angles.

The expansion of (1.3) gives the equations of motion of the system—a set of coupled second-order differential equations in q_j.

Hamilton also showed that the equations of Euler and Lagrange can be transformed into a state-space model. By defining the generalised momentum $p_j \triangleq \partial\mathcal{L}/\partial\dot{q}_j$, he found that

$$\dot{p}_j = -\frac{\partial\mathcal{H}}{\partial q_j}, \qquad \dot{q}_j = \frac{\partial\mathcal{H}}{\partial p_j}, \quad i = 1, 2, \ldots, n, \tag{1.4}$$

where $\mathcal{H}(\mathbf{p}, \mathbf{q})$ is the sum of the kinetic and potential energy of the system (total energy), and $\mathbf{p}$ and $\mathbf{q}$ are called energy variables.

The Hamiltonian model (1.4) proved to be a significant step in the development of mechanics and then other sciences. Jacobi (1804-1851) called Hamilton equations **Canonical Equations** and developed the theory of canonical transformations, which, allows one to solve the equations not by integration, but by coordinate transformations. This theory was used to relate problems in mechanics to other scientific problems, like for example Fermat's principle of least time in relation to the propagation of a ray of light.

The approach to the study of mechanics initiated by the ideas of Leibniz, put to solid foundations by Lagrange's calculus of variations, and then further elaborated by Hamilton and Jacobi is known as **Analytical Mechanics**—this name derives from Lagrange's book *Mécanique Analytique* published in 1788, in which all the developments are purely based on calculus (analysis).

For systems of free particles, both the vectorial and analytical approaches are equivalent. The advantage of the analytical approach is most evident for systems with kinematic constraints. Because Newton's approach isolates particles, the forces of constraints need to be considered, but these forces are unknown, and need to be found as part of the solution to the problem. Newton's third Law of action and reaction takes care of this for some systems, but fails, for example, to accommodate forces of electric and magnetic nature, and it is cumbersome to apply to complex systems. The analytical approach provides a mechanism to deal with the forces of constraint, and in some cases, these forces can be eliminated from the problem altogether. Furthermore, the variational principles are invariant to changes of generalised coordinates—it can be seen, for example, that a change of coordinates from q_j to q'_j should not affect the value of the integral (4.127). This is a very powerful property, for it is known that certain problems in mechanics are easier to solve if one choses an appropriate set of coordinates. Let us illustrate some of these concepts with a simple example.

Example 1 (Fixed Pendulum) *Consider the case of a pendulum made of a massless rigid bar and a mass particle. Let the pendulum hang from a fixed pivot as shown in Figure 1.1.*

Method 1 - Newton's Law

We can choose the cartesian coordinates of the mass to describe its position and write the following equations of motion based on Newton's Second Law:

$$m\,\ddot{x} = -F_T \sin\theta, \tag{1.5}$$
$$m\,\ddot{y} = mg - F_T \cos\theta, \tag{1.6}$$

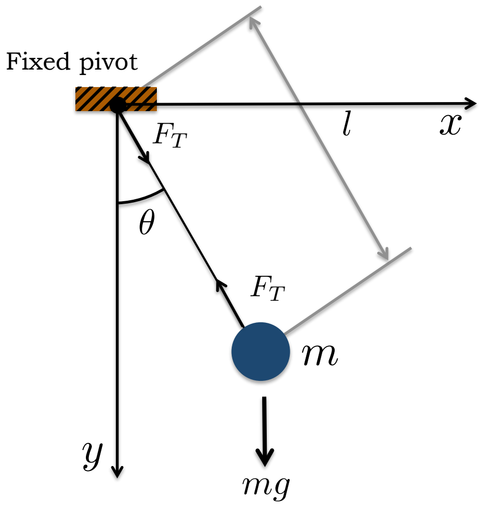


Figure 1.1: Pendulum hanging from a fixed point.

where F_T is the tension in the bar—force of constraint. The coordinates x and y, satisfy the following kinematic constraint:

$$l^2 = x^2 + y^2. \tag{1.7}$$

By using the fact that $\sin\theta = x/l$, $\cos\theta = y/l$, *and taking the first and second time derivatives of* (1.7), *namely,*

$$0 = x\,\dot{x} + y\,\dot{y}, \tag{1.8}$$

$$0 = \dot{x}^2 + x\,\ddot{x} + \dot{y}^2 + y\,\ddot{y}, \tag{1.9}$$

we can, after some algebraic work, eliminate the constraint force and obtain the following model:

$$\ddot{x} = -\frac{x(t)\,\dot{x}^2(t)}{l^2 - x^2} - \frac{x\sqrt{l^2 - x^2}}{l^2}g, \tag{1.10}$$

which is valid provided that $x \neq l$; namely $\theta \neq \pm\pi/2$. Note that if we simulate this model and obtain $x(t)$, we can then compute $y(t)$ from the constraint (1.7)*:*

$$y = \sqrt{l^2 - x^2}. \tag{1.11}$$

If we are designing the system, we may then wish to compute the force of constraint F_T using either (1.5) *or* (1.6)*.*

Method 2 - Lagrangian with dependent coordinates

Let us solve the same problem using Lagrangian mechanics. Indeed, let

$$q_1 \triangleq x, \qquad q_2 \triangleq y,$$

and the constraint

$$f(q_1, q_2) = \sqrt{q_1^2 + q_2^2} - l = 0. \tag{1.12}$$

The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2) + mg\,q_2. \tag{1.13}$$

Since q_1 and q_2 are constrained, we can use the Lagrange-multiplier method with the fact that the constraint (1.12) is holonomic (independent of $\dot{q}_i$):

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = \lambda\,\frac{\partial f}{\partial q_j}, \qquad j = 1, 2, \tag{1.14}$$

where λ is the undetermined multiplier. Note that λ has units of force—it is indeed the force of constraint.

The expansion of (1.14) leads to

$$m\,\ddot{q}_1 - \lambda\frac{q_1}{l} = 0, \tag{1.15}$$

$$m\,\ddot{q}_2 - mg - \lambda\frac{q_2}{l} = 0, \tag{1.16}$$

These equations together with the derivative of the constraint

$$\dot{q}_1^2 + q_1\ddot{q}_1 + \dot{q}_2^2 + q_2\ddot{q}_2 = 0, \tag{1.17}$$

form a system of equations that can be integrated. Substitution of (1.15) and (1.16) into (1.17) give

$$\lambda = -\frac{m}{l}(\dot{q}_1^2 + \dot{q}_2^2 + q_2\,g). \tag{1.18}$$

If we eliminite q_2 with the help of the constraint (1.17), we obtain exactly (1.10). However, (1.15) , (1.16), and (1.18) can be simulated.

Method 3 - Lagrangian with generalised coordinates (independent)

Since the constraint (1.12) is holonomic[3], we can find a reduced set of independent coordinates that describe the configuration. Indeed, since the system is a 1DOF system, we just need one coordinate. Hence, can choose the angle θ as a generalised coordinate q. The kinetic and potential energy can then be expressed in terms of q and $\dot{q}$:

$$\mathcal{T} = \frac{1}{2}\,m\,\vec{v}\cdot\vec{v} = \frac{1}{2}\,ml^2\dot{q}^2, \qquad \mathcal{V} = m\,g\,l\,(1-\cos q).$$

Hence the Lagrangian is

$$\mathcal{L} = \frac{1}{2}\,ml^2\dot{q}^2 - m\,g\,l\,(1-\cos q), \tag{1.19}$$

and we can use the Euler-Lagrange equation (1.3) since there are no constraints on q. This gives

$$\ddot{q} = -\frac{g}{l}\,\sin q. \tag{1.20}$$

In this case, we were able to ignore the constraint force altogether, which is just one of the nice features of the analytical approach. This method, however, does not provide any information about the force of constraint.

Method 4 - Hamiltonian

[3] We will define this properly as we advance in the text—Section 4.2.

From the Lagrangian (1.19), we can define the generalised momentum:

$$p \triangleq \frac{\partial \mathcal{L}}{\partial \dot{q}} = ml^2\,\dot{q}.$$

and the total energy becomes

$$\mathcal{H}(p,q) = \frac{1}{2}\,\frac{1}{ml^2}\,\dot{p}^2 + m\,g\,l\,(1-\cos q).$$

Hamilton Equations—see (1.4)—give the Hamiltonian model

$$\dot{p} = -mgl\,\sin q, \tag{1.21}$$

$$\dot{q} = \frac{1}{ml^2}\,p. \tag{1.22}$$

From a modelling point of view, the Hamiltonian model is not superior to the Lagrangian model. However, the Hamiltonian model highlights properties of the system that can be exploited for control.

□

As we can see in the example above, the analytical approach can be more economical when it comes to dealing with constrained systems (like a robot manipulator), and its invariance to coordinate changes can be used to choose coordinates that simplify the solution of the equations of motion. The analytical approach, however, can only deal with monogenic forces—forces that derive from the gradient of a potential function—and this excludes, for example, friction forces. The vectorial approach does not restrict the nature of the force; and therefore, it has no problems in dealing with non-monogenic forces. One should not take a sides on either approach, but rather attempt to master them both, for there are problems in which one approach is better suited than the other.

In the rest of these lecture notes, we will review and elaborate on both vectorial and analytical approaches. If you did not understand all details in Example 1, do not worry, for we will come back to it as we progress with the lecture notes.

Chapter 2

Kinematics

"Geometry existed before the Creation. It is co-eternal with the mind of God. Geometry provided God with a model for the Creation. Geometry is God Himself."

Johannes Kepler (1571-1630)
German Mathematician and Astronomer

Kinematics refers to the geometrical description of motion regardless of the forces causing the motion. This chapter introduces rigid-body kinematics. In particular, we define the reference frames from which the motion can be described and their associated coordinate systems; define the different vector quantities that characterise the motion; and describe the transformations that link the representation of vectors in different coordinate systems. We then consider specific examples of vehicle navigation and robot kinematics.

2.1 Vector Magnitudes and Basic Operations

The description of motion of a body requires the use of vector magnitudes in a three-dimensional space: position, velocity, acceleration, and force.

We will denote a physical vector magnitude as $\vec{u}$. Vectors admit the operations of addition and scalar multiplication:

$$\vec{w} = \vec{u} + \vec{v}, \qquad \vec{v} = a\,\vec{u}.$$

A **basis** is a maximal set of linear independent vectors. In the Euclidean three-dimensional space, every basis has 3 vectors. That is, a basis in the three-dimensional space, denoted by $\{a\}$, is a set $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ for which

$$\alpha\,\vec{a}_1 + \beta\,\vec{a}_2 + \gamma\,\vec{a}_3 = \vec{0} \quad \Rightarrow \quad \alpha = \beta = \gamma = 0.$$

If the vectors in a basis are mutually perpendicular, the basis is **orthogonal**.

Any vector $\vec{u}$ can then be expressed as a **linear combination** of the elements of the basis:

$$\vec{u} = u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3. \tag{2.1}$$

The terms on the right-hand side of (2.1) are called the **components** of the vector in the basis, and the scalars u_i^a are called the **coordinates** of the vector in the basis $\{a\}$. The right-upper script ^{a} indicates that the coordinates are associated with the basis $\{a\}$. If we choose a different basis, in general, the coordinates will change.

The **Euclidean norm** or **Cartesian metric** of a vector expressed in terms of an orthogonal basis is defined as

$$\|\vec{u}\| = \sqrt{(u_1^a)^2 + (u_2^a)^2 + (u_3^a)^2}. \tag{2.2}$$

A vector with norm equal to one is called a **unit vector**. If an orthogonal basis is made of unit vectors, then the basis is called **orthonormal**.

The norm (2.2) is induced by the following **inner product**:

$$\vec{u} \cdot \vec{v} = (u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3) \cdot (v_1^a\,\vec{a}_1 + v_2^a\,\vec{a}_2 + v_3^a\,\vec{a}_3) = u_1^a\,v_1^a + u_2^a\,v_2^a + u_3^a\,v_3^a. \tag{2.3}$$

That is, (2.2) can be expressed as

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}.$$

Using norms, the inner product (2.3) can be also expressed as

$$\vec{u} \cdot \vec{v} = \|\vec{u}\|\,\|\vec{v}\|\,\cos\varphi,$$

where φ denotes the angle between the two vectors. If two vectors are **orthogonal**, their inner product is zero.

For physical vectors, we can also define the **cross product**:

$$\vec{u} = u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3,$$

$$\vec{v} = v_1^a\,\vec{a}_1 + v_2^a\,\vec{a}_2 + v_3^a\,\vec{a}_3,$$

$$\vec{w} = \vec{u} \times \vec{v} = \begin{vmatrix} u_2^a & u_3^a \\ v_2^a & v_3^a \end{vmatrix} \vec{a}_1 - \begin{vmatrix} u_1^a & u_3^a \\ v_1^a & v_3^a \end{vmatrix} \vec{a}_2 + \begin{vmatrix} u_1^a & u_2^a \\ v_1^a & v_2^a \end{vmatrix} \vec{a}_3 \tag{2.4}$$

where $|\cdot|$ denotes the determinant. This operation takes two vectors, $\vec{u}$ and $\vec{v}$, and transforms them into a vector $\vec{w}$ which is perpendicular to both $\vec{u}$ and $\vec{v}$, and

$$\|\vec{w}\| = \|\vec{u}\|\,\|\vec{v}\|\,\sin\varphi,$$

where φ denotes the angle between $\vec{u}$ and $\vec{v}$.

2.2 Matrix Representation of Vectors

Once a basis is chosen, vectors can be represented as column matrices in $\mathbb{R}^3$ (space of column matrices) and operations be performed using matrix algebra. That is, the vectors $\vec{u}$ and $\vec{v}$ expressed the basis $\{a\}$, can be represented as

$$\mathbf{u}^a = \begin{bmatrix} u_1^a \\ u_2^a \\ u_3^a \end{bmatrix}, \quad \mathbf{v}^a = \begin{bmatrix} v_1^a \\ v_2^a \\ v_3^a \end{bmatrix}.$$

The inner product can be computed as

$$\vec{u}\cdot\vec{v} \qquad \equiv \qquad (\mathbf{u}^a)^{\mathsf{T}}\,\mathbf{v}^a = \begin{bmatrix} u_1^a & u_2^a & u_3^a \end{bmatrix} \begin{bmatrix} v_1^a \\ v_2^a \\ v_3^a \end{bmatrix}. \tag{2.5}$$

Note that in order to make computations, all coordinates must be related to the same basis.

Similarly, the cross product can be computed as

$$\vec{w} = \vec{u}\times\vec{v} \qquad \equiv \qquad \mathbf{w}^a = \mathbf{S}(\mathbf{u}^a)\,\mathbf{v}^a, \tag{2.6}$$

where $\mathbf{S}(\cdot)$ is the **Skew-symmetric operator** for vectors in $\mathbb{R}^3$, that is $\mathbf{S}(\cdot) = -\mathbf{S}^{\mathsf{T}}(\cdot)$:

$$\mathbf{S}(\mathbf{u}) = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}. \tag{2.7}$$

Remarks:

- Vector magnitudes can be expressed in different basis. The notation $\vec{u}$ is independent of the basis chosen; however, the matrix representation $\mathbf{u}^a$ is always relative to a basis, and this is indicated explicitly in our notation.

- The advantage of using a the notation $\vec{u}$ is that we can express equations that hold independent of the basis chosen to represent the vectors. For example, the equation $\vec{w} = \vec{u}\times\vec{v}$ holds in any basis we may choose to express the three vector magnitudes. Similarly, the norm of a vector $\|\vec{u}\|$ and the inner product of two vectors $\vec{u}\cdot\vec{v}$ are independent of the basis chosen to express the vectors. This will prove to be useful to derive equations of motion.

- The matrix representation simplifies the tasks of developing computational models for simulation and analysing properties related to the motion of mechanical systems.

2.3 Coordinate Transformations

A vector $\vec{u}$ can be represented in a basis $\{a\}$ and also in a basis $\{b\}$:

$$\vec{u} = u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3 = u_1^b\,\vec{b}_1 + u_2^b\,\vec{b}_2 + u_3^b\,\vec{b}_3. \tag{2.8}$$

A natural question to ask at this point is *how are the coordinates related?* To answer this question, we can start by expressing the coordinates in $\{b\}$ as follows:

$$u_1^b = \vec{u}\cdot\vec{b}_1, \quad u_2^b = \vec{u}\cdot\vec{b}_2, \quad u_3^b = \vec{u}\cdot\vec{b}_3, \tag{2.9}$$

which expands to

$$u_1^b = (u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3)\cdot\vec{b}_1, \tag{2.10}$$
$$u_2^b = (u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3)\cdot\vec{b}_2, \tag{2.11}$$
$$u_3^b = (u_1^a\,\vec{a}_1 + u_2^a\,\vec{a}_2 + u_3^a\,\vec{a}_3)\cdot\vec{b}_3. \tag{2.12}$$

The latter can be arranged as follows:

$$\begin{bmatrix} u_1^b \\ u_2^b \\ u_3^b \end{bmatrix} = \begin{bmatrix} (\vec{a}_1 \cdot \vec{b}_1) & (\vec{a}_2 \cdot \vec{b}_1) & (\vec{a}_3 \cdot \vec{b}_1) \\ (\vec{a}_1 \cdot \vec{b}_2) & (\vec{a}_2 \cdot \vec{b}_2) & (\vec{a}_3 \cdot \vec{b}_2) \\ (\vec{a}_1 \cdot \vec{b}_3) & (\vec{a}_2 \cdot \vec{b}_3) & (\vec{a}_3 \cdot \vec{b}_3) \end{bmatrix} \begin{bmatrix} u_1^a \\ u_2^a \\ u_3^a, \end{bmatrix} \tag{2.13}$$

or in matrix notation

$$\mathbf{u}^b = \mathbf{C}_a^b \, \mathbf{u}^a, \tag{2.14}$$

where $\mathbf{C}_a^b$ is called a **coordinate transformation matrix** and the adopted notation is $\mathbf{C}_{from}^{to}$. The columns of $\mathbf{C}_a^b$ are the direction cosines of the vectors $\vec{a}_i$ in the basis $\{b\}$.

To derive (2.14), we have expressed u_i^b in terms of u_i^a, $\vec{a}_i$ and $\vec{b}_i$. If we alternatively express u_i^a in terms of u_i^b, $\vec{a}_i$ and $\vec{b}_i$, we obtain

$$\mathbf{u}^a = \mathbf{C}_b^a \, \mathbf{u}^b = (\mathbf{C}_a^b)^{-1} \, \mathbf{u}^b. \tag{2.15}$$

It we do this, we find that

$$\mathbf{C}_b^a = (\mathbf{C}_a^b)^{-1} = (\mathbf{C}_a^b)^T. \tag{2.16}$$

The last equality shows that the coordinate transformation matrix $\mathbf{C}_a^b$ is **orthogonal**:

$$\mathbf{C}_a^b \, (\mathbf{C}_a^b)^T = \mathbf{I}. \tag{2.17}$$

This is a nice property: if we want the inverse of a coordinate transformation matrix, we just need to transpose it.

We can now extend our high-school definitions of scalars and vector magnitudes:

A magnitude whose coordinates in different basis are the same is called a **scalar** or **zeroth-order tensor**.

A magnitude whose coordinates in different bases are linearly related according to

$$u_i^b = \sum_j C_{ij} \, u_j^a,$$

where C_{ij} are the elements of the coordinate transformation matrix, is called a **vector** or **first-order tensor**.

As we advance in these lecture notes, we will also find second-order tensors. When second-order tensors are expressed in terms of bases, they become 3×3 matrices. An example is the cross-product tensor:

$$\vec{w} = \vec{S}_u \cdot \vec{v} \qquad \equiv \qquad \vec{w} = \vec{u} \times \vec{v} \qquad \equiv \qquad \mathbf{w}^a = \mathbf{S}(\mathbf{u}^a) \, \mathbf{v}^a,$$

Here, the product with the tensor $\vec{S}_u$ maps the vector $\vec{v}$ into $\vec{w}$. Another important second-order tensor that we will use it the inertia tensor, which relates the angular velocity to the angular momentum of a rigid body:

$$\vec{L} = \vec{I} \cdot \vec{\omega}.$$

Choosing a basis and expressing this last equation in matrix form will allow us, for example, to find the principal axis of inertia, which are given by the eigenvectors of the inertia matrix.

2.4 Reference Frames and Coordinate Systems

In order to describe the motion of rigid bodies, we need a reference frame and a coordinate system.

> A **reference frame** is perspective from which the motion is described by an observer. A reference frame can be defined by a set of at least 3 non-colinear points that are rigidly connected.

Hence a rigid body can be considered a reference frame, but not every reference frame is a rigid body. For example, in most robotic applications, the robot motion is described using the Earth as a reference frame (a rigid body). In space-craft navigation, however, it is common to use a reference frame defined in terms of the sun and other stars of our galaxy (not a rigid body). Reference frames will be denoted with calligraphic font: $\mathcal{A}$, $\mathcal{B}$, *etc.*

To quantify motion, we associate coordinate systems to reference frames:

> A **coordinate system** is a mathematical entity that allows us to establish a one-to-one correspondence between vector magnitudes and scalars called coordinates. A basis defines a coordinate system.

In an Euclidean space (flat space), a **Cartesian coordinate system** is a basis of vectors $\{a\}$ such that the modulus of a vector $\vec{u}$ expressed as a linear combination of the basis vectors is given by the Cartesian metric (2.2). That is, the basis of a Cartesian coordinate system is orthogonal. The Cartesian coordinate system allows us to obtain the coordinates associated with $\vec{u}$, which we normally represent as an element in $\mathbb{R}^3$ (a column matrix). Note that we can have non Cartesian coordinate systems in the Euclidean space; examples of these are cylindrical and spherical coordinate systems. If we express a vector using these coordinate systems, the norm of the vector is not given by (2.2) anymore.

> Remember that a reference frame is not the same as a coordinate system, one is physical and the other is mathematical. Note also that coordinate systems are bases and they have no origins.

In certain applications we may use different sets of coordinate systems associated with a single reference frame. This is common in vehicles, where the vehicle is a reference frame, and the different sensors used for measuring motion magnitudes can have a different coordinate systems.

2.5 Time-Derivatives of Vectors

> The time derivative or rate of change of a scalar magnitude is the same in every frame. The time derivative or rate of change of a vector magnitude depends, in general, on the reference frame in which the vector is being observed.

We will use a notation that indicates this explicitly:

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \lim_{\Delta t \to 0}^{\mathcal{A}} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{\Delta t}, \tag{2.18}$$

where $\lim^{\mathcal{A}}_{\Delta t \to 0}$ indicates that the increment $\vec{r}(t+\Delta t) - \vec{r}(t)$ is observed in $\{a\}$.

Because of (2.18), we can talk about the *time derivative of* $\vec{r}$ *in* $\mathcal{A}$ and the *time derivative of* $\vec{r}$ *in* $\mathcal{B}$. Consider the scenario depicted in Figure 2.1. Let $\{a\}$ be a basis in $\mathcal{A}$, $\{b\}$ a basis in $\mathcal{B}$. Then the time derivatives of $\vec{r}$ in $\mathcal{A}$ and $\mathcal{B}$ are

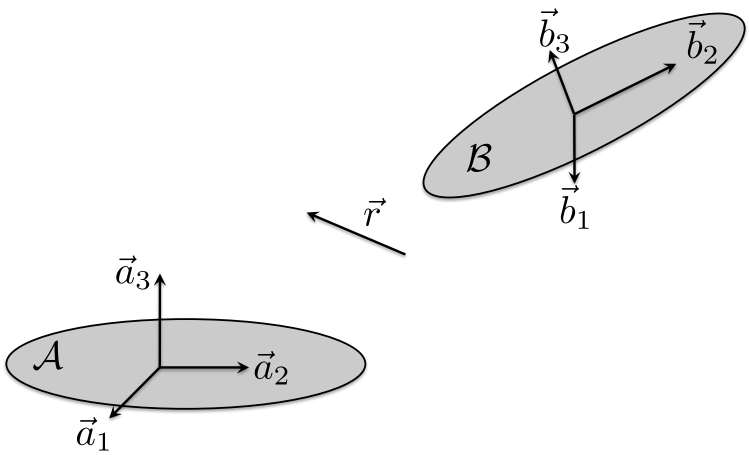


Figure 2.1: A vector observed in two different frames.

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{dr_1^a}{dt}\vec{a}_1 + \frac{dr_2^a}{dt}\vec{a}_2 + \frac{dr_3^a}{dt}\vec{a}_3, \tag{2.19}$$

$$\frac{{}^{\mathcal{B}}d\vec{r}}{dt} = \frac{dr_1^b}{dt}\vec{b}_1 + \frac{dr_2^b}{dt}\vec{b}_2 + \frac{dr_3^b}{dt}\vec{b}_3. \tag{2.20}$$

In general, the rate of change of a vector in different frames will be different, that is,

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} \neq \frac{{}^{\mathcal{B}}d\vec{r}}{dt}.$$

Example 2 (Derivatives in different frames) *Suppose that* $\mathcal{B}$ *in Figure 2.1 rotates with respect to* $\mathcal{A}$*, and that* $\vec{r}$ *is a vector between two points fixed in* $\mathcal{A}$*, then*

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \vec{0}, \qquad \frac{{}^{\mathcal{B}}d\vec{r}}{dt} \neq \vec{0}.$$

□

This simple example shows us that the time derivative of a vector depends on the reference frame. Therefore, we should always make explicit or clarify in which frame the derivatives are observed.

The following relations hold for the derivatives of vectors:

$$\begin{aligned}
\frac{{}^{\mathcal{A}}d}{dt}(\vec{u}+\vec{v}) &= \frac{{}^{\mathcal{A}}d\vec{u}}{dt} + \frac{{}^{\mathcal{A}}d\vec{v}}{dt},\\
\frac{{}^{\mathcal{A}}d}{dt}(k\,\vec{u}) &= k\,\frac{{}^{\mathcal{A}}d\vec{u}}{dt},\\
\frac{{}^{\mathcal{A}}d}{dt}(\vec{u}\cdot\vec{v}) &= \frac{{}^{\mathcal{A}}d\vec{u}}{dt}\cdot\vec{v} + \vec{u}\cdot\frac{{}^{\mathcal{A}}d\vec{v}}{dt},\\
\frac{{}^{\mathcal{A}}d}{dt}(\vec{u}\times\vec{v}) &= \frac{{}^{\mathcal{A}}d\vec{u}}{dt}\times\vec{v} + \vec{u}\times\frac{{}^{\mathcal{A}}d\vec{v}}{dt}
\end{aligned}$$

2.6 Angular Velocity and Transport Theorem

The following theorem defines the angular velocity, and it is fundamental to derive kinematic relationships and equations of motion of mechanical systems.

Theorem 1 (Rate of Change Transport Theorem) *Consider the scenario depicted in the Figure 2.1, and assume that $\mathcal{B}$ only rotates with respect to $\mathcal{A}$. Then there exist a unique vector $\vec{\omega}_{\mathcal{B}/\mathcal{A}}$ called the* **angular velocity** *of $\mathcal{B}$ with respect to $\mathcal{A}$ such that*

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{{}^{\mathcal{B}}d\vec{r}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}}\times\vec{r}. \tag{2.21}$$

We will demonstrate (2.21)—it is a good exercise, which reviews the use several properties of vector algebra. The following proof is similar to that by Rao (2006):

Proof 1 (Rate of Change Transport Theorem) *Let us express $\vec{r}$ in $\{b\}$,*

$$\vec{r} = r_1^b\,\vec{b}_1 + r_2^b\,\vec{b}_2 + r_3^b\,\vec{b}_3,$$

and observe it rate of change in $\mathcal{A}$:

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{dr_1^b}{dt}\,\vec{b}_1 + \frac{dr_2^b}{dt}\,\vec{b}_2 + \frac{dr_3^b}{dt}\,\vec{b}_3 + r_1^b\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} + r_2^b\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} + r_3^b\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt}. \tag{2.22}$$

Note that since we are making observations in $\mathcal{A}$, the vectors of the basis $\{b\}$ are not constant due to the rotation of $\mathcal{B}$ with respect to $\mathcal{A}$. Note also that r_i^b are scalar so their time derivatives do not depend on the frame in which they are observed.

The first three terms on the right-hand side of (2.22) are the components of the derivative of $\vec{r}$ observed in $\mathcal{B}$. Thus,

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{{}^{\mathcal{B}}d\vec{r}}{dt} + r_1^b\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} + r_2^b\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} + r_3^b\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt}, \tag{2.23}$$

So now we need to concentrate on the last three terms. The time derivative of a vector with constant magnitude is perpendicular to the vector itself. For example, consider the following scalar product:

$$\vec{b}_1\cdot\vec{b}_1 = \|\vec{b}\|^2 = 1,$$

and let us take the time derivative:

$$\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} \cdot \vec{b}_1 + \vec{b}_1 \cdot \frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} = 2\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} \cdot \vec{b}_1 = 0.$$

This implies that ${}^{\mathcal{A}}d\vec{b}_1/dt$ *is perpendicular to* $\vec{b}_1$. *Therefore,* ${}^{\mathcal{A}}d\vec{b}_1/dt$ *must be in the plane that contains* $\vec{b}_2$ *and* $\vec{b}_3$. *Then,* ${}^{\mathcal{A}}d\vec{b}_1/dt$ *can be expressed as a linear combination of* $\vec{b}_2$ *and* $\vec{b}_3$. *The same applies for the other two vectors that form the basis* $\{b\}$. *Then,*

$$\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} = c_{12}\,\vec{b}_2 + c_{13}\,\vec{b}_3, \tag{2.24}$$
$$\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} = c_{21}\,\vec{b}_1 + c_{23}\,\vec{b}_3, \tag{2.25}$$
$$\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} = c_{31}\,\vec{b}_1 + c_{32}\,\vec{b}_2, \tag{2.26}$$

where c_{ij} *are scalar functions that we seek to determine. Because the basis* $\{b\}$ *is orthogonal, we have that*

$$\vec{b}_i \cdot \vec{b}_j = 0, \quad i \neq j. \tag{2.27}$$

Taking the time derivative of (2.27), *we obtain*

$$\frac{{}^{\mathcal{A}}d\vec{b}_i}{dt} \cdot \vec{b}_j + \vec{b}_i \cdot \frac{{}^{\mathcal{A}}d\vec{b}_j}{dt} = 0, \quad i \neq j,$$

or alternatively,

$$\frac{{}^{\mathcal{A}}d\vec{b}_i}{dt} \cdot \vec{b}_j = -\vec{b}_i \cdot \frac{{}^{\mathcal{A}}d\vec{b}_j}{dt}, \quad i \neq j. \tag{2.28}$$

Let us apply this to (2.24) *to* (2.26),

$$\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} \cdot \vec{b}_2 = c_{12} = \;-\vec{b}_1 \cdot \frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} = -c_{21}, \tag{2.29}$$

$$\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} \cdot \vec{b}_3 = c_{23} = \;-\vec{b}_2 \cdot \frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} = -c_{32}, \tag{2.30}$$

$$\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} \cdot \vec{b}_1 = c_{31} = \;-\vec{b}_3 \cdot \frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} = -c_{13}, \tag{2.31}$$

Let us define,

$$\omega_1^b \triangleq c_{23} = -c_{32}, \tag{2.32}$$
$$\omega_2^b \triangleq c_{31} = -c_{13}, \tag{2.33}$$
$$\omega_3^b \triangleq c_{12} = -c_{21}, \tag{2.34}$$

and

$$\vec{\omega}_{\mathcal{B}/\mathcal{A}} \triangleq \omega_1^b\,\vec{b}_1 + \omega_2^b\,\vec{b}_2 + \omega_3^b\,\vec{b}_3. \tag{2.35}$$

Then, (2.24) to (2.26) can be expressed

$$\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} = \omega_3^b\,\vec{b}_2 - \omega_2^b\,\vec{b}_3 = \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{b}_1, \tag{2.36}$$

$$\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} = -\omega_3^b\,\vec{b}_1 + \omega_1^b\,\vec{b}_3 = \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{b}_2, \tag{2.37}$$

$$\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} = \omega_2^b\,\vec{b}_1 - \omega_1^b\,\vec{b}_2 = \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{b}_3. \tag{2.38}$$

Substituting these expressions into (2.23), we obtain

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{{}^{\mathcal{B}}d\vec{r}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times (r_1^b\,\vec{b}_1 + r_2^b\,\vec{b}_2 + r_3^b\,\vec{b}_3), \tag{2.39}$$

which gives the result (2.21):

$$\frac{{}^{\mathcal{A}}d\vec{r}}{dt} = \frac{{}^{\mathcal{B}}d\vec{r}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}. \tag{2.40}$$

■

Let us investigate a bit more some properties of the angular velocity vector. First let us note that (2.21) is valid in general for frames that rotate relative to each other. Therefore, we can also write

$$\frac{{}^{\mathcal{B}}d\vec{r}}{dt} = \frac{{}^{\mathcal{A}}d\vec{r}}{dt} + \vec{\omega}_{\mathcal{A}/\mathcal{B}} \times \vec{r}. \tag{2.41}$$

A comparison of (2.40) and (2.41), gives

$$\vec{\omega}_{\mathcal{B}/\mathcal{A}} = -\vec{\omega}_{\mathcal{A}/\mathcal{B}}. \tag{2.42}$$

We can use (2.36) to (2.38) to determine the coordinates of $\vec{\omega}_{\mathcal{B}/\mathcal{A}}$ in $\{b\}$ by taking inner products:

$$\omega_1 = \frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} \cdot \vec{b}_3, \tag{2.43}$$

$$\omega_2 = \frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} \cdot \vec{b}_1, \tag{2.44}$$

$$\omega_3 = \frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} \cdot \vec{b}_2, \tag{2.45}$$

and hence

$$\vec{\omega}_{\mathcal{B}/\mathcal{A}} = \left(\frac{{}^{\mathcal{A}}d\vec{b}_2}{dt} \cdot \vec{b}_3\right)\vec{b}_1 + \left(\frac{{}^{\mathcal{A}}d\vec{b}_3}{dt} \cdot \vec{b}_1\right)\vec{b}_2 + \left(\frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} \cdot \vec{b}_2\right)\vec{b}_3. \tag{2.46}$$

Some books take (2.46) as a definition.

Example 3 (Rotation about a coordinate axis) *Consider the scenario depicted in Figure 2.2, where the frame $\mathcal{B}$ (with associated basis $\{b\}$) rotates relative to the frame $\mathcal{A}$ (with associated basis $\{a\}$). The rotation is is at a constant angular rate ω_3 about the coordinate direction given by $\vec{a}_3$.*

The derivative ${}^{\mathcal{A}}d\vec{b}_2/dt$ is normal to $\vec{b}_3$, so the first term of (2.46) is zero. The derivative ${}^{\mathcal{A}}d\vec{b}_3/dt$ is zero, so the second term of (2.46) is also zero. The derivative ${}^{\mathcal{A}}d\vec{b}_1/dt$ is in the direction of $\vec{b}_2$, so the scalar product in the last term is non zero. This indicates that the vector $\vec{\omega}_{b/a}$ is directed along $\vec{b}_3$, which aligns with the axis of rotation. The differential ${}^{\mathcal{A}}d\vec{b}_1$ can be expressed in terms of the arc length (angle $\times$ radius), and we note that it is in the direction of $\vec{b}_2$:

$$ {}^{\mathcal{A}}d\vec{b}_1 = \left(\omega_3 dt \, \|\vec{b}_1\|\right) \vec{b}_2, $$

and therefore

$$ \frac{{}^{\mathcal{A}}d\vec{b}_1}{dt} = \omega_3 \, \vec{b}_2. $$

Substituting this in the last term of (2.46) and doing the dot product with $\vec{b}_2$, we obtain

$$ \vec{\omega}_{\mathcal{B}/\mathcal{A}} = \omega_3 \, \vec{b}_3. $$

The latter is in agreement of what you have seen in elementary courses of mechanics, namely, that in a single axis rotation, the vector of angular velocity is directed along the axis of rotation.

□

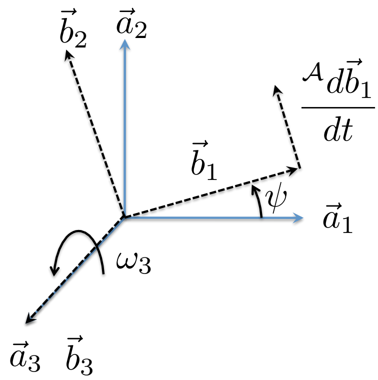


Figure 2.2: Rotation of $\mathcal{B}$ relative to $\mathcal{A}$.

2.7 Addition of Angular Velocities

Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ be four different reference frames. Then,

$$ \frac{{}^{\mathcal{A}}d\vec{u}}{dt} = \frac{{}^{\mathcal{B}}d\vec{u}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{u}, \tag{2.47} $$

$$ \frac{{}^{\mathcal{B}}d\vec{u}}{dt} = \frac{{}^{\mathcal{C}}d\vec{u}}{dt} + \vec{\omega}_{\mathcal{C}/\mathcal{B}} \times \vec{u}, \tag{2.48} $$

$$ \frac{{}^{\mathcal{C}}d\vec{u}}{dt} = \frac{{}^{\mathcal{D}}d\vec{u}}{dt} + \vec{\omega}_{\mathcal{D}/\mathcal{C}} \times \vec{u}. \tag{2.49} $$

If we substitute these backwards from the last expression, we obtain

$$\frac{{}^{\mathcal{A}}d\vec{u}}{dt} = \frac{{}^{\mathcal{D}}d\vec{u}}{dt} + (\vec{\omega}_{\mathcal{D}/\mathcal{C}} + \vec{\omega}_{\mathcal{C}/\mathcal{B}} + \vec{\omega}_{\mathcal{B}/\mathcal{A}}) \times \vec{u}, \tag{2.50}$$

which leads to

$$\vec{\omega}_{\mathcal{D}/\mathcal{A}} = \vec{\omega}_{\mathcal{D}/\mathcal{C}} + \vec{\omega}_{\mathcal{C}/\mathcal{B}} + \vec{\omega}_{\mathcal{B}/\mathcal{A}}. \tag{2.51}$$

To remember this rule, note that with the adopted notation, the multiplication of the subscripts on the right-hand side simplifies to that on the left-hand side.

2.8 Rotation Tensors and Matrices

Consider the case depicted in Figure 2.3, where the vector $\vec{v}$ is obtained by rotating the vector $\vec{u}$ an angle μ about the rotation axis given by the direction of the unit vector $\vec{n}$. The angle μ is positive according to the right-hand-side convention about $\vec{n}$.

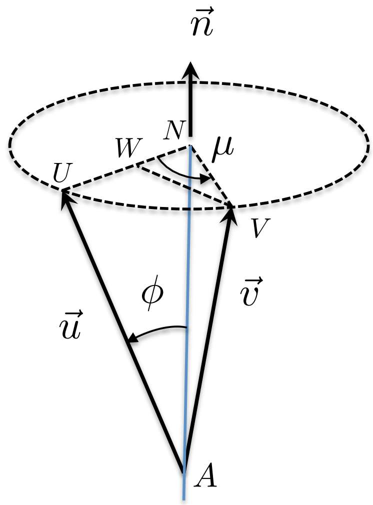


Figure 2.3: Rotation of a vector $\vec{u}$ about the axis $\overline{AN}$.

We can express $\vec{v}$ as follows:

$$\begin{aligned}\vec{v} &= \vec{r}_{N/A} + \vec{r}_{W/N} + \vec{r}_{V/W}, \\ &= (\vec{u}\cdot\vec{n})\,\vec{n} + \frac{\vec{u} - (\vec{u}\cdot\vec{n})\vec{n}}{\|\vec{u} - (\vec{u}\cdot\vec{n})\vec{n}\|}\|\vec{r}_{V/N}\|\,\cos\mu + \frac{\vec{n}\times\vec{u}}{\|\vec{u}\|\,\sin\phi}\|\vec{r}_{V/N}\|\,\sin\mu.\end{aligned}$$

Since,

$$\|\vec{r}_{V/N}\| = \|\vec{u} - (\vec{u}\cdot\vec{n})\vec{n}\| = \|\vec{u}\|\,\sin\phi,$$

then

$$\vec{v} = (\vec{u}\cdot\vec{n})\,\vec{n} + (\vec{u} - (\vec{u}\cdot\vec{n})\vec{n})\,\cos\mu + (\vec{n}\times\vec{u})\,\sin\mu.$$

Re-arranging terms, we obtain the following **Rotation Formula**:

$$\vec{v} = (1 - \cos\mu)\,(\vec{u} \cdot \vec{n})\,\vec{n} + \vec{u}\cos\mu + (\vec{n} \times \vec{u})\,\sin\mu, \tag{2.52}$$

where μ is positive according to the right-hand rule about $\vec{n}$.

Expression (2.52) is a **Rotation Tensor**, which takes the vector $\vec{u}$ and maps it into the vector $\vec{v}$:

$$\vec{v} = \vec{R}(\vec{n}, \mu) \cdot \vec{u}. \tag{2.53}$$

We should interpret the product in (2.53) as a tensor operation on a vector.

If we express all vectors of Figure 2.3 in the basis $\{a\}$, we obtain

$$\mathbf{v}^a = [(1 - \cos\mu)\mathbf{n}^a(\mathbf{n}^a)^\mathsf{T} + \cos\mu\,\mathbf{I} + \mathbf{S}(\mathbf{n}^a)\,\sin\mu]\mathbf{u}^a, \tag{2.54}$$

the rotation tensor then becomes a **Rotation Matrix**, which relates the coordinates of $\vec{u}$ in $\{a\}$ to the coordinates of $\vec{v}$ in $\{a\}$:

$$\mathbf{v}^a = \mathbf{R}(\mathbf{n}^a, \mu)\mathbf{u}^a, \tag{2.55}$$

$$\mathbf{R}(\mathbf{n}^a, \mu) \triangleq [(1 - \cos\mu)\mathbf{n}^a(\mathbf{n}^a)^\mathsf{T} + \cos\mu\,\mathbf{I} + \mathbf{S}(\mathbf{n}^a)\,\sin\mu]. \tag{2.56}$$

Note that by expressing the tensor in a basis, the product in (2.53) becomes a matrix multiplication.

Let us consider the **single rotations** about the coordinate axis,

- A rotation of an angle ψ about the z-axis ($\mathbf{n}^a = [0, 0, 1]^\mathsf{T}$ and $\mu = \psi$ in (2.56)):

$$\mathbf{R}_{z,\psi} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.57}$$

- A rotation of an angle θ about the y-axis ($\mathbf{n}^a = [0, 1, 0]^\mathsf{T}$ and $\mu = \theta$ in (2.56)):

$$\mathbf{R}_{y,\theta} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}. \tag{2.58}$$

- A rotation of an angle ϕ about the x-axis ($\mathbf{n}^a = [1, 0, 0]^\mathsf{T}$ and $\mu = \phi$ in (2.56)):

$$\mathbf{R}_{x,\phi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}. \tag{2.59}$$

The following are simple rules for building single-axis rotation matrices:

For a rotation of an angle μ about the axis i, (i =1,2, or 3),

1. Place a 1 on the diagonal entry R_{ii}.
2. Complete with zeros the row i and column $j = i$,
3. Complete the remaining diagonal entries with $\cos\mu$,
4. Complete the two remaining entries with $\sin\mu$
5. Put a minus in front of the $\sin\mu$ that is in the row directly below the 1—if the row containing the one is the bottom row, the minus sign goes in the top row.

Example 4 (Rotation about a single axis) *Consider the vector $\vec{u}$ coplanar with the basis vectors $\vec{a}_1$ and $\vec{a}_2$ as depicted in Figure 2.4. If we want to rotate $\vec{u}$ and angle ψ about the coordinate axis $\vec{a}_3$, then the coordinates of the resulting vector, $\vec{v}$, are related to the coordinates of $\vec{u}$ as follows:*

$$\begin{bmatrix} v_1^a \\ v_2^a \\ v_3^a \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1^a \\ u_2^a \\ u_3^a \end{bmatrix}. \tag{2.60}$$

Expanding this matrix multiplication gives the standard formulas of planar rotations or 2D-rotations.

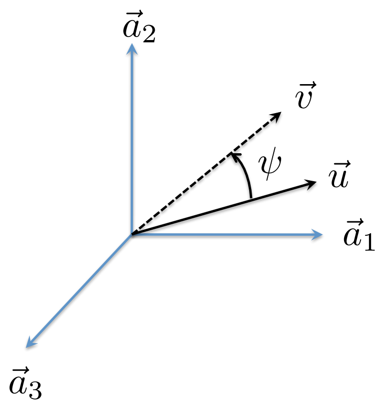


Figure 2.4: Rotation of a vector $\vec{u}$ about the z-coordinate axis.

□

2.9 Consecutive Rotations and Euler Angles

When we introduced the rotation matrix in (2.56), we used four parameters to define the matrix, namely, the coordinates of the unit vector that indicates the axis of rotation and the angle to be rotated. Since a unit vector has norm equal to one, such a vector can be specified in terms of two parameters only (3 coordinates and one constraint). This indicates that a rotation matrix could be expressed in terms of only three parameters.

The **attitude** (or **orientation**) of a basis $\{b\}$ relative to a basis $\{a\}$ can be described by the three consecutive rotations about the main axes, which take $\{a\}$ into $\{b\}$. These rotations can be performed in a different order (there are 12 different ways of doing this), and each triplet of rotated angles is called a set of **Euler angles**. The most commonly used Euler angles in robotics and vehicle dynamics are **yaw**, **pitch** and **roll**:

Yaw, Pitch and **Roll** are defined by the following three consecutive rotations (Z-Y-X sequence) that take $\{a\}$ into the orientation of $\{b\}$:

1. Rotation about the z axis of $\{a\}$ an angle ψ (yaw angle) resulting in $\{a'\}$;
2. Rotation about the y axis of $\{a'\}$ an angle θ (pitch angle) resulting in $\{a''\}$;
3. Rotation about the x axis of $\{a''\}$ an angle ϕ (roll angle) resulting in $\{b\}$.

This is illustrated in Figure 2.5.

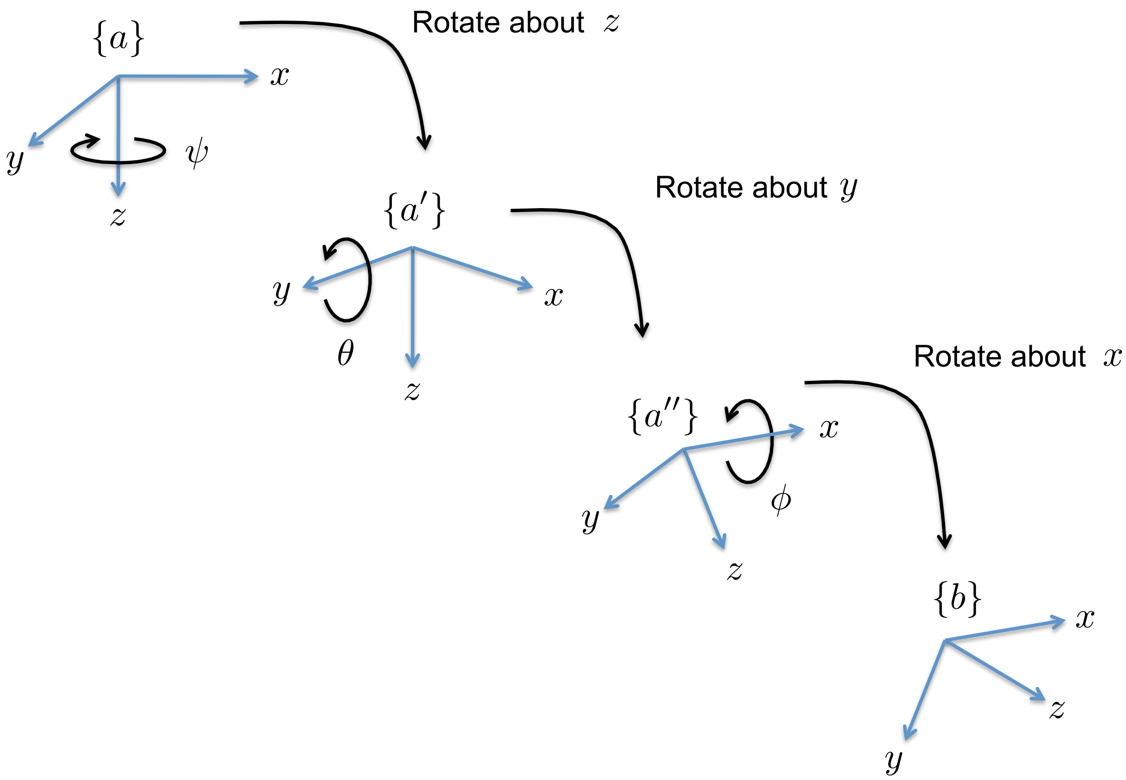


Figure 2.5: Roll, Pitch and yaw.

The positive angle convention corresponds to a right-handed screw advancing in the positive direction of the axis about which the rotation is performed.

The vector of Euler angles that takes $\{a\}$ into the orientation of $\{b\}$ will be denoted by

$$\boldsymbol{\Theta}_b^a \triangleq [\phi, \theta, \psi]^{\mathsf{T}}. \tag{2.61}$$

Note that (2.61) is not the representation of a physical vector magnitude in a basis, but a convenient short hand notation.

Several books on mechanics use the traditional Euler angles, which result from the consecutive rotations with the Z-Y-Z sequence. That is, first rotates about the z axis, then about the y axis, followed by another rotation about the z axis. In this course, we will use the Z-Y-X sequence, namely, yaw, pitch, and roll.

2.10 Rotation Matrix in Terms of Euler Angles

Using the consecutive single rotations (2.57)-(2.59) corresponding to roll pitch and yaw, the rotation matrix that take $\{a\}$ into $\{b\}$ can be expressed as

$$\mathbf{R}_b^a = \mathbf{R}_{z,\psi}\, \mathbf{R}_{y,\theta}\, \mathbf{R}_{x,\phi}. \tag{2.62}$$

Note that the order in this multiplication reflects the rotation sequence starting from the left matrix, and the reason for this order is that

$$\mathbf{u}^a = \mathbf{R}_b^a\, \mathbf{u}^b. \tag{2.63}$$

After multiplication, we obtain

$$\mathbf{R}_b^a = \begin{bmatrix} c\psi c\theta & -s\psi c\phi + c\psi s\theta s\phi & s\psi s\phi + c\psi c\phi s\theta \\ s\psi c\theta & c\psi c\phi + s\phi s\theta s\psi & -c\psi s\phi + s\psi c\phi s\theta \\ -s\theta & c\theta s\phi & c\theta c\phi \end{bmatrix}. \tag{2.64}$$

where $s \equiv \sin(\cdot)$ and $c \equiv \cos(\cdot)$.

2.11 Rotations vs Coordinate Transformations

Rotation matrices are related to the coordinate transformation matrices we introduced in Section 2.3.

If $\mathbf{R}_b^a$ is the matrix that rotates the vectors of the basis $\{a\}$ into the vectors of the basis $\{b\}$, then $\mathbf{R}_b^a$ is the coordinate tranformation matrix that relates the coordinates of a vector in $\{b\}$ to its components in $\{a\}$:

$$\mathbf{u}^a = \mathbf{R}_b^a\, \mathbf{u}^b. \tag{2.65}$$

Be aware that many texts on mechanics use coordinate transformation matrices $\mathbf{C}_b^a$ as in (2.14). The difference between the coordinate transformation matrix $\mathbf{C}_b^a$ in (2.14) and $\mathbf{R}_b^a$ in (2.65) is that the former is defined in terms of the Euler angles that take $\{b\}$ into the orientation of $\{a\}$, whereas the latter is defined in terms of the Euler angles that take $\{a\}$ into the orientation of $\{b\}$. It is easy to become confused because one matrix looks like the transpose of the other, and thus the inverse, but one should bear in mind that the Euler angle convention is different. **In this course, we will use only rotation matrices**, which is the standard convention in robotics.

Consider the case depicted in Figure 2.6. If we use rotation matrix that rotates $\vec{a}_1$ into $\vec{b}_1$, we can then relate the coordinates of the vector $\vec{u}$ in the two basis as follows:

$$\begin{bmatrix} u_1^a \\ u_2^a \\ u_3^a \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1^b \\ u_2^b \\ u_3^b. \end{bmatrix} \tag{2.66}$$

You can always check that you have used the correct transformation by looking at some simple cases. For example, if we take $\psi = \pi/2$, and $\mathbf{u}^a = [1, 0, 0]^\mathsf{T}$, then

$$\begin{bmatrix} 1^a \\ 0^a \\ 0^a \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0^b \\ -1^b \\ 0^b. \end{bmatrix}, \tag{2.67}$$

which is what we should expect.

Therefore, a rotation matrix can be used to obtain the coordinates of a rotated vector in a basis—see (2.55), and also to relate the coordinates of a vector in two different basis that are rotated from each other—see (2.65).

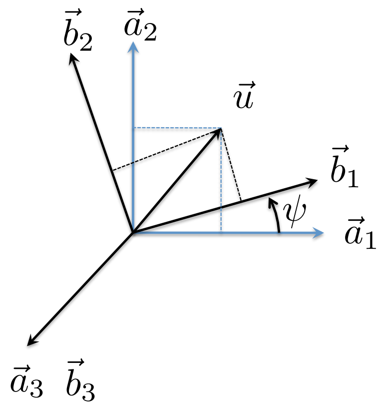


Figure 2.6: Vector coordinates in two basis.

2.12 Properties of Rotation Matrices

Rotation matrices are **orthogonal** (the collumns are orthogonal vectors),

$$\mathbf{R}\mathbf{R}^\mathsf{T} = \mathbf{I}_{3\times3}.$$

This means that the columns are orthogonal vectors in $\mathbb{R}^3$, and also that the inverse is just the transposed:

$$\mathbf{R}^{-1} = \mathbf{R}^\intercal.$$

In addition, $\det(\mathbf{R}) = 1$.

Matrices that satisfy the properties described above belong to the **special orthogonal group of order** 3, namely,

$$SO(3) : \{\mathbf{R} \in \mathbb{R}^{3\times3} | \mathbf{R}\mathbf{R}^\intercal = \mathbf{I}_{3\times3}, \det(\mathbf{R}) = 1\}.$$

2.13 Time-derivative of a Rotation Matrix

Since the rotation matrix $\mathbf{R}_b^a$ is orthogonal,

$$\mathbf{R}_b^a(\mathbf{R}_b^a)^\intercal = \mathbf{I}_{3\times3}, \tag{2.68}$$

then

$$\frac{d}{dt}[\mathbf{R}_b^a(\mathbf{R}_b^a)^\intercal] = \dot{\mathbf{R}}_b^a(\mathbf{R}_b^a)^\intercal + \mathbf{R}_b^a(\dot{\mathbf{R}}_b^a)^\intercal = \mathbf{0}. \tag{2.69}$$

This indicates that $\dot{\mathbf{R}}_b^a(\mathbf{R}_b^a)^\intercal$ is a skew-symmetric. Thus, $\dot{\mathbf{R}}_b^a(\mathbf{R}_b^a)^\intercal$ can be described by a column vector as in (2.7). such a vector is the angular velocity vector $\vec{\omega}_{b/a}$, and it satisfies the following relationship:

$$\mathbf{S}(\boldsymbol{\omega}_{b/a}^a) = \dot{\mathbf{R}}_b^a\,(\mathbf{R}_b^a)^\intercal. \tag{2.70}$$

Note also that

$$\mathbf{S}(\boldsymbol{\omega}_{b/a}^a) = \mathbf{R}_b^a\,\mathbf{S}(\boldsymbol{\omega}_{b/a}^b)(\mathbf{R}_b^a)^{-1}. \tag{2.71}$$

Equation (2.71) indicates that the cross-product operator $\mathbf{S}(\cdot)$ is a **second order tensor**: the components of a second order tensor in two different basis are transformed by pre and post multiplying by a rotation matrix. To demonstrate that (2.71) holds, let us start with the cross product $\vec{w} = \vec{\omega} \times \vec{r}$. This can be expressed in $\{a\}$ and $\{b\}$:

$$\mathbf{w}^a = \mathbf{S}(\boldsymbol{\omega}^a)\mathbf{r}^a, \tag{2.72}$$

$$\mathbf{w}^b = \mathbf{S}(\boldsymbol{\omega}^b)\mathbf{r}^b. \tag{2.73}$$

Then,

$$\mathbf{w}^a = \mathbf{R}_b^a\mathbf{w}^b \quad \Rightarrow \quad \mathbf{S}(\boldsymbol{\omega}^a)\mathbf{r}^a = \mathbf{R}_b^a\mathbf{S}(\boldsymbol{\omega}^b)\mathbf{r}^b, \tag{2.74}$$

$$= \mathbf{R}_b^a\mathbf{S}(\boldsymbol{\omega}^b)\mathbf{R}_a^b\mathbf{r}^a, \tag{2.75}$$

and hence

$$\mathbf{S}(\boldsymbol{\omega}^a) = \mathbf{R}_b^a\mathbf{S}(\boldsymbol{\omega}^b)\mathbf{R}_a^b. \tag{2.76}$$

Using the relation (2.71), we can express the **derivative of the rotation matrix** as

$$\dot{\mathbf{R}}_b^a = \mathbf{S}(\boldsymbol{\omega}_{b/a}^a)\,\mathbf{R}_b^a = \mathbf{R}_b^a\,\mathbf{S}(\boldsymbol{\omega}_{b/a}^b). \tag{2.77}$$

Also, since $\mathbf{R}_a^b = (\mathbf{R}_b^a)^T$, it follows that

$$\dot{\mathbf{R}}_a^b = -\mathbf{R}_a^b\,\mathbf{S}(\boldsymbol{\omega}_{b/a}^a) \ = -\mathbf{S}(\boldsymbol{\omega}_{b/a}^b)\,\mathbf{R}_a^b\,. \tag{2.78}$$

2.14 Angular Velocity and Euler-angle Derivatives

Let us consider the Euler angles that take $\{a\}$ into $\{d\}$ via a composite roll, pitch and yaw rotations

$$\{a\} \to \{b\} \to \{c\} \to \{d\}.$$

Then we have the following angular velocities:

$$\boldsymbol{\omega}_{b/a}^a = \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix}, \quad \boldsymbol{\omega}_{c/b}^b = \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix}, \quad \boldsymbol{\omega}_{d/c}^c = \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix}. \tag{2.79}$$

The addition of angular velocities (2.51) establishes that

$$\vec{\omega}_{d/a} = \vec{\omega}_{b/a} + \vec{\omega}_{c/b} + \vec{\omega}_{d/c}. \tag{2.80}$$

In order to perform this sum we must express all the angular velocities in the same basis; and therefore,

$$\boldsymbol{\omega}_{d/a}^a = \boldsymbol{\omega}_{b/a}^a + \mathbf{R}_b^a\,\boldsymbol{\omega}_{c/b}^b + \mathbf{R}_b^a\,\mathbf{R}_c^b\,\boldsymbol{\omega}_{d/c}^c. \tag{2.81}$$

Because the rotations are single rotations, $\mathbf{R}_b^a = \mathbf{R}_{z,\psi}$ and $\mathbf{R}_c^b = \mathbf{R}_{y,\theta}$, and then

$$\boldsymbol{\omega}_{d/a}^a = \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} + \mathbf{R}_{z,\psi} \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \mathbf{R}_{z,\psi}\,\mathbf{R}_{y,\theta} \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix}. \tag{2.82}$$

If we express the angular velocity in $\{d\}$,

$$\boldsymbol{\omega}_{d/a}^d = \mathbf{R}_{x,-\phi}\mathbf{R}_{y,-\theta} \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} + \mathbf{R}_{x,-\phi} \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix}. \tag{2.83}$$

After multiplication, we obtain

$$\boldsymbol{\omega}_{d/a}^a = \mathbf{E}_a(\boldsymbol{\Theta}_d^a)\dot{\boldsymbol{\Theta}}_d^a = \begin{bmatrix} c\psi c\theta & -s\psi & 0 \\ s\psi c\theta & c\psi & 0 \\ -s\theta & 0 & 1 \end{bmatrix} \dot{\boldsymbol{\Theta}}_d^a. \tag{2.84}$$

$$\boldsymbol{\omega}^d_{d/a} = \mathbf{E}_d(\boldsymbol{\Theta}^a_d)\dot{\boldsymbol{\Theta}}^a_d = \begin{bmatrix} 1 & 0 & -s\theta \\ 0 & c\phi & s\phi c\theta \\ 0 & -s\phi & c\phi c\theta \end{bmatrix} \dot{\boldsymbol{\Theta}}^a_d. \tag{2.85}$$

The inverse relationships are obtained via matrix inversion:

$$\dot{\boldsymbol{\Theta}}^a_d = \mathbf{E}_a^{-1}(\boldsymbol{\Theta}^a_d)\boldsymbol{\omega}^a_{d/a} = \begin{bmatrix} c\psi/c\theta & s\phi/c\theta & 0 \\ -s\psi & c\psi & 0 \\ c\psi t\theta & s\psi t\theta & 1 \end{bmatrix} \boldsymbol{\omega}^a_{d/a}. \tag{2.86}$$

$$\dot{\boldsymbol{\Theta}}^a_d = \mathbf{E}_d^{-1}(\boldsymbol{\Theta}^a_d)\boldsymbol{\omega}^d_{d/a} = \begin{bmatrix} 1 & s\phi t\theta & c\phi t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi/c\theta & c\phi/c\theta \end{bmatrix} \boldsymbol{\omega}^d_{d/a}. \tag{2.87}$$

Since neither $\boldsymbol{\Theta}^a_d$, nor $\dot{\boldsymbol{\Theta}}^a_d$ is a physical vector, the above transformations are not rotations: for example, the $\det(\mathbf{E}_d) = \cos\theta$, and that $\mathbf{E}_d^{-1} \neq \mathbf{E}_d^{\mathsf{T}}$.

The transformation $\mathbf{E}_d^{-1}$ has a singularity for $\theta = \pm\pi/2$—known as the Euler-angle singularity. This singularity is not a problem in many applications, but it can be a problem for aircraft and some underwater vehicles. To avoid this singularity, one can work directly with the matrix differential equation

$$\dot{\mathbf{R}}^d_a = \mathbf{S}(\boldsymbol{\omega}^d_{d/a})\,\mathbf{R}^d_a, \tag{2.88}$$

and then obtain the Euler angles from

$$\theta = -\sin^{-1}\mathbf{R}^d_a(1,3), \tag{2.89}$$

$$\phi = \operatorname{atan2}(\mathbf{R}^d_a(2,3), \mathbf{R}^d_a(3,3)), \tag{2.90}$$

$$\psi = \operatorname{atan2}(\mathbf{R}^d_a(1,2), \mathbf{R}^d_a(1,1)). \tag{2.91}$$

Equation (2.88) is known as the **Poisson Kinematic Equation** (Stevens and Lewis, 2003).

The Euler singularity can be avoided by using a redundant parameterisation of the rotation matrix. One of such parameterisations can be obtained in terms of **quaternions**, which use 4 parameters to describe the rotation matrix instead of the 3 Euler angles. This is used, for example, in aircraft and spacecraft simulation and control design. This description goes beyond the scope of these notes, and the reader can see, for example, Stevens and Lewis (2003) and references there in.

2.15 Position, Velocity, and Acceleration

Figure 2.7 shows a particle[1] that moves relative to a frame $\mathcal{A}$ to with we associate a basis $\{a\}$.

[1] A point with no volume but with a mass associated to it.

The **position** vector denoted by $\vec{r}_{P/A}$ indicates "*the position of the point P with respect to the point A.*"

The **rate of change of position** observed in $\mathcal{A}$ will be denoted by $\frac{{}^{\mathcal{A}}d\vec{r}_{P/A}}{dt}$, which refers to "*the rate of change of the position of P with respect to A observed in $\mathcal{A}$.*"

The **Law of Inertia** of Galilei and Newton states that a particle sufficiently removed from other particles and bodies will retain its state of motion: either rest or uniform rectilinear motion. A reference frame in which the Law of Inertia holds it is called a **Newtonian** or **inertial frame**.

If $\mathcal{A}$ is considered an inertial frame, then the rate of change of position of a point P with respect to a point A fixed in $\mathcal{A}$ is called **velocity**,

$$\vec{v}_{P/\mathcal{A}} \triangleq \frac{{}^{\mathcal{A}}d\vec{r}_{P/A}}{dt}.$$

Note that we can use the notation $\vec{v}_{P/\mathcal{A}}$ because the velocity will be the same for any reference point A we chose in $\mathcal{A}$. Note also that we reserve the term velocity only for rates of change of position observed in inertial frames.

The modulus or norm of a velocity is called a **speed**:

$$v_{P/\mathcal{A}} \triangleq \|\vec{v}_{P/\mathcal{A}}\|.$$

The **rate of change of velocity** of a point P with respect to a point A fixed in an inertial frame $\mathcal{A}$ is called **acceleration** and will be denoted by

$$\vec{a}_{P/\mathcal{A}} = \frac{{}^{\mathcal{A}}d\vec{v}_{P/\mathcal{A}}}{dt} = \frac{{}^{\mathcal{A}}d^2\vec{r}_{P/A}}{dt^2},$$

where again we can use the notation $\vec{a}_{P/\mathcal{A}}$ because the acceleration will be the same for any reference point A we chose $\mathcal{A}$. Note that we reserve the term acceleration only for rates of change of velocity in inertial frames.

2.16 Motion in Different Frames

Consider the case depicted in Figure 2.8, where the point A is fixed in $\mathcal{A}$ and the point B is fixed in $\mathcal{A}$. Assume also that the particle P is moving. Then a question of interest is how to relate the motion of P in the two frames $\mathcal{A}$ and $\mathcal{B}$. We start by establishing the relative positions:

$$\vec{r}_{P/A} = \vec{r}_{B/A} + \vec{r}_{P/B}. \tag{2.92}$$

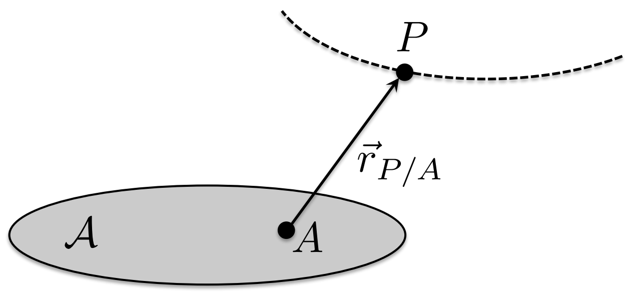


Figure 2.7: Position of a particle P with respect to a point A fixed in a frame $\mathcal{A}$.

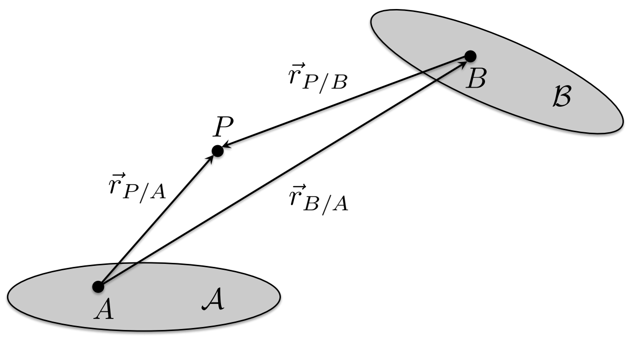


Figure 2.8: Relative positions of a particle and two reference frames.

Now we can take the time derivative in $\mathcal{A}$:

$$\frac{{}^{\mathcal{A}}d\vec{r}_{P/A}}{dt} = \frac{{}^{\mathcal{A}}d\vec{r}_{B/A}}{dt} + \frac{{}^{\mathcal{A}}d\vec{r}_{P/B}}{dt}. \tag{2.93}$$

If we further assume that $\mathcal{A}$ is inertial, then

$$\vec{v}_{P/\mathcal{A}} = \vec{v}_{B/\mathcal{A}} + \frac{{}^{\mathcal{A}}d\vec{r}_{P/B}}{dt}. \tag{2.94}$$

Note that in the last term we have made use of the fact that B is fixed in $\mathcal{B}$.

Now we can apply the Transport Theorem (2.41) to the last term of (2.94) to obtain the sought relation between velocities in the two frames:

$$\vec{v}_{P/\mathcal{A}} = \vec{v}_{B/\mathcal{A}} + \frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B}. \tag{2.95}$$

To obtain the accelerations, again we start from the positions (2.92) and take the time-derivatives in $\mathcal{A}$:

$$\frac{{}^{\mathcal{A}}d^2\vec{r}_{P/A}}{dt^2} = \frac{{}^{\mathcal{A}}d^2\vec{r}_{B/A}}{dt^2} + \frac{{}^{\mathcal{A}}d^2\vec{r}_{P/B}}{dt^2}, \tag{2.96}$$

which we can write as

$$\vec{a}_{P/\mathcal{A}} = \vec{a}_{B/\mathcal{A}} + \frac{{}^{\mathcal{A}}d}{dt}\left(\frac{{}^{\mathcal{A}}d\vec{r}_{P/B}}{dt}\right). \tag{2.97}$$

We can apply the Transport Theorem (2.21) to the third term on the right-hand side:

$$\vec{a}_{P/\mathcal{A}} = \vec{a}_{B/\mathcal{A}} + \frac{{}^{\mathcal{A}}d}{dt}\left(\frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B}\right). \tag{2.98}$$

We can further apply the Transport Theorem to the term in backets in (2.98):

$$\vec{a}_{P/\mathcal{A}} = \vec{a}_{B/\mathcal{A}} + \frac{{}^{\mathcal{B}}d}{dt}\left(\frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B}\right) + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \left(\frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B}\right). \tag{2.99}$$

Let is define the **angular acceleration**:

$$\vec{\alpha}_{\mathcal{B}/\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\vec{\omega}_{\mathcal{B}/\mathcal{A}} = \frac{{}^{\mathcal{B}}d}{dt}\vec{\omega}_{\mathcal{B}/\mathcal{A}},$$

where the last equality follows from the Transport Theorem. Distributing (2.99), we obtain,

$$\vec{a}_{P/\mathcal{A}} = \vec{a}_{B/\mathcal{A}} + \frac{{}^{\mathcal{B}}d^2\vec{r}_{P/B}}{dt^2} + \vec{\alpha}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B} + 2\vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} + \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B}. \tag{2.100}$$

Expression (2.100) gives the acceleration of a particle P in a frame a using observations of the motion in a frame b which rotates and translated with respect to a. Some of the terms in (2.100) have names:

$$\begin{aligned} &\vec{\alpha}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B} && \text{(Transverse or Euler acceleration)} \\ &2\vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \frac{{}^{\mathcal{B}}d\vec{r}_{P/B}}{dt} && \text{(Coriolis acceleration)} \\ &\vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{\omega}_{\mathcal{B}/\mathcal{A}} \times \vec{r}_{P/B} && \text{(Centripetal acceleration)} \end{aligned}$$

It is very easy to make mistakes when trying to apply the final formulae (2.95) and (2.100) directly to different problems. It is safer instead to consider first the relative positions, and then apply the transport theorem once to get the velocities and twice to get the accelerations. Let us see an application example.

Example 5 (Planar Robotic Manipulator) *Consider the 2-DOF planar robotic manipulator shown in Figure 2.9. Let us obtain the position, velocity, and acceleration of the end effector at the point B. We consider three reference frames: base $\mathcal{B}_0$ and two rigid-body links $\mathcal{B}_1$ and $\mathcal{B}_2$. Associated with each reference frame we consider the points of interest O, A, and B. To each frame we associate a basis $\{0\}$, $\{1\}$, and $\{2\}$ respectively such that the rotation is about the z-direction. The lengths of the links are given by*

$$\begin{aligned} l_1 &= \|\vec{r}_{A/O}\|, \\ l_2 &= \|\vec{r}_{B/A}\|. \end{aligned}$$

<u>Position</u>
The position of the effector relative to the base is given by

$$\vec{r}_{B/O} = \vec{r}_{A/O} + \vec{r}_{B/A}, \tag{2.101}$$

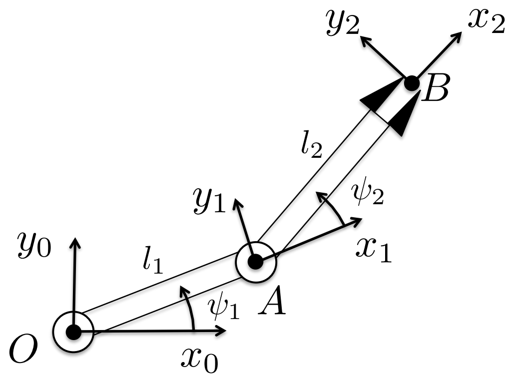


Figure 2.9: Two-degree-of-freedom planar robotic manipulator.

which can be expressed as

$$\mathbf{r}^0_{B/O} = \mathbf{R}^0_1\,\mathbf{r}^1_{A/O} + \mathbf{R}^0_1\,\mathbf{R}^1_2\,\mathbf{r}^2_{B/A}, \tag{2.102}$$

where

$$\mathbf{r}^1_{A/O} = \begin{bmatrix} l_1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{r}^2_{B/A} = \begin{bmatrix} l_2 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{R}^0_1 = \begin{bmatrix} c\psi_1 & -s\psi_1 & 0 \\ s\psi_1 & c\psi_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{R}^1_2 = \begin{bmatrix} c\psi_2 & -s\psi_2 & 0 \\ s\psi_2 & c\psi_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.103}$$

Velocity
The velocity of the effector relative to the base is given by

$$\vec{v}_{B/O} \triangleq {}^0\dot{\vec{r}}_{B/O} = {}^0\dot{\vec{r}}_{A/O} + {}^0\dot{\vec{r}}_{B/A}, \tag{2.104}$$

where we have used the notation ${}^0\dot{\vec{r}} \equiv {}^0d\vec{r}/dt$. *Application of the transport theorem leads to*

$$\begin{aligned}\vec{v}_{B/O} &= {}^0\dot{\vec{r}}_{A/O} + {}^0\dot{\vec{r}}_{B/A}, \\ &= \vec{\omega}_{1/0} \times \vec{r}_{A/O} + {}^1\dot{\vec{r}}_{B/A} + \vec{\omega}_{1/0} \times \vec{r}_{B/A}, \quad = \vec{\omega}_{1/0} \times \vec{r}_{A/O} + \vec{\omega}_{2/1} \times \vec{r}_{B/A} + \vec{\omega}_{1/0} \times \vec{r}_{B/A},\end{aligned}$$

and hence

$$\vec{v}_{B/O} = \vec{\omega}_{1/0} \times \vec{r}_{A/O} + (\vec{\omega}_{2/1} + \vec{\omega}_{1/0}) \times \vec{r}_{B/A}. \tag{2.105}$$

The latter can be expressed as

$$\mathbf{v}^0_{B/O} = \mathbf{S}(\boldsymbol{\omega}^0_{1/0})\,\mathbf{R}^0_1\,\mathbf{r}^1_{A/O} + \mathbf{S}(\boldsymbol{\omega}^0_{2/1} + \boldsymbol{\omega}^0_{1/0})\,\mathbf{R}^0_1\,\mathbf{R}^1_2\,\mathbf{r}^2_{B/A}, \tag{2.106}$$

where

$$\boldsymbol{\omega}^0_{1/0} = \begin{bmatrix} 0 \\ 0 \\ \dot{\psi}_1 \end{bmatrix}, \quad \boldsymbol{\omega}^0_{2/1} = \begin{bmatrix} 0 \\ 0 \\ \dot{\psi}_2 \end{bmatrix}. \tag{2.107}$$

Acceleration
The velocity of the effector relative to the base is given by

$$\vec{a}_{B/O} \triangleq {}^0\dot{\vec{v}}_{B/O}. \tag{2.108}$$

Application of the transport theorem to (2.105) *leads to*

$$\begin{aligned}\vec{a}_{B/O} &= {}^0\dot{\vec{\omega}}_{1/0} \times \vec{r}_{A/O} + \vec{\omega}_{1/0} \times {}^0\dot{\vec{r}}_{A/O} + ({}^0\dot{\vec{\omega}}_{2/1} + {}^0\dot{\vec{\omega}}_{1/0}) \times \vec{r}_{B/A} + (\vec{\omega}_{2/1} + \vec{\omega}_{1/0}) \times {}^0\dot{\vec{r}}_{B/A} \\ &= \vec{\alpha}_{1/0} \times \vec{r}_{A/O} + \vec{\omega}_{1/0} \times \vec{\omega}_{1/0} \times \vec{r}_{A/O} + (\vec{\alpha}_{2/1} + \vec{\alpha}_{1/0}) \times \vec{r}_{B/A} \\ &\quad + (\vec{\omega}_{2/1} + \vec{\omega}_{1/0}) \times (\vec{\omega}_{2/1} + \vec{\omega}_{1/0}) \times \vec{r}_{B/A},\end{aligned} \tag{2.109}$$

where $\vec{\alpha}_{1/0} \triangleq {}^0\dot{\vec{\omega}}_{1/0}$ *and* $\vec{\alpha}_{2/1} \triangleq {}^0\dot{\vec{\omega}}_{2/1}$. *It is left as an exercise to express* (2.109) *in* $\{0\}$.

□

2.17 Navigation and Vehicle Kinematics

For general vehicle navigation, we consider three reference frames: stars, Earth and the vehicle. Associated to these frames we consider different coordinate systems as illustrated in Figure 2.10. These coordinate systems are specified as follows:

- **Earth-Centred Inertial (ECI)** $\{i\}$: The vector $\vec{i}_1$ points towards the centre of mass of the sun in the direction of the line that passes through the centre of the Earth and the centre of the sun when the sun crosses the Equator at the vernal equinox (first day of the spring in the northern hemisphere). The Earth axis of rotation is used to define $\vec{i}_3$ and $\vec{i}_2$ completes the right-hand basis. This coordinate system is used when the point of reference considered is at the centre of mass of the Earth. The reference frame is conformed by the centre of mass of the Earth together with the stars to which the basis vectors point. This reference frame is considered inertial.

- **Earth-Centred Earth-Fixed (ECEF)** $\{e\}$: The vector $\vec{e}_1$ points towards the intersection of the Equator and the Greenwich meridian. The Earth axis of rotation is used to define $\vec{e}_3$ and $\vec{e}_2$ completes the right-hand basis. This coordinate system, associated with the Earth frame, is used when the point of reference considered is at the centre of mass of the Earth.

- **Local geographical North-East-Down (NED)** $\{n\}$: the shape of the Earth can be modelled using a ellipsoid of revolution (spheroid) centred at the Earth's centre of mass and rotating about an axis through the north and south poles. The parameters of this spheroid have been documented by the US department of defence World Geodesic System in 1984 (WGS-84). The NED is based on reference WGS-84 spheroid. The vectors $\vec{n}_1$ and $\vec{n}_2$ are in a tangential plane to the spheroid, and $\vec{n}_1$ points towards the geometric North, $\vec{n}_2$ points towards East, and $\vec{n}_3$ points downwards completing the right-hand basis.

- **Body-fixed** $\{b\}$: The body-fixed coordinate system is associated with a vehicle: aircraft, cart, marine vessel, etc. The basis vectors $\vec{b}_i$ have a prescribed directions. For aircraft, $\vec{b}_1$ points forward, $\vec{b}_2$ points towards the right wing, and $\vec{b}_3$ points down. For marine vehicles $\vec{b}_1$ points forward, $\vec{b}_2$ points towards the starboard (right-hand side when facing towards the bow), and $\vec{b}_3$ points down. For land vehicles, $\vec{b}_1$ points forward, $\vec{b}_2$ points to the left, and $\vec{b}_3$ points up.

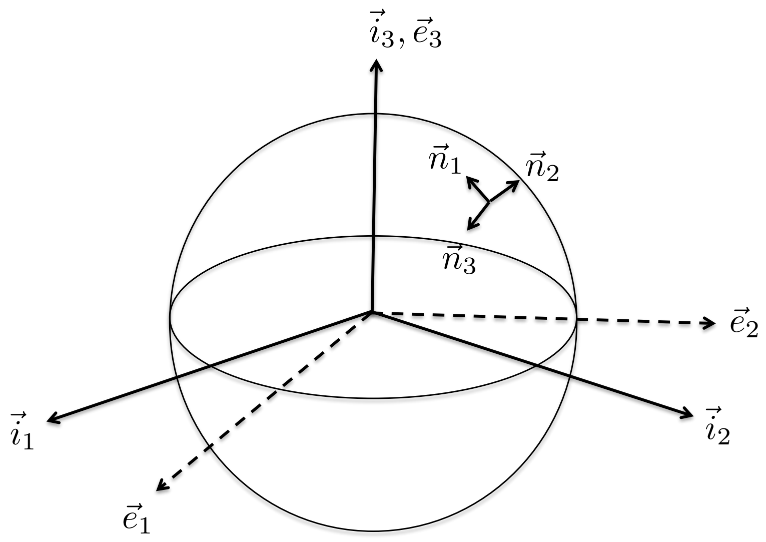


Figure 2.10: Coordinate systems used in navigation: $\{i\}$-ECI, $\{e\}$-ECEF, and $\{n\}$-NED.

For many applications of mobile robots the earth is consider an inertial frame as we discuss in the following example.

Example 6 (Accelerations observed on Earth) *Consider the inertial frame $\mathcal{I}$ made of the centre of mass of the Earth and the stars with a point of reference O at the centre of mass of the Earth and the associated ECI $\{i\}$ coordinate system. Consider also the Earth as a frame $\mathcal{N}$ with reference point N on the surface of the Earth and the NED $\{n\}$ coordinate system.*

Consider the motion of a point P. We can use (2.100) to obtain its acceleration:

$$\vec{a}_{P/\mathcal{I}} = \vec{a}_{N/\mathcal{I}} + \frac{{}^{\mathcal{N}}d^2\vec{r}_{P/N}}{dt^2} + 2\vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \frac{{}^{\mathcal{N}}d\vec{r}_{P/N}}{dt} + \vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \vec{r}_{P/N}. \tag{2.110}$$

Note that because the angular velocity $\vec{\omega}_{\mathcal{N}/\mathcal{I}}$ is constant, then $\vec{\alpha}_{\mathcal{N}/\mathcal{I}} = \vec{0}$; hence, we have omitted the transverse acceleration.

An observer standing on the surface of the Earth close to N would perceive

$$\frac{{}^{\mathcal{N}}d^2\vec{r}_{P/N}}{dt^2} = \vec{a}_{P/\mathcal{I}} - \vec{a}_{N/\mathcal{I}} - 2\vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \frac{{}^{\mathcal{N}}d\vec{r}_{P/N}}{dt} - \vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \vec{\omega}_{\mathcal{N}/\mathcal{I}} \times \vec{r}_{P/N}. \tag{2.111}$$

The Coriolis term is what makes, for example, the trajectory of the a rocket curve after its launch, and also the wind spiral towards the low pressure areas in the atmosphere. The negative of the centripetal term is called the centrifugal acceleration.

For most of the experiments that we do in the laboratory, we operate at low speed and in proximity to our local geographical frame $\mathcal{N}$. This together with the fact that the Earth spin rate is small ($\|\vec{\omega}_{\mathcal{N}/\mathcal{I}}\| = 7.2921150 \times 10^{-5}$ rad/s), result in negligible Coriolis and centripetal accelerations due to the rotation of the Earth and also in a negligible $\vec{a}_{N/\mathcal{I}}$. Hence, for lab experiments, we consider that $\mathcal{N}$ is approximately inertial:

$$\vec{a}_{P/\mathcal{N}} \triangleq \frac{{}^{\mathcal{N}}d^2\vec{r}_{P/N}}{dt^2} \approx \vec{a}_{P/\mathcal{I}}. \tag{2.112}$$

This assumption is also made for vehicle navigation in local geographical areas.

□

GPS receivers provide coordinates in terms of latitude, longitude and height relative to the WGS-84 datum model. Figure 2.11 shows the geodetic latitude μ, geodetic height h, and longitude l relative to the WGS-84 ellipsoid of revolution. The ECEF coordinates of a point are given by

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} = \begin{bmatrix} (N+h)\cos\mu\,\cos l \\ (N+h)\cos\mu\,\sin l \\ (\frac{r_p^2}{r_e^2}N+h)\sin\mu \end{bmatrix}, \tag{2.113}$$

where $r_p = 6356752\text{m}$ and $r_e = 6378137\text{m}$ are the polar and equatorial semi-axes of the WGS-84 ellipsoid, and

$$N = \frac{r_e^2}{\sqrt{r_e^2\,\cos^2\mu + r_p^2\sin^2\mu}}.$$

The rotation matrix that relates coordinates in ECEF and NED is given by

$$\mathbf{R}_n^e(l,\mu) = \mathbf{R}_{z,l}\,\mathbf{R}_{y,-\mu-\pi/2} = \begin{bmatrix} -\cos l\,\sin\mu & -\sin l & -\cos l\cos\mu \\ -\sin l\,\sin\mu & \cos l & -\sin l\,\cos\mu \\ 0 & 0 & -\sin\mu \end{bmatrix}. \tag{2.114}$$

In applications of mobile robot positioning relative to a reference point O, we can find the coordinates of a point B on the robot from measurements of l, μ, h as follows:

$$\mathbf{r}_{B/O}^n = \mathbf{R}_n^e(l_O,\mu_O)^{\mathsf{T}}(\mathbf{r}_B^e - \mathbf{r}_O^e).$$

For further details about GPS and applications to vehicle navigation see, for example, Hofmann-Wellenhof et al. (2003), Farrel and Barth (1999), and Rogers (2007).

Figure 2.12 shows the two coordinate systems and reference points used in vehicle motion description. The position of the vehicle is given by the relative position of B with respect to N. The components of this vector are North, East and Down positions and are denoted

$$\mathbf{r}_{B/N}^n \triangleq [N,\, E,\, D]^T.$$

The orientation of the of the vehicle is given by the Euler angles that take $\{n\}$ into $\{b\}$. The *Euler-angle vector* is given by $\mathbf{\Theta}_b^n \triangleq [\phi,\, \theta,\, \psi]^T$. The **position-orientation vector** is defined by

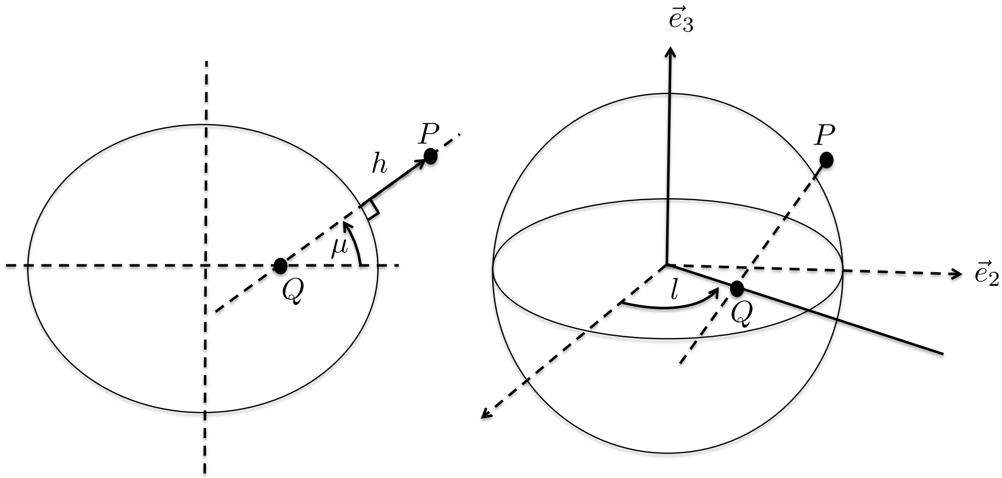


Figure 2.11: Geodetic latitude μ, geodetic height h, and longitude l relative to the WGS-84 ellipsoid of revolution.

$$\boldsymbol{\eta} \triangleq \begin{bmatrix} \mathbf{r}^n_{B/N} \\ \boldsymbol{\Theta}^n_b \end{bmatrix} = [N,\, E,\, D,\, \phi,\, \theta,\, \psi]^\mathsf{T}. \tag{2.115}$$

The velocities are expressed in terms of body-fixed coordinates and denoted by the **body-fixed velocity vector** (linear-angular),

$$\boldsymbol{\nu} \triangleq \begin{bmatrix} \mathbf{v}^b_{B/\mathcal{N}} \\ \boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} \end{bmatrix} = [u,\, v,\, w,\, p,\, q,\, r]^\mathsf{T}. \tag{2.116}$$

The linear-velocity vector $\mathbf{v}^b_{B/\mathcal{N}} = {}^{\mathcal{N}}\dot{\mathbf{r}}^b_{B/N} = [u,\, v,\, w]^T$ is the time derivative of the position vector as seen from the frame $\mathcal{N}$, and the components correspond to expressing the vector in $\{b\}$. These components are the **surge**, **sway**, and **heave velocities** respectively. The vector $\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} = [p,\, q,\, r]^T$ is the angular velocity of the body with respect to Earth with components corresponding to expressing the vector in $\{b\}$. These components are the **roll**, **pitch**, and **yaw rates** respectively. Table 2.1 summarises the notation.

The **trajectory** of a vehicle is given by the time evolution of the position-orientation vector $\boldsymbol{\eta}$ defined in (2.115),

$$\boldsymbol{\eta}(t) = \boldsymbol{\eta}(t_0) + \int_{t_0}^{t} \dot{\boldsymbol{\eta}}(t')\, dt',$$

Note that the time integral of $\boldsymbol{\nu}$ has no physical meaning.

The time-derivative of the positions is related to the body-fixed velocities via a kinematic transformation,

$$\dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\eta})\,\boldsymbol{\nu}, \qquad \mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \mathbf{R}^n_b(\boldsymbol{\Theta}) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}(\boldsymbol{\Theta}) \end{bmatrix} \tag{2.117}$$

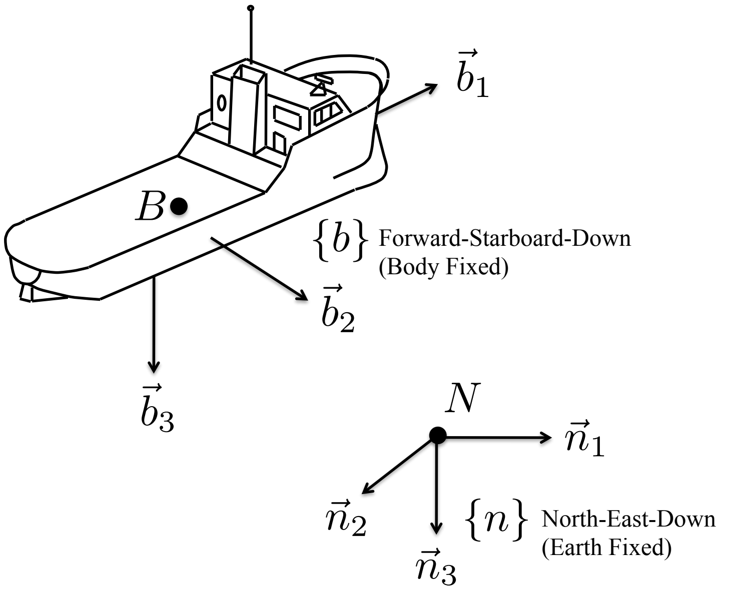


Figure 2.12: Coordinate systems used for vehicle motion description.

Table 2.1: Summary of vehicle motion variables.

Variable	**Name**	**Frame**	**Units**
N	North position	Earth-fixed	m
E	East position	Earth-fixed	m
D	Down position	Earth-fixed	m
ϕ	Roll angle	-	rad
θ	Pitch angle	-	rad
ψ	Yaw angle	-	rad
u	Surge speed	Body-fixed	m/s
v	Sway speed	Body-fixed	m/s
w	Heave speed	Body-fixed	m/s
p	Roll rate	Body-fixed	rad/s
q	Pitch rate	Body-fixed	rad/s
r	Yaw rate	Body-fixed	rad/s
$\mathbf{r}^n_{B/N} = [N,\ E,\ D]^\mathsf{T}$	Position vector	Earth-fixed	
$\mathbf{v}^b_{\mathcal{B}/\mathcal{N}} = {}^{\mathcal{N}}\dot{\mathbf{r}}^b_{B/\mathcal{N}} = [u,\ v,\ w]^\mathsf{T}$	Linear-velocity vector	Body-fixed	
$\boldsymbol{\Theta} = [\phi, \theta, \psi]^\mathsf{T}$	Euler-angle vector	-	
$\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} = [p, q, r]^\mathsf{T}$	Angular-velocity vector	Body-fixed	
$\boldsymbol{\eta} = [(\mathbf{r}^n_{B/N})^\mathsf{T}, \boldsymbol{\Theta}^T]^\mathsf{T}$	Position-orientation vector	-	
$\boldsymbol{\nu} = [(\mathbf{v}^b_{B/\mathcal{N}})^\mathsf{T}, (\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})^\mathsf{T}]^\mathsf{T}$	Body-fixed velocity vector	Body-fixed	

where the rotation matrix $\mathbf{R}^n_b(\boldsymbol{\Theta})$ is given by (2.64) and the transformation from the angular velocity expressed in the body-fixed coordinates to the time derivatives of the Euler angles given in (2.87), that is

$$\mathbf{R}^n_b(\boldsymbol{\Theta}) = \begin{bmatrix} c_\psi c_\theta & -s_\psi c_\phi + c_\psi s_\theta s_\phi & s_\psi s_\phi + c_\psi c_\phi s_\theta \\ s_\psi c_\theta & c_\psi c_\phi + s_\phi s_\theta s_\psi & -c_\psi s_\phi + s_\psi c_\phi s_\theta \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{bmatrix},$$

and

$$\mathbf{T}(\mathbf{\Theta}) = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & \frac{s_\phi}{c_\theta} & \frac{c_\phi}{c_\theta} \end{bmatrix}, \quad t_\theta \equiv \tan(\theta), \qquad \cos(\theta) \neq 0.$$

Chapter 3

Newton-Euler Equations of Motion

"To myself, I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me."

Sir Isaac Newton (1642-1727)
English Physicist and Mathematician

This chapter presents the rigid-body equations of motion in terms of Newton-Euler theory. These equations relate the forces acting on a body to its motion. Together with the kinematic transformation presented in the previous chapter, the equations of motion provide the basis for vehicle dynamic models and some simple robotic manipulator models—the topic of multi-body system dynamics, however, is often better dealt with using variational approaches in terms of D'Alembert and Lagrangian formulations, which are discussed in the next chapter.

3.1 Vectors in Mechanics

When we describe magnitudes of mechanical systems, we can distinguish two types of vector magnitudes (Kane and Levinson, 1985):

- Free vectors,
- Sliding vectors,

A free-vector magnitude has no specified line of action. Examples of these magnitudes are angular velocity and torque. A sliding-vector magnitude, on the other hand, has line of action. An example of this type of magnitude is a force acting on a body.

The **moment of a sliding vector** $\vec{u}$ about a point O is defined as

$$\vec{M}_{\vec{u}/O} = \vec{r}_{P/O} \times \vec{u}, \tag{3.1}$$

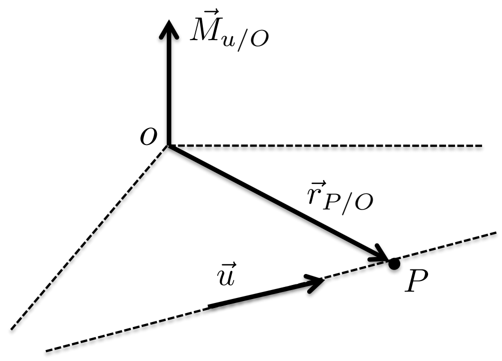


Figure 3.1: Moment of a vector $\vec{u}$ about a point O.

where P is any point on the line of action of $\vec{u}$ as illustrated in Figure 3.1.

Two important moments used in the mechanics are

- Moment of a force,
- Moment of momentum of a particle; known as angular momentum.

A contact force $\vec{F}$ acting on on a rigid body will have a line of action. The **moment of the force about a point** O is

$$\vec{M}_{\vec{F}/O} = \vec{r}_{P/O} \times \vec{F}. \tag{3.2}$$

The **angular momentum of a particle about a point** O is defined as the moment of momentum $\vec{p} = m\vec{v}$ about the point:

$$\vec{L}_O \triangleq \vec{M}_{\vec{p}/O} = \vec{r}_{P/O} \times \vec{p}, \tag{3.3}$$

where $\vec{r}_{P/O}$ can be taken as position of the particle relative to the point O.

Two forces that have a null resultant but do not have the same line of action are called a **couple**—see Figure 3.2. The resultant moment of a couple is the same about *any* point. That is, if $\vec{F}_1$ and $\vec{F}_2$ are a couple, then

$$\vec{M}_{\vec{F}1/P} + \vec{M}_{\vec{F}2/P} = \vec{M}_{\vec{F}1/Q} + \vec{M}_{\vec{F}2/Q}, \tag{3.4}$$

for any pair of points P and Q.

The resultant moment of a couple is called a **torque**, and it will be denoted $\vec{T}(t)$.

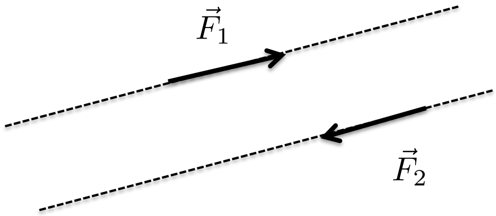


Figure 3.2: Couple.

3.2 Equivalent Sets of Forces

Rigid bodies are often acted upon by sets of forces. In order to implement mathematical models, it convenient to identify equivalent sets of forces.

A set S of forces acting on a rigid body will have a resultant force and moment about a point P,

$$\vec{F}_S = \sum_j \vec{F}_j, \tag{3.5}$$

$$\vec{M}_{S/P} = \sum_j \vec{r}_{j/P} \times \vec{F}_j, \tag{3.6}$$

where $\vec{r}_{j/P}$ is a vector from P to any point on the line of action of $\vec{F}_j$.

The moment of the set of forces S about another point Q is

$$\begin{aligned} \vec{M}_{S/Q} &= \sum_j \vec{r}_{j/Q} \times \vec{F}_j, & (3.7)\\ &= \sum_j (\vec{r}_{P/Q} + \vec{r}_{j/P}) \times \vec{F}_j, & (3.8)\\ &= \vec{r}_{P/Q} \times \vec{F}_S + \sum_j \vec{r}_{j/P} \times \vec{F}_j. & (3.9) \end{aligned}$$

This result is known as the **Moment Transport Theorem** (Rao, 2006):

$$\vec{M}_{S/Q} = \vec{M}_{S/P} + \vec{r}_{P/Q} \times \vec{F}_S. \tag{3.10}$$

The resultant force $\vec{F}_S$ of a system of forces S does not have a line of action. Hence, the second term in (3.10) is a bit strange, for it makes sense only if $\vec{F}_S$ has a line of action through P. Because of this, it is convenient to find an equivalent representation of S in terms of well defined vector forces with line of action.

Two sets of forces S and Σ are equivalent if they have the same resultant and moment about any point.

If a set of forces S has a resultant $\vec{F}_S$ and a moment $\vec{M}_{S/B}$, then we can define the following equivalent set Σ:

- A force $\vec{F}_\Sigma = \vec{F}_S$ but with line of action through B,
- A torque $\vec{T}_\Sigma = \vec{M}_{S/B}$.

Note that due to the moment transport theorem

$$\vec{M}_{S/Q} = \vec{T}_\Sigma + \vec{r}_{B/Q} \times \vec{F}_\Sigma, \tag{3.11}$$

which is well defined because $\vec{F}_\Sigma$ has line of action.

This is illustrated in Figure 3.3.

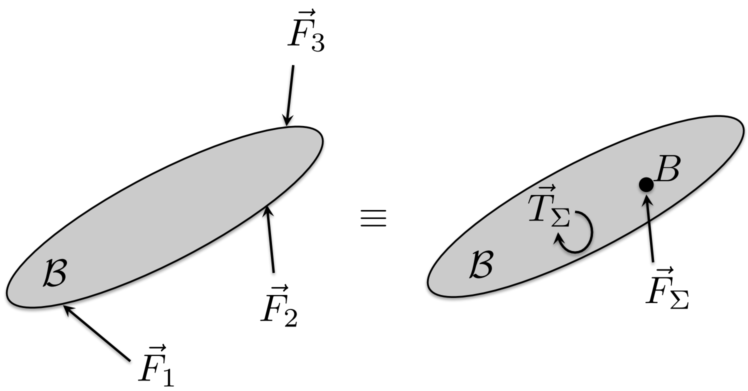


Figure 3.3: Equivalent systems of forces.

3.3 Centre of Mass and Forces on a Rigid Body

The location of the **centre of mass** of a rigid body $\mathcal{B}$ with mass m is given by

$$\vec{r}_C \triangleq \frac{1}{m}\int_{\mathcal{B}} \vec{r}_m \, dm, \tag{3.12}$$

where $\vec{r}_m$ is the position of the mass element dm with respect to a chosen reference frame. From (3.12), it follows that

$$\int_{\mathcal{B}} \vec{r}_{m/C} \, dm = \int_{\mathcal{B}} (\vec{r}_m - \vec{r}_C) \, dm = 0. \tag{3.13}$$

In order to compute the location of the centre of mass, it is convenient to express (3.12) relative to a reference point B in the body:

$$\mathbf{r}^b_{C/B} = \frac{1}{m}\int_{\mathcal{B}} \mathbf{r}^b_{m/B} \, dm. \tag{3.14}$$

The resultant of gravity forces on a rigid body is called **weight**, and it is given by

$$\vec{W} = \int_{\mathcal{B}} \vec{g}\, dm, \tag{3.15}$$

where $\vec{g}$ is the acceleration of gravity. The moment the weight produces about a point B is

$$\vec{M}_{\vec{W}/B} = \int_{\mathcal{B}} \vec{r}_{m/B} \times \vec{g}\, dm = \int_{\mathcal{B}} \vec{r}_{m/B}\, dm \times \vec{g} = \int_{\mathcal{B}} \vec{r}_{C/B} \times \vec{g}\, dm = \vec{r}_{C/B} \times \vec{W}. \tag{3.16}$$

Note that if B coincides with the centre of mass C, then $\vec{M}_{\vec{W}/B}$ vanishes.

Expression (3.16) indicates that the resultant of gravity forces (weight) acting on a body produces a moment as if the weight had line of action through the centre of mass. For this reason, the centre of mass $\vec{r}_C$ is also called the **centre of gravity**; however, we do not need gravity to define $\vec{r}_C$.

The following are equivalent sets of forces on a rigid body (Kane and Levinson, 1985):

1. A set S with resultant force $\vec{F}_S$ without line of action and moment $\vec{M}_{S/B}$ about a point B.
2. A force $\vec{F}_C = \vec{F}_S$ with line of action through the centre of mass and torque $\vec{T}_C = \vec{M}_{S/C}$.
3. A force $\vec{F}_B = \vec{F}_C$ with line of action through the the point B and torque $\vec{T}_B = \vec{M}_{S/B}$, where from (3.11), it follows that

$$\vec{T}_B = \vec{T}_C + \vec{r}_{C/B} \times \vec{F}_C. \tag{3.17}$$

The representations 2 and 3 are commonly used in robotics and multi-body dynamic simulation software. Note that (3.11) can be used to find moments about particular points given the above force representations.

3.4 Angular Momentum and Inertia Tensor

The angular momentum of a rigid body $\mathcal{B}$ about its centre of mass is defined as

$$\vec{L}_C \triangleq \int_{\mathcal{B}} \vec{r}_{m/C} \times \vec{v}_m\, dm, \tag{3.18}$$

where $\vec{r}_{m/C}$ is position the mass element dm relative to the centre of mass, and $\vec{v}_m$ is the velocity of the mass element dm with respect to an inertial frame.

Let us consider a point of reference N in the inertial frame $\mathcal{N}$. The position of dm can be expressed as

$$\vec{r}_{m/N} = \vec{r}_{C/N} + \vec{r}_{m/C}. \tag{3.19}$$

Using the rate of transport theorem,

$$\vec{v}_m = \vec{v}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{m/C}. \tag{3.20}$$

The angular momentum (3.18) can be then be expressed as

$$\begin{aligned}
\vec{L}_C &= \int_{\mathcal{B}} \vec{r}_{m/C} \times (\vec{v}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{m/C})\, dm, \\
&= \int_{\mathcal{B}} \vec{r}_{m/C}\, dm \times \vec{v}_C - \int_{\mathcal{B}} \vec{r}_{m/C} \times \vec{r}_{m/C} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}\, dm, \\
&= -\int_{\mathcal{B}} \vec{r}_{m/C} \times \vec{r}_{m/C} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}\, dm,
\end{aligned} \tag{3.21}$$

where in the second step, we have made use of (3.13).

Expression (3.21) transforms the vector $\vec{\omega}_{\mathcal{B}/\mathcal{N}}$ into the vector $\vec{L}_C$. This transformation can be expressed as

$$\vec{L}_C = \vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}} \qquad \equiv \qquad \vec{L}_C = -\int_{\mathcal{B}} \vec{r}_{m/C} \times \vec{r}_{m/C} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}\, dm, \tag{3.22}$$

where $\vec{I}_C$ is the **inertia tensor about the centre of gravity**.

Note that "·" in (3.22) is not a scalar product, but a tensor product. Think of the inertia tensor as an operator that takes the angular velocity and transform it into the angular momentum. In this course, we will not elaborate on Tensor Calculus, but use the tensor notation simply as a short hand notation that allows us to be independent of coordinate systems. The student should bear in mind that there is much more to it. As we will see in the next section, when a an inertia tensor is expressed in a coordinate system, it becomes a 3 by 3 matrix called the inertia matrix, and the tensor product is simply a matrix multiplication.

The angular momentum can be taken about and arbitrary point B fixed to the body:

$$\vec{L}_B \triangleq \int_{\mathcal{B}} \vec{r}_{m/B} \times \vec{v}_m\, dm. \tag{3.23}$$

Using the kinematic relationships

$$\vec{r}_{m/N} = \vec{r}_{B/N} + \vec{r}_{m/B}, \tag{3.24}$$

$$\vec{v}_m = \vec{v}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{m/B}, \tag{3.25}$$

we can express (3.23), as

$$\vec{L}_B = \int_{\mathcal{B}} \vec{r}_{m/B} \times (\vec{v}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{m/B})\, dm, \tag{3.26}$$

$$= \int_{\mathcal{B}} \vec{r}_{m/B}\, dm \times \vec{v}_B - \int_{\mathcal{B}} \vec{r}_{m/B} \times \vec{r}_{m/B} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}\, dm, \tag{3.27}$$

$$= \vec{r}_{C/B} \times m\, \vec{v}_B - \int_{\mathcal{B}} \vec{r}_{m/B} \times \vec{r}_{m/B} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}\, dm. \tag{3.28}$$

The second term in (3.28) can be expressed in terms of the inertia tensor about B; then,

$$\vec{L}_B = \vec{r}_{C/B} \times m\,\vec{v}_B + \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{3.29}$$

The relationship between $\vec{L}_B$ and $\vec{L}_C$ can

$$\vec{L}_B = \int_{\mathcal{B}} \vec{r}_{m/B} \times \vec{v}_m \, dm, \tag{3.30}$$

$$= \int_{\mathcal{B}} (\vec{r}_{m/C} + \vec{r}_{C/B}) \times \vec{v}_m \, dm, \tag{3.31}$$

$$= \vec{L}_C + \int_{\mathcal{B}} \vec{r}_{C/B} \times \vec{v}_m \, dm, \tag{3.32}$$

from which it follows, **the transport theorem for angular momentum**:

$$\vec{L}_B = \vec{L}_C + \vec{r}_{C/B} \times m\,\vec{v}_C. \tag{3.33}$$

3.5 Inertia Matrix

Let us consider a coordinate system $\{b\}$ fixed to the body. From (3.22), it follows that

$$\mathbf{L}_C^b = \mathbf{I}_C^b\,\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b, \qquad \mathbf{I}_C^b = -\int_{\mathcal{B}} \mathbf{S}(\mathbf{r}_{m/C}^b)\mathbf{S}(\mathbf{r}_{m/C}^b)\, dm, \tag{3.34}$$

Alternatively, (3.34) can be expressed as

$$\mathbf{I}_C^b = \int_{\mathcal{B}} (\mathbf{r}_{m/C}^b \cdot \mathbf{r}_{m/C}^b)\mathbf{I}_{3\times 3} - \mathbf{r}_{m/C}^b (\mathbf{r}_{m/C}^b)^\mathsf{T}\, dm, \tag{3.35}$$

where $\mathbf{I}_{3\times 3}$ is the 3 by 3 identity matrix. If the components of $\mathbf{r}_{m/C}^b$ are denoted x, y, z, then

$$\mathbf{I}_C^b = \int_{\mathcal{B}} \begin{bmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{bmatrix} dm. \tag{3.36}$$

Note that while the inertia tensor depends only on the point about which the angular momentum is considered, the inertia matrix depends not only on this point, but also on the basis $\{b\}$ chosen.

The angular momentum in matrix form can then be expressed as

$$\mathbf{L}_C^b = \mathbf{I}_C^b\,\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b. \tag{3.37}$$

If we express the above in the coordinate system $\{n\}$ fixed to the inertial frame $\mathcal{N}$, we obtain

$$\mathbf{R}_b^n \mathbf{L}_C^b = \mathbf{R}_b^n \mathbf{I}_C^b\, \mathbf{R}_n^b \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n, \tag{3.38}$$

$$\mathbf{L}_C^n = \mathbf{I}_C^n\,\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n, \tag{3.39}$$

where

$$\mathbf{L}_C^n = \mathbf{I}_C^n\,\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n, \qquad \mathbf{I}_C^n = \mathbf{R}_b^n \mathbf{I}_C^b\,\mathbf{R}_n^b. \tag{3.40}$$

That is, if we express the angular momentum of the body in a coordinate system $\{n\}$ and the body rotates with respect to $\{b\}$, then the inertia matrix is, in general, time varying.

Let us consider now what happens if we translate the point of reference to from C to a point B in the body is

$$\begin{aligned}
\mathbf{I}_B^b &= -\int_{\mathcal{B}} \mathbf{S}(\mathbf{r}_{m/B}^b)\mathbf{S}(\mathbf{r}_{m/B}^b)\,dm,\\
&= -\int_{\mathcal{B}} \mathbf{S}(\mathbf{r}_{m/C}^b + \mathbf{r}_{C/B}^b)\mathbf{S}(\mathbf{r}_{m/C}^b + \mathbf{r}_{C/B}^b)\,dm,\\
&= -\int_{\mathcal{B}} \mathbf{S}(\mathbf{r}_{m/C}^b)\mathbf{S}(\mathbf{r}_{m/C}^b)\,dm\\
&- \mathbf{S}(\mathbf{r}_{C/B}^b)\int_{\mathcal{B}} \mathbf{r}_{m/C}^b\,dm - \mathbf{S}\left(\int_{\mathcal{B}} \mathbf{r}_{m/C}^b\,dm\right)\mathbf{r}_{C/B}^b\\
&- \mathbf{S}(\mathbf{r}_{C/B}^b)\mathbf{S}(\mathbf{r}_{C/B}^b)\int_{\mathcal{B}} dm.
\end{aligned}$$

From (3.13), it follows that

$$\int_{\mathcal{B}} \mathbf{r}_{m/C}^b\,dm = 0, \tag{3.41}$$

which leads to a result known as the **Parallel Axis Theorem**:

$$\mathbf{I}_B^b = \mathbf{I}_C^b - m\mathbf{S}^2(\mathbf{r}_{C/B}^b) = \mathbf{I}_C^b - m[(\mathbf{r}_{C/B}^b)^T\mathbf{r}_{C/B}^b\mathbf{I}_{3\times 3} + \mathbf{r}_{C/B}^b(\mathbf{r}_{C/B}^b)^{\mathsf{T}}]. \tag{3.42}$$

Note that because we have made use of (3.41), (3.42) must only be applied between the point B and the centre of gravity C.

So if we need to translate the reference point from B' to B, we can use (3.42) in a sequence: we first translate B' to C and then translate C to B.

The inertia matrix is always symmetric and since it is also real, the **Spectral Theorem** of linear algebra establishes that is can be diagonalised using an orthogonal transformation (Strang, 1988).

For a real-symmetric matrix, the eigenvalues λ_i are non-repeated and the corresponding eigenvectors $\mathbf{V}_i$ are orthogonal. Therefore, if we form the matrix $\mathbf{R} = [\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3]$, Then, we can obtain a diagonal inertia matrix:

$$\mathbf{I}_B^{b'} = \mathbf{R}^T\mathbf{I}_B^b\mathbf{R} = \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix}. \tag{3.43}$$

Given an inertia matrix $\mathbf{I}_O^b$, its eigenvectors $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ can be used to form a rotation matrix $\mathbf{R}_{b'}^b = [\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3]$. In the new basis $\{b'\}$, the inertia matrix is diagonal,

$$\mathbf{I}_B^{b'} = \mathbf{R}_b^{b'} \, \mathbf{I}_B^b \, \mathbf{R}_{b'}^b = \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} \tag{3.44}$$

and the diagonal elements called principal moments of inertia are the eigenvalues of $\mathbf{I}_B^b$.

This can sometimes be used to simplify the equations of motion. However, we may end up with an inconvenient rotated basis $\{b'\}$.

A different approach is to use the parallel axis theorem (3.42) and find $\mathbf{r}_{C/B}^b = [x_C, y_C, z_C]^\mathsf{T}$ such that $\mathbf{I}_B^b$ is diagonal. Namely, solve the equations

$$mI_{yz}^C \, x_C^2 = -I_{xy}^C \, I_{xz}^C, \tag{3.45}$$
$$mI_{xz}^C \, y_C^2 = -I_{xy}^C \, I_{yz}^C, \tag{3.46}$$
$$mI_{xy}^C \, z_C^2 = -I_{xz}^C \, I_{yz}^C. \tag{3.47}$$

3.6 Newton-Euler Equations of Motion

In his book *Philosophiæ Naturalis Principia Mathematica*, first published in 1687, Newton formulated his three laws of motion:

- Law I: Every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed.

- Law II: The change of momentum of a body is proportional to the impulse (time-integral of the force) impressed on the body, and happens along the straight line on which that impulse is impressed.

- Law III: To every action there is always an equal and opposite reaction: or the forces of two bodies on each other are always equal and are directed in opposite directions.

Although the laws refer to bodies, they apply to particles, for Newton was not considering rotation. Euler (1707-1783) contributed to the extension of these laws to the case of rigid bodies. It was Euler also who first used the angular velocity vector as an auxiliary kinematic variable to express the change in angular momentum. Because a rigid body experiences both translation and rotation, the conservation of momentum needs to be considered for both characteristics of motion (Rao, 2006):

Euler's 1st Law: In an inertial frame $\mathcal{N}$, the resultant of a set of forces S applied to a rigid body equals the product of the mass times the acceleration of the centre of mass:

$$\vec{F} = m\,\vec{a}_C \quad \equiv \quad \mathbf{F}^n = m\,\mathbf{a}_C^n. \tag{3.48}$$

Euler's 2nd Law: In an inertial frame $\mathcal{N}$, the resultant moment of a set of forces S applied to the rigid body of mass m equals the rate of change of the angular momentum about the centre of mass:

$$\vec{M}_C = \frac{{}^{\mathcal{N}}d}{dt}\,\vec{L}_C \quad \equiv \quad \mathbf{M}_C^n = {}^{\mathcal{N}}\dot{\mathbf{L}}_C^n, \tag{3.49}$$

where we have simplified the notation for the resultant moment $\vec{M}_C \equiv \vec{M}_{S/C}$.

To apply the above laws of motion to the analysis of motion of vehicles, it is convenient to express them in terms of body-fixed coordinates. There are three reasons for doing this:

- First, the inertia matrix is constant in body-fixed coordinates, which simplifies the analysis and simulation of the equations of motion and the design of control laws.
- Second, the forces imparted by the actuators located on the vehicle are more conveniently described in body-fixed coordinates.
- Third, measurements taken onboard using inertial navigation systems provide quantities in body-fixed coordinates.

Equations of Motion about the Centre of Mass

Let us consider an inertial frame $\mathcal{N}$ with an associated coordinate system $\{n\}$, a coordinate system $\{b\}$ fixed to the body frame $\mathcal{B}$ and use the centre of mass as a point of reference in the body. By applying the transport theorem to $\vec{v}_C$, (3.48) by can be written as

$$\vec{F} = m[{}^{\mathcal{B}}\dot{\vec{v}}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_C]. \tag{3.50}$$

For the rotational motion, we can consider the transport theorem applied to the angular velocity:

$$\begin{aligned}
\vec{M}_C &= \frac{{}^{\mathcal{N}}d}{dt}\,\vec{L}_C = \frac{{}^{\mathcal{N}}d}{dt}(\vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}), && (3.51)\\
&= \frac{{}^{\mathcal{B}}d}{dt}(\vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}) + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times (\vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}), && (3.52)\\
&= \vec{I}_C \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times (\vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}), && (3.53)
\end{aligned}$$

where we have use the fact that $\frac{{}^{\mathcal{B}}d}{dt}\vec{I}_C = \vec{0}$.

Therefore, the **rigid body equations of motion bout the centre of mass** can be expressed as

$$m[{}^{\mathcal{B}}\dot{\vec{v}}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_C] = \vec{F}, \tag{3.54}$$
$$\vec{I}_C \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times (\vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}) = \vec{T}_C, \tag{3.55}$$

where $\vec{F}$ has line of action through the centre of mass, and $\vec{T}_C$ is a torque equal to the resultant moment about the centre of mass.

Equation (3.55) is called Euler equation, and hence (3.54)-(3.55) are referred to as the **Newton-Euler Equations of motion about the centre of mass**.

Expressing all variables in (3.54)-(3.55) in $\{b\}$, the above equations become

$$m[{}^{\mathcal{B}}\dot{\mathbf{v}}_C^b + \mathbf{S}(\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b)\mathbf{v}_C^b] = \mathbf{F}^b, \tag{3.56}$$
$$\mathbf{I}_C^b\, {}^{\mathcal{B}}\dot{\boldsymbol{\omega}}_{\mathcal{B}/\mathcal{N}}^b + \mathbf{S}(\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b)\, \mathbf{I}_C^b \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b = \mathbf{T}_C^b. \tag{3.57}$$

or alternatively

$$\begin{bmatrix} m\mathbf{I}_{3\times 3} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_C^b \end{bmatrix} \begin{bmatrix} {}^{\mathcal{B}}\dot{\mathbf{v}}_C^b \\ {}^{\mathcal{B}}\dot{\boldsymbol{\omega}}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix} + \begin{bmatrix} m\mathbf{S}(\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b) & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}(\mathbf{I}_C^b \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b) \end{bmatrix} \begin{bmatrix} \mathbf{v}_C^b \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix} = \begin{bmatrix} \mathbf{F}^b \\ \mathbf{T}_C^b \end{bmatrix}. \tag{3.58}$$

Equations of Motion about a Point

The equations can also be expressed at a point B in the body different from the centre of mass. This is advantageous in many vehicle and robotic applications.

Let express the position of the centre of mass as follows

$$\vec{r}_{C/N} = \vec{r}_{B/N} + \vec{r}_{C/B}, \tag{3.59}$$

where N is a point of reference in the inertial frame $\mathcal{N}$. By taking time derivatives in the inertial frame leads to

$$\vec{v}_C = \vec{v}_B + \frac{{}^{\mathcal{N}}d}{dt}\vec{r}_{C/B}, \tag{3.60}$$
$$= \vec{v}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B}. \tag{3.61}$$

Taking a second derivative in the inertial frame

$${}^{\mathcal{N}}\dot{\vec{v}}_C = {}^{\mathcal{N}}\dot{\vec{v}}_B + \frac{{}^{\mathcal{N}}d}{dt}(\vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B}), \tag{3.62}$$
$$= {}^{\mathcal{B}}\dot{\vec{v}}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_B + {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B}. \tag{3.63}$$

We can now substitute this in (3.48) to obtain the translation equation of motion:

$$\vec{F} = m[{}^{\mathcal{B}}\dot{\vec{v}}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_B + {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B}]. \tag{3.64}$$

For the rotation, we can start by taking the time derivative of (3.33) in the inertial frame, namelly,

$$
\begin{aligned}
{}^{\mathcal{N}}\dot{\vec{L}}_B &= {}^{\mathcal{N}}\dot{\vec{L}}_C + \frac{{}^{\mathcal{N}}d}{dt}\vec{r}_{C/B} \times m\,\vec{v}_C + \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_C, &(3.65)\\
&= {}^{\mathcal{N}}\dot{\vec{L}}_C + (\vec{v}_C - \vec{v}_B) \times m\,\vec{v}_C + \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_C, &(3.66)\\
&= {}^{\mathcal{N}}\dot{\vec{L}}_C - \vec{v}_B \times m\,\vec{v}_C + \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_C. &(3.67)
\end{aligned}
$$

On the other hand, the time derivative of (3.29) leads to

$${}^{\mathcal{N}}\dot{\vec{L}}_B = \vec{v}_C \times m\vec{v}_B + \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_B + \vec{I}_B \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times (\vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}) \tag{3.68}$$

Combining the last two expressions,

$${}^{\mathcal{N}}\dot{\vec{L}}_C = \vec{r}_{C/B} \times m({}^{\mathcal{N}}\dot{\vec{v}}_B - {}^{\mathcal{N}}\dot{\vec{v}}_C) + \vec{I}_B \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{3.69}$$

Using (3.69) in Euler's 2nd law (3.49), we obtain

$$\vec{M}_C = \vec{r}_{C/B} \times m({}^{\mathcal{N}}\dot{\vec{v}}_B - {}^{\mathcal{N}}\dot{\vec{v}}_C) + \vec{I}_B \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{3.70}$$

Using (3.10) and (3.48), it follows that

$$\vec{M}_C = \vec{M}_B - \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_C. \tag{3.71}$$

Substituting this in (3.70) gives

$$\vec{M}_B = \vec{r}_{C/B} \times m\,{}^{\mathcal{N}}\dot{\vec{v}}_B + \vec{I}_{BO} \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{3.72}$$

Application of the transport theorem to the first term on the right-hand side of (3.72) leads to the **Euler Equation about the about** B:

$$\vec{M}_B = \vec{r}_{C/B} \times m\,{}^{\mathcal{B}}\dot{\vec{v}}_B + \vec{r}_{C/B} \times m\,(\vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_{BO}) + \vec{I}_B \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{3.73}$$

Doing a replacement with an equivalent set of forces, the **rigid-body equations of motion about a point** B are

$$
\begin{aligned}
m[{}^{\mathcal{B}}\dot{\vec{v}}_B + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_B + {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{C/B}] = \vec{F}, &\quad(3.74)\\
\vec{r}_{C/B} \times m\,{}^{\mathcal{B}}\dot{\vec{v}}_B + \vec{r}_{C/B} \times m\,(\vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{v}_B) + \vec{I}_B \cdot {}^{\mathcal{B}}\dot{\vec{\omega}}_{\mathcal{B}/\mathcal{N}} + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}} = \vec{T}_B. &\quad(3.75)
\end{aligned}
$$

Expressing all variables in (3.74)-(3.75) in $\{b\}$, leads to the following equations in matrix form:

$$\begin{bmatrix} m\mathbf{I}_{3\times3} & -m\mathbf{S}(\mathbf{r}^b_{C/B}) \\ m\mathbf{S}(\mathbf{r}^b_{C/B}) & \mathbf{I}^b_B \end{bmatrix} \begin{bmatrix} {}^{\mathcal{B}}\dot{\mathbf{v}}^b_B \\ {}^{\mathcal{B}}\dot{\boldsymbol{\omega}}^b_{\mathcal{B}/\mathcal{N}} \end{bmatrix} + \begin{bmatrix} m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})[\mathbf{v}^b_B + \mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{r}^b_{C/B}] \\ \mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\,\mathbf{I}^b_B\,\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} + m\mathbf{S}(\mathbf{r}^b_{C/B})\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{v}^b_B \end{bmatrix} = \begin{bmatrix} \mathbf{F}^b \\ \mathbf{T}^b_B \end{bmatrix}. \tag{3.76}$$

The second term on the left-hand side of (3.76) are known as D'Alembert forces.

3.7 Vehicle Dynamics

In Section 2.17, we discussed the variables commonly used for vehicle motion description and their kinematic transformation. If we combine these results with the rigid-body equations of motion, we obtain a the basis of a vehicle dynamic model:

$$\dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\eta})\,\boldsymbol{\nu}, \tag{3.77}$$
$$\mathbf{M}_{RB}\dot{\boldsymbol{\nu}} + \mathbf{C}_{RB}(\boldsymbol{\nu})\boldsymbol{\nu} = \boldsymbol{\tau}. \tag{3.78}$$

where the variables are

- $\boldsymbol{\eta} = [(\mathbf{r}^n_{B/N})^\mathsf{T}, (\boldsymbol{\Theta}^n_b)^\mathsf{T}]^\mathsf{T}$—defined in (2.115)—is the vehicle-position orientation vector. The vector $\mathbf{r}^n_{B/N} = [N, E, D]^\mathsf{T}$ gives the components in the NED coordinate system of the position of the point of reference O of the vehicle relative to the Earth frame. The vector $\boldsymbol{\Theta}^n_b = [\phi, \theta, \psi]^\mathsf{T}$ is the vector of Euler angles that take the NED coordinate system into the body-fixed coordinate system.

- $\boldsymbol{\nu} = [({}^{\mathcal{N}}\dot{\mathbf{r}}^b_{B/N})^\mathsf{T}, (\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})^\mathsf{T}]^\mathsf{T}$—defined in (2.116)—the linear and angular velocity vector expressed in the body-fixed coordinate system. This vector is usually called the body-fixed velocity vector. The vector ${}^{\mathcal{N}}\dot{\mathbf{r}}^b_{B/N} \equiv \mathbf{v}^b_B = [u, v, w]^\mathsf{T}$ gives the components of the velocity of the the point B expressed in the body-fixed coordinate system. The vector $\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} = [p, q, r]^\mathsf{T}$ gives the components of the angular velocity of the vehicle relative to the inertial frame expressed in the body-fixed coordinate system.

- $\boldsymbol{\tau} = [(\mathbf{F}^b)^\mathsf{T}, (\mathbf{T}^b_B)]^\mathsf{T}$ is the vector of forces and torques. The vector $\mathbf{F}^b$ gives the components in body-fixed coordinates of a force with line of action through B that is equal to the resultant of external forces. The vector $\mathbf{T}^b_B$ is a torque that is equal to the resultant moment of the external forces about B.

The kinematic transformation $\mathbf{J}(\boldsymbol{\eta})$ is defined in (2.117). The **mass matrix** $\mathbf{M}_{RB}$ is given by

$$\mathbf{M}_{RB} \triangleq \begin{bmatrix} m\mathbf{I}_{3\times3} & -m\mathbf{S}(\mathbf{r}^b_{C/B}) \\ m\mathbf{S}(\mathbf{r}^b_{C/B}) & \mathbf{I}^b_B \end{bmatrix}, \tag{3.79}$$

which follows from (3.76). The mass matrix is always symmetric and positive definite:

$$\mathbf{M}_{RB} = \mathbf{M}_{RB}^T > \mathbf{0}.$$

The **Coriollis-Centripetal matrix** $\mathbf{C}_{RB}(\boldsymbol{\nu})$ follow from a factorisation of the second term on the left-hand side of (3.76). There are different ways of doing this factorisation, and one of these factorisation gives

$$\begin{bmatrix} m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})[\mathbf{v}^b_B + \mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{r}^b_{C/B}] \\ \mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\,\mathbf{I}^b_B\,\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} + m\mathbf{S}(\mathbf{r}^b_{C/B})\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{v}^b_B \end{bmatrix} = \begin{bmatrix} m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) & -m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{S}(\mathbf{r}^b_{C/B}) \\ m\mathbf{S}(\mathbf{r}^b_{C/B})\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) & -\mathbf{S}(\mathbf{I}^b_B\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) \end{bmatrix} \begin{bmatrix} \mathbf{v}^b_B \\ \boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}} \end{bmatrix},$$

from which it follows that

$$\mathbf{C}_{RB}(\boldsymbol{\nu}) = \begin{bmatrix} m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) & -m\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}})\mathbf{S}(\mathbf{r}^b_{C/B}) \\ m\mathbf{S}(\mathbf{r}^b_{C/B})\mathbf{S}(\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) & -\mathbf{S}(\mathbf{I}^b_B\boldsymbol{\omega}^b_{\mathcal{B}/\mathcal{N}}) \end{bmatrix}. \tag{3.80}$$

The Coriollis and centripetal matrix is skew-symmetric, that is

$$\mathbf{C}_{RB}(\boldsymbol{\nu}) = -\mathbf{C}_{RB}(\boldsymbol{\nu})^{\mathsf{T}}$$

If we denote $\mathbf{r}^b_{C/B} = [x_C, y_C, z_C]^{\mathsf{T}}$ and expand the model (3.77)-(3.78) component-wise, the equations of motion become:

$$\begin{aligned}
m\left[\dot{u} - vr + wq - x_C(q^2 + r^2) + y_C(pq - \dot{r}) + z_C(pr + \dot{q})\right] &= \tau_1 \\
m\left[\dot{v} - wp + ur - y_C(r^2 + p^2) + z_C(qr - \dot{p}) + x_C(qp + \dot{r})\right] &= \tau_2 \\
m\left[\dot{w} - uq + vp - z_C(p^2 + q^2) + x_C(rp - \dot{q}) + y_C(rq + \dot{p})\right] &= \tau_3 \\
I_x\dot{p} + (I_z - I_y)qr - (\dot{r} + pq)I_{xz} + (r^2 - q^2)I_{yz} + (pr - \dot{q})I_{xy} & \\
+m\left[y_C(\dot{w} - uq + vp) - z_C(\dot{v} - wp + ur)\right] &= \tau_4 \\
I_y\dot{q} + (I_x - I_z)rp - (\dot{p} + qr)I_{xy} + (p^2 - r^2)I_{zx} + (qp - \dot{r})I_{yz} & \\
+m\left[z_C(\dot{u} - vr + wq) - x_C(\dot{w} - uq + vp)\right] &= \tau_5 \\
I_z\dot{r} + (I_y - I_x)pq - (\dot{q} + rp)I_{yz} + (q^2 - p^2)I_{xy} + (rq - \dot{p})I_{zx} & \\
+m\left[x_C(\dot{v} - wp + ur) - y_C(\dot{u} - vr + wq)\right] &= \tau_6
\end{aligned} \tag{3.81}$$

A simple way to implement the model (3.77)-(3.78) in Matlab or in Matlab/Simulnk is to re-write the model in state-space form and write a Matlab function that takes as input $\boldsymbol{\eta}, \boldsymbol{\nu}, \boldsymbol{\tau}$ and computes $\dot{\boldsymbol{\eta}}, \dot{\boldsymbol{\nu}}$.

$$\begin{bmatrix} \dot{\boldsymbol{\eta}} \\ \dot{\boldsymbol{\nu}} \end{bmatrix} = \mathbf{f}(\boldsymbol{\eta}, \boldsymbol{\nu}, \boldsymbol{\tau}), \qquad \equiv \qquad \begin{bmatrix} \dot{\boldsymbol{\eta}} \\ \dot{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{J}(\boldsymbol{\eta}) \\ \mathbf{0} & -\mathbf{M}_{RB}^{-1}\mathbf{C}_{RB}(\boldsymbol{\nu}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{M}_{RB}^{-1} \end{bmatrix} \boldsymbol{\tau} \tag{3.82}$$

3.8 Examples of Vehicle Dynamics

The rigid body vehicle model (3.77)-(3.78) is the basis for models of

- Aircraft,
- Satellites,

- Land vehicles,
- Surface and underwater marine vehicles,

Different disciplines use various conventions for the Earth and the body coordinate system and also the point of reference in the body. In order to implement a full model of a vehicle, one needs to consider the nature of the external forces $\boldsymbol{\tau}$, which may include interactions with the environment, disturbances, and control.

Example 7 (Aircraft Dynamics) *For aircraft conducting local missions, the motion is often described relative to the NED coordinate system. The equations of motion (3.77)-(3.78) usually referred to the centre of mass ($\mathbf{r}^b_{C/B} = \mathbf{0}$), and the body-fixed coordinate system $\{b\}$ has positive directions forward, right wing and down.*

The motion variables used are $\boldsymbol{\eta}$ and $\boldsymbol{\nu}$ as defined in Table 2.1, the forces-and torques $\boldsymbol{\tau}$ are related to gravity, propulsion, and aerodynamic effects:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_G + \boldsymbol{\tau}_P + \boldsymbol{\tau}_A. \tag{3.83}$$

Zipfel (2007), for example, gives the following general model:

$$\boldsymbol{\tau}_G = \begin{bmatrix} -mg\sin\theta \\ mg\sin\phi\cos\theta \\ mg\cos\phi\cos\theta \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\tau}_P = \begin{bmatrix} X_T \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\tau}_A = \begin{bmatrix} \bar{q}SC_X(M, V_T, \alpha, \beta, \delta_e, q) \\ \bar{q}SC_Y(M, V_T, \alpha, \beta, \delta_r, \delta_a, p, r) \\ \bar{q}SC_Z(M, V_T, \alpha, \beta, \delta_e, q) \\ \bar{q}SbC_L(M, V_T, \alpha, \beta, \delta_a, p, r), \\ \bar{q}S\bar{c}C_M(M, V_T, \alpha, \beta, \delta_e, q), \\ \bar{q}SbC_N(M, V_T, \alpha, \beta, \delta_a, \delta_r, p, r), \end{bmatrix},$$

where $\bar{q} = \frac{\rho V_T^2}{2}$ is the dynamic pressure, V_T is the airspeed relative to the aircraft, M is the Mach number, α is the angle of attack and β the sideslip angle relative to the air velocity vector. X_T is the thrust produced by the propulsion system, S is the wing characteristic area, b is the wing span, $\bar{c}$ is the wing mean cord. The coefficients C_X, C_Y, C_Z are the longitudinal, transverse and normal force coefficients and C_L, C_M, C_N are the roll, pitch, and yaw coefficients. The normalised control surface angles are δ_a-aileron, δ_e-elevator, and δ_r-rudder, for which a positive deflection produces a negative moment about the unit vectors of $\{b\}$. Figure 3.4 shows the main control surfaces of the aircraft. In nominal operational conditions, the elevator is used to produce a pitch moment, the ailerons produce roll moment, and the rudder produces lateral forces and yaw moment. Note that the aircraft turns by a combination of roll and pitch angles and the rudder is used to initiate the turning manoeuvre, and produce lateral forces for landing in cross wind conditions.

The aerodynamic coefficients are complex nonlinear functions. Some flight simulators use look-up tables to implement these functions based on data recorded in wind tunnels—a non-parametric representation of the forces. Alternatively, parametric models based on series expansions can be obtained as an approximation. Larger aircraft models should also incorporate gyroscopic torques in the propulsion forces due to the rotors of the engines. For further details about aircraft dynamics see Stevens and Lewis (2003), and Zipfel (2007).

□

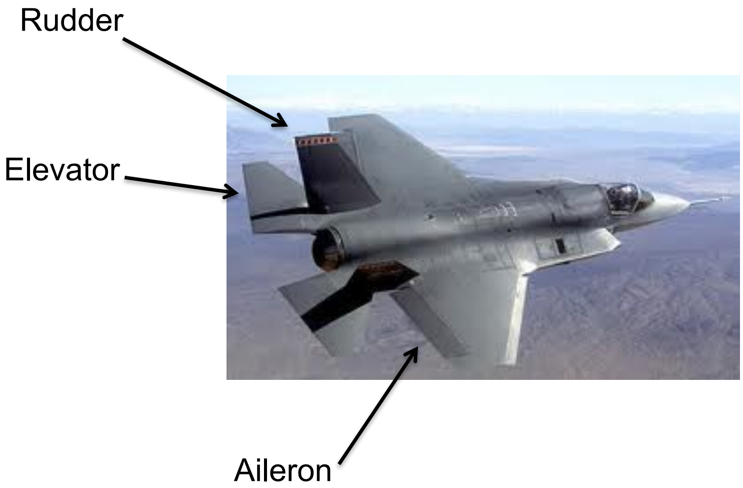


Figure 3.4: Aircraft main control surfaces.

Example 8 (Marine surface vessel manoeuvring model in 4 degrees of freedom) *For marine surface vessels, it is common to study manoeuvrability using a 4-degree-of-freedom model. The model in this example corresponds to a high-speed trimaran ferry. The model is adapted from Perez et al. (2007). Figure 3.5 shows a picture of the vessel.*

The model considered is given (3.77)-(3.78) *with the coordinate systems shown in Figure 2.12. Only the degrees of freedom of surge, sway, roll and yaw, are considered that is,*

$$\boldsymbol{\eta} = [x, y, \phi, \psi]^{\mathsf{T}}$$

$$\boldsymbol{\nu} = [u, v, p, r]^{\mathsf{T}}$$

$$\boldsymbol{\tau} = [X, Y, K, N]^{\mathsf{T}}.$$

The motion and force variables associated with the the other degrees of freedom are considered to be zero.

The forces have three components, hydrodynamic, propulsion, and control:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_H + \boldsymbol{\tau}_P + \boldsymbol{\tau}_C.$$

The rigid-body mass and Coriollis-centripetal matrices are given by

$$\mathbf{M}_{RB} = \begin{bmatrix} m & 0 & 0 & -my_C \\ 0 & m & -mz_C & mx_C \\ 0 & -mz_C & I_x & -I_{xz} \\ -my_C & mx_C & -I_{zx} & I_z \end{bmatrix},$$

and

$$\mathbf{C}_{RB}(\nu) = \begin{bmatrix} 0 & 0 & mz_C r & -m(x_C r + v) \\ 0 & 0 & -my_C p & -m(y_C r - u) \\ -mz_C r & my_C p & 0 & I_{yz} r + I_{xy} p \\ m(x_C r + v) & m(y_C r - u) & -I_{yz} r - I_{xy} p & 0 \end{bmatrix},$$

Figure 3.5: Austal's hull H260 "Benchijigua Express." Picture courtesy of Austal Ships.

where m is the mass of the vessel, $\mathbf{r}_C^b = [x_C, y_C, z_C]^\mathsf{T}$ gives position the centre of mass relative to B, and I_{ik} are the moments and products of inertia about B.

The kinematic transformation (3.77) reduces to

$$\mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 & 0 \\ \sin(\psi) & \cos(\psi)\cos(\phi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cos(\phi) \end{bmatrix}.$$

The hydrodynamic forces due to the interaction of the vessel with the fluid are pressure induced forces that depend on both velocity and acceleration, and there are also restoring forces due to gravity and buoyancy that depend on the displacements. These forces are approximated by

$$\boldsymbol{\tau}_H = -\mathbf{M}_A\dot{\boldsymbol{\nu}} - \mathbf{C}_A(\boldsymbol{\nu})\boldsymbol{\nu} - \mathbf{D}(\boldsymbol{\nu})\boldsymbol{\nu} - \mathbf{g}(\boldsymbol{\eta}). \tag{3.84}$$

The first two terms on the right-hand side of (3.84) can be explained by considering the motion of the craft in an irrotational flow and for an ideal fluid (no viscosity). As the vessel moves, it changes momentum of the fluid. The kinetic energy of the ideal fluid due to the motion of the vessel can be expressed as

$$\mathcal{T}_A = \frac{1}{2}\boldsymbol{\nu}^T\mathbf{M}_A\boldsymbol{\nu}, \tag{3.85}$$

where the constant matrix $\mathbf{M}_A$ is called the matrix of added mass *coefficients. The first term on the right-hand side of (3.84) represents pressure-induced forces proportional to the accelerations of the vessel. The second term corresponds to Coriollis and centripetal forces due to the added mass. Note that the term "added mass" does not mean that the vessel moves a finite mass of fluid with it, rather that there are forces due to the acceleration of the fluid, and this is reflected in the mathematical model as a finite additional mass. The third term in (3.84) corresponds to damping forces, which have the following origins:*

Potential Damping: *This damping force is the result of a body passing through a fluid and wake wave making. The word* potential *indicates that this damping force can be obtained from a study of irrotational flow in an ideal fluid (no viscosity). In such a fluid, one can define a potential function of the space, such that its gradient gives* *the vector field of flow velocity, and the pressure can be computed from the potential function. As the fluid is displaced, there is a buildup of pressure in front, and a decrease of pressure behind, the body. This pressure difference induces a force that opposes the motion.*

Skin Friction: *This viscous effect is caused by the passage of water over the wetted surface of the hull. At low speeds, it is dominated by a linear component, while at higher speeds, nonlinear terms dominate.*

Vortex Shedding: *This damping force arises from the vortex generation due to flow separation (viscous effects), which occurs at the sharp edges commonly found at the bow and stern of a marine craft, or the control surfaces.*

Lifting Forces: *The hull of a craft can be modelled as a low aspect ratio wing. The forces generated by the hull at some angle of attack due to a manoeuvre can be resolved into two components: lift and drag. The former acts perpendicular to the direction of motion of the vessel, while the latter acts in the opposite direction of the motion. This is depicted in Figure 3.6.Note that the lift force is not actually a damping force as it is oriented perpendicular to the motion of the vessel, it does not serve to dissipate energy. With an abuse of terminology, all lift forces are generically grouped into a damping term. The drag component that arises due to lift is known as* parasitic drag*, or* lift-induced drag*. This force acts directly in opposition to motion, and is properly referred to as a damping force.*

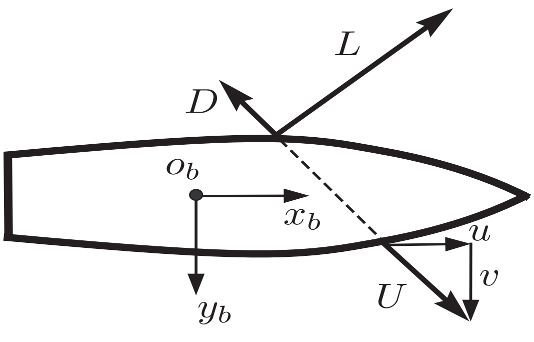


Figure 3.6: Lift and Drag forces experienced during manoeuvring.

The last term in (3.84) represents forces due to gravity and buoyancy. These forces tend to restore the up-right equilibrium of the vessel; and therefore, are called restoring forces. *These forces depend on the displacement volume of the vessel, its shape and heave, pitch, and roll angles.*

The added mass matrix and the Coriollis-centripetal matrix due to added mass are given by

$$\mathbf{M}_A = \mathbf{M}_A^T = -\begin{bmatrix} X_{\dot{u}} & 0 & 0 & 0 \\ 0 & Y_{\dot{v}} & Y_{\dot{p}} & Y_{\dot{r}} \\ 0 & K_{\dot{v}} & K_{\dot{p}} & K_{\dot{r}} \\ 0 & N_{\dot{v}} & N_{\dot{p}} & N_{\dot{r}} \end{bmatrix},$$

$$\mathbf{C}_A(\boldsymbol{\nu}) = \begin{bmatrix} 0 & 0 & 0 & Y_{\dot{v}}v + Y_{\dot{p}}p + Y_{\dot{r}}r \\ 0 & 0 & 0 & -X_{\dot{u}}u \\ 0 & 0 & 0 & 0 \\ -Y_{\dot{v}}v - Y_{\dot{p}}p - Y_{\dot{r}}r & X_{\dot{u}}u & 0 & 0 \end{bmatrix}.$$

The adopted damping terms take into account lift, drag, and viscous effects.

$$\mathbf{D}(\boldsymbol{\nu}) = \mathbf{D}_{LD}(\boldsymbol{\nu}) + \mathbf{D}_{VIS}(\boldsymbol{\nu}),$$

where

$$\mathbf{D}_{LD}(\boldsymbol{\nu}) = \begin{bmatrix} 0 & 0 & 0 & X_{rv}\,v \\ 0 & Y_{uv}\,u & 0 & Y_{ur}\,u \\ 0 & K_{uv}\,u & 0 & K_{ur}\,u \\ 0 & N_{uv}\,u & 0 & N_{ur}\,u \end{bmatrix}. \tag{3.86}$$

$$\mathbf{D}_{VIS}(\boldsymbol{\nu}) = \begin{bmatrix} X_{u|u|} & 0 & 0 & 0 \\ 0 & Y_{|v|v}\,|v| + Y_{|r|v}\,|v| & 0 & Y_{|v|v}\,|v| + Y_{|r|r}\,|r| \\ 0 & 0 & K_{p|p|} + Y_p & 0 \\ 0 & N_{|v|v}\,|v| + N_{|r|v}\,|v| & 0 & N_{|v|v}\,|v| + N_{|r|r}\,|r| \end{bmatrix}. \tag{3.87}$$

Finally the restoring term reduces to

$$\mathbf{g}(\boldsymbol{\eta}) = [0,\, 0,\, M_g(\boldsymbol{\eta}),\, 0]^T.$$

The propulsion forces take the form

$$\boldsymbol{\tau}_P = \mathbf{B}_P\,\mathbf{u}_P,$$

where $\mathbf{u}_P$ *are the forces of the individual propulsion units and* $\mathbf{B}_P$ *is a matrix that maps the forces* $\mathbf{u}_P$ *into* $\boldsymbol{\tau}_P$*. The matrix* $\mathbf{B}_P$ *depends on the location of the propulsion units on the vessel. The control forces have a similar form:*

$$\boldsymbol{\tau}_C = \mathbf{B}_C\,\mathbf{u}_C,$$

where $\mathbf{u}_C$ *are the control surface forces (rudders, fins, trim tabs, interceptors, etc.), and* $\mathbf{B}_C$ *is a matrix that maps the forces* $\mathbf{u}_C$ *into* $\boldsymbol{\tau}_C$*.*

Figure 3.7 shows model validation data for the velocities based on a full scale zig-zag sea trial. The hydrodynamic parameters of the model were partially obtained from computational fluid dynamic software, and partially by optimising model prediction errors using full scale trial data. For further details see Perez et al. (2007).

$\square$

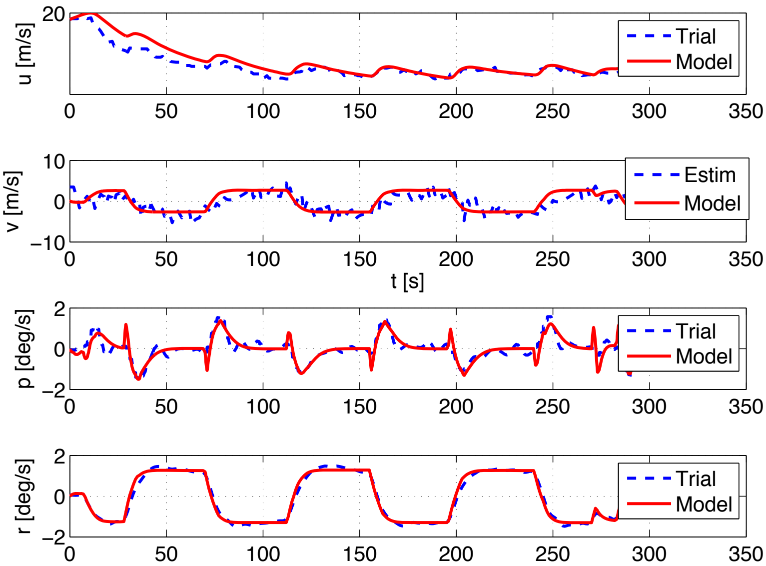


Figure 3.7: Model validation against a zig-zag sea trial data for a high speed trimaran ferry.

Chapter 4

Analytical Mechanics

"Put off thy shoes from off thy feet,
for the place whereon thou standest is holy ground."

EXODUS III, 5

The approach to study of mechanics based the on the idea of using energy as fundamental quantity rather than force and momentum was pushed by Leibniz, put to solid foundations by Euler and Lagrange, and then further elaborated by Hamilton and Jacobi. The techniques that follow from this line of thought are known as **Analytical Mechanics**—this name derives from Lagrange's book *Mécanique Analytique* published in 1788, in which all the developments are purely based on calculus (mathematical analysis) (Lanczos, 1970).

Analytical mechanics provides a neat way to handle the modeling and analysis of mechanical systems subject to constraints, which has not parallel in the vectorial approach. The modeling process is constructive:

1. Choose a set of generalized coordinates,

2. Establish the system kinetic and potential energy relative to an inertial frame,

3. Derive the equations of motion from either Lagrange, Lagrange-D'Alembert, or Hamilton equations.

4.1 Generalised Coordinates and Configuration Space

A central concept of analytical mechanics is that of coordinates used to describe the configuration of the system.

The **configuration** of a mechanical system is specified by the location of all its particles. The configuration can thus be fully specified by the Cartesian coordinates of all the particles:

$$x_i, y_i, z_i, \qquad i = 1, 2, \ldots, N,$$

where N is the number of particles.

If the N particles move freely in the three-dimensional space, then we need $3N$ coordinates to specify the configuration. If the particles are constrained in their positions or velocities, it is usually possible to specify the configuration of the system by fewer than $3N$ coordinates.

A minimal set of $n \leq 3N$ coordinates that fully specify the configuration of a system of N particles is called a set of **generalized coordinates** and denoted q_i with $i = 1, 2, \ldots, n$.

The generalized coordinates are related to the cartesian coordinates of any point of interest in the system by transformation equations:

$$\begin{aligned} x_i &= x_i(q_1, q_2, ..., q_n, t), \\ y_i &= y_i(q_1, q_2, ..., q_n, t), \\ z_i &= y_i(q_1, q_2, ..., q_n, t). \end{aligned}$$

Example 9 (Planar Robotic Manipulator) *Consider the planar robotic manipulator shown in Figure 2.9. For this system, we can choose as generalized coordinates $q_1 = \psi_1$ and $q_2 = \psi_2$. Then, the location of any point of the system can be determined from knowledge of these two coordinates and the geometry of the manipulator (dimensions of the links). For example the location of the tool is given by (2.102):*

$$\begin{aligned} x_{B/O} &= l_1 \cos q_1 + l_2 \cos q_1 \, \cos q_2 - l_2 \sin q_1 \, \sin q_2, \\ y_{B/O} &= l_1 \sin q_1 + l_2 \sin q_1 \, \cos q_2 + l_2 \, \cos q_1 \, \sin q_2, \\ z_{B/O} &= 0. \end{aligned}$$

■

A key characteristics of the analytical approach is that the choice of generalised coordinates is irrelevant. In the example, above we could have chosen the angles of the two links relative to the horizontal instead of their relative angles and obtain a mapping between these two generalized coordinates and the location of the tool.

The number of generalized coordinates of a mechanical system is a unique characteristic of the system, and it is equal to the number of **degrees of freedom (DOF)** of the system.

The number of states required for a dynamic model of a mechanical system is, in general, twice the number of DOF.

The fact the the number of DOF is often less than the number of particles in the system is a consequence of the constraints. If a system is described by N coordinates $\tilde{q}_i$, but these coordinates are subject to m constraints of the form

$$g_k(\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_N) = 0, \qquad k = 1, 2, \ldots, m,$$

then we can, in general, choose a new set of $n = N - m$, generalized coordinates q_i, which are independent and describe the motion of the system in harmony with the constraints. In such case, n is the number of DOF of the system. If we go back to Example 1 of the pendulum in Chapter 1, the coordinates $\tilde{q}_1 = x, \tilde{q}_2 = y$ of the pendulum must satisfy the constraint imposed by the fact that the bar of the pendulum is rigid: $\tilde{q}_1^2 + \tilde{q}_2^2 - l^2 = 0$. Hence, the system has 1DOF. The angle of the pendulum can be chosen as a generalised coordinate, namely $q = \theta$.

Example 10 (Two particles rigidly attached) *Consider two particles moving in the 3-dimensional space, which are rigidly linked such that their distance remains constant. In this case, the Cartesian coordinates of the system, must satisfy a constraint:*

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = a^2. \tag{4.1}$$

In term of these coordinates, we have that the number of DOF is $6 - 1 = 5$.

We may chose to use the constraint (4.1) *to eliminate one of the cartesian coordinates. This approach is seldom recommended because it breaks symmetry (which variable are we going to eliminate and why?) and often leads to models with singularities. Instead, we could try to find a set of generalized coordinates in harmony with the constraints. The following sets of generalized coordinates are but two particular choices for this example:*

1. *Location of one particle (3 coordinates) and two angles that give the orientation of the line segment that joints it to the other particle.*

2. *Location of the centre of mass of the two particles and two angles that give the orientation of the line segment that joints them.*

■

The following are examples of systems with different degrees of freedom:

- 1DOF: A piston, a single pendulum,
- 2DOF: A particle moving on a surface,
- 3DOF: A rigid body moving on a surface, a particle moving in space,
- 4DOF: A robotic manipulator with 4 links with rotary joints,
- 5DOF: Two particles rigidly liked moving in space,
- 6DOF: A free rigid-body moving in space.

It follows from the Example 10, that the generalized coordinates may not be homogenous in nature, that is, they are usually a mix of positions and angles.

For a system with n generalised coordinates, we can consider the vector

$$\mathbf{q} = [q_1, q_2, \dots, q_n]^\top.$$

This vector belongs of a space called **configuration space** and the generalised coordinates are often referred to as **configuration variables**. As the mechanical system moves in the 3-dimensional space, the vector $\mathbf{q}$ traces a curve in the n-dimensional configuration space. Therefore, the use of generalised coordinates allows us to transfer the study of motion of a system to that of the motion of a single point in the configuration space. This picture provides a useful link between mechanics and geometry.

The time derivative of the generalised coordinates are called **generalised velocities**, namely $\dot{q}_i$, and we also consider the vector of generalised velocities:

$$\dot{\mathbf{q}} = [\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n]^\top.$$

4.2 Constraints

Kinematic constraints are relations among coordinates,

$$f_k(q_1, q_2, \dots, q_n, t) = 0, \qquad k = 1, 2, \dots, m, \tag{4.2}$$

or relations between the coordinates and their differentials,

$$g_k(q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, t) = 0, \qquad k = 1, 2, \dots, m. \tag{4.3}$$

Constraints of the form (4.2) and constraints of the form (4.3) that can be integrated into the form (4.2) are called **holonomic**. A system that is subject to holonomic constraints is called a **holonomic system**.

Constraints of the form (4.3) that cannot be integrated into the form (4.2) are called **nonholonomic**. A system that is subject to nonholonomic constraints is called a **nonholonomic system**.

Constraints of the form (4.3) are often affine in the generalized velocities $\dot{q}_i$, that is

$$g_k(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_{i=1}^{n} G_{ki}(\mathbf{q}, t)\, \dot{q}_i + G_{kt}(\mathbf{q}, t) = 0, \qquad k = 1, 2, \dots, m. \tag{4.4}$$

or in differential form

$$\sum_{i=1}^{n} G_{ki}(\mathbf{q}, t)\, dq_i + G_{kt}(\mathbf{q}, t)\, dt = 0, \qquad k = 1, 2, \dots, m. \tag{4.5}$$

The constraints (4.4) and (4.5) are called **Pfaffian constraints** (Greenwood, 2003).

Note that a holonomic constraint of the form (4.2) can be differentiated

$$\sum_{i=1}^{n} \frac{\partial f_k(\mathbf{q},t)}{\partial q_i}\,\dot{q}_i + \frac{\partial f_k(\mathbf{q},t)}{\partial t}\,dt = 0, \qquad k = 1,2,\ldots,m, \tag{4.6}$$

which leads to the forms (4.4) and (4.5) by taking $G_{ki}(\mathbf{q},t) \triangleq \partial f_k(\mathbf{q},t)/\partial q_i$ and $G_{kt}(\mathbf{q},t) \triangleq \partial f_k(\mathbf{q},t)/\partial t$. Constrains of the form (4.4) and (4.5), however, may not be integrated into the form (4.2). Necessary and sufficient conditions for the constraints (4.4) to be integrable are

$$\frac{d}{dt}\left(\frac{\partial g_k}{\partial \dot{q}_i}\right) - \frac{\partial g_k}{\partial q_i} = 0 \quad i = 1,2,\ldots,n, \tag{4.7}$$

or that $M(\mathbf{q},t)\,g_k(\mathbf{q},\dot{\mathbf{q}},t)$, where $M(\mathbf{q},t)$ is an integration factor, satisfy (4.7)—see Greenwood (2003) for details.

Example 11 (Integrable differential constraint) *Consider a cylinder of radius r rolling without slipping on an inclined plane—see Figure 4.1. We can choose as coordinates $\tilde{q}_1 = x$ as the location of the point of contact of the cylinder along the slope and $\tilde{q}_2 = \theta$ the angle of rotation of the cylinder. For this system, the following constraint holds*

$$\dot{\tilde{q}}_1 = r\,\dot{\tilde{q}}_2. \tag{4.8}$$

This constraint satisfy (4.7); and it is therefore it is integrable. Hence, the constraint (4.8) is holonomic. Indeed, it can be integrated to

$$\tilde{q}_1 = r\,\tilde{q}_2. \tag{4.9}$$

This indicates that the system configuration could be described by a single configuration, and thus, it has 1DOF. As a generalised coordinate, we could choose, for example the angle θ. ■

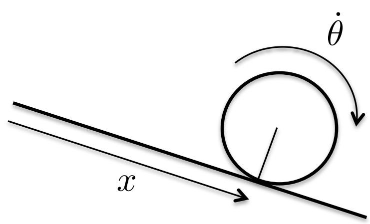


Figure 4.1: Rolling Cylinder.

Example 12 (Nonholonomic constraint) *Consider the rolling vertical disk of radius r shown in Figure 4.2. The configuration of the system can be described using the following generalised coordinates:*

$$q_1 = x_P, \qquad q_2 = y_P, \qquad q_3 = \psi, \qquad q_4 = \theta.$$

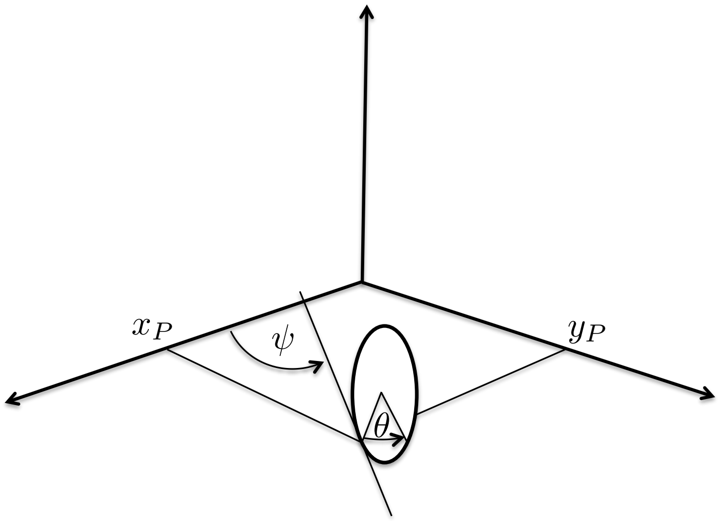


Figure 4.2: Rolling vertical disk.

If the disc rolls without slipping, then the following constraints hold:

$$\dot{q}_1 - (r\dot{q}_4 \cos q_3) = 0,$$
$$\dot{q}_2 - (r\dot{q}_4 \sin q_3) = 0,$$

which are non-integrable—there is no equation for $\dot{q}_4$. ■

Holonomic systems are constrained in their configuration. If we use n coordinates and we have m holonomic constraints, then the admissible coordinates are confined to a manifold[1] of $n-m$ dimensions. For these systems, we can work with a surplus of coordinates and the constraints, or find a set of $n-m$ generalized coordinates in harmony with the constraints; and therefore, eliminate the constraints from the equations of motion. In Example 11, we have the option to work with $\tilde{q}_1$ and $\tilde{q}_2$ subject to the constraint (4.9). In this case the manifold is simply a line in the plane $\tilde{q}_1$-$\tilde{q}_2$. Alternatively we can choose the rolling angle as a generalised variable $q = \theta$. A similar situation occurs in Example 1 of Chapter 1, in this case the manifold is a circle centred at the origin of the Euclidean space of the cartesian coordinates.

A nonholonomic system is not constrained in the configuration, but in the trajectories followed by the system when moving from one configuration to another. Unlike holonomic systems, it is not possible to find a set independent generalised coordinates; and therefore, the constraints cannot be eliminated from the equations of motion.

Example 13 (Nonholonomic system) *Mobile robots and vehicles with non-slipping wheels are nonholonomic systems. These systems move in the plane, and therefore we can describe their configuration with the position of a point in the vehicle and its heading angle as generalised coordinates, namely,*

$$q_1 = x, \qquad q_2 = y, \qquad q_3 = \psi,$$

[1]In general terms, a manifold is a curved space that resembles locally the Euclidean space (flat space). For example a circle is a manifold in $\mathbb{R}^2$ which locally resembles a line, and a sphere is a manifold in $\mathbb{R}^3$ which locally resembles a plane.

We can position a vehicle on the plane at any desired location and with any desired heading, but the constraints on the velocities

$$\dot{q}_1 = u \cos q_3,$$
$$\dot{q}_2 = u \sin q_3,$$

where u is the forward speed of the vehicle, restrict the admissible trajectories and are not integrable. For example if the vehicle is parked it cannot move sideways to a different location, but it can manoeuvre to end up parked sideways at the desired location. ■

If the constraints (4.2) and (4.3) are independent of time they are also called **scleronomic**, otherwise they are called **rehonomic**. If a mechanical system is subject to scleronomic constraints, then the energy can be conserved.

4.3 Principle of Virtual Work

We now enter the realm of analytical mechanics with a principle that describes the equilibtium of mechanical systems—statics. If we consider a configuration variable $q(t)$ as a function of time, then the differential of the function is the increment due to a differential increment in dt, that is

$$dq(t) = \dot{q}(t)\, dt.$$

Let us now consider a virtual displacement; a concept introduced by Lagrange:

A **virtual displacement** or **variation**, $\delta q(t)$, is a change of $q(t)$ in harmony with the constraints occurring at t and not due to dt. This displacement is something that we impose as a mathematical experiment and hence it is called virtual whereas $dq(t)$ is an actual displacement.

Consider a function $r(\mathbf{q}, t)$, then

$$\delta r = \frac{\partial r}{\partial q_1}\delta q_1 + \cdots + \frac{\partial r}{\partial q_n}\delta q_n. \tag{4.10}$$

In (4.10), the virtual displacements δq_i are at our disposal, but δr is prescribed once δq_i are chosen.

Suppose that a set of forces $\vec{F}_j$ is acting on a system at the points $\vec{r}_j$, then we can define the **virtual work** of the set as

$$\delta W \triangleq \sum_j \vec{F}_j \cdot \delta \vec{r}_j. \tag{4.11}$$

If we use generalised coordinates, this can be further expressed as

$$\delta W = \sum_j \vec{F}_j \cdot \left(\sum_i \frac{\partial \vec{r}_j}{\partial q_1}\delta q_1 + \cdots + \frac{\partial \vec{r}_j}{\partial q_n}\delta q_n \right). \tag{4.12}$$

We can define the **generalised forces**

$$Q_j \triangleq \sum_{i=1}^{n} \vec{F}_j \cdot \frac{\partial \vec{r}_j}{\partial q_i}, \tag{4.13}$$

and then, the virtual work (4.11) can be expressed as

$$\delta W = \sum_{i=1}^{n} Q_i \, \delta q_i. \tag{4.14}$$

Note that the units of each generalised force depends on the units of the associated generalised coordinate such that their product has units of energy.

We are now ready to state our first principle of analytical mechanics:

Principle 1 (Principle Virtual Work - PVW) *A mechanical system is in equilibrium if, and only if, the virtual work of the set of forces acting on the system is zero, namely,*

$$\delta W = \sum_{j} \vec{F}_j \cdot \delta \vec{r}_j = \sum_{i=1}^{n} Q_i \, \delta q_i = \mathbf{Q}^{\mathsf{T}} \, \delta \mathbf{q} = 0. \tag{4.15}$$

Equation (4.15) establishes that the forces must be perpendicular to the generalised displacements. If the system is free of any constraints, this means that the sum of the forces must vanish, for the displacements can be chosen in any direction, and zero is the only vector perpendicular to any choice of virtual displacements. This leads to the usual result in statics establishing that the condition for equilibrium is that the sum of acting forces and moments must vanish.

If there are constraints, however, then the virtual displacements must be in harmony the constraints, namely,

$$\delta f_k = \sum_i \frac{\partial f_k}{\partial q_i} \delta q_i = 0 \quad \text{(Holonomic)}, \tag{4.16}$$

$$\delta g_k = \sum_i G_{ki}(\mathbf{q}, t) \delta q_i = 0, \quad \text{(Nonholonomic-Pfaffian)}. \tag{4.17}$$

In these cases, the forces need only be perpendicular to the virtual displacements. Furhermore, since the system is in equilibrium, $\vec{F}_j = -\vec{F}_j^c$, where $\vec{F}_j^c$ are the forces of constraints. Hence,

$$\sum_j \vec{F}_j^c \cdot \delta \vec{r}_j = 0. \tag{4.18}$$

Equation (4.18) is, thus, a corollary that follows from the Principle of Virtual Work (Lanczos, 1970):

Corollary 1 (Virtual work of constraint forces) *The virtual work of the forces of constraint is always zero for virtual displacements in harmony with constraints; therefore, the virtual displacements are always perpendicular to the forces of constraints.*

This fundamental result extends beyond the realm of statics, which allows the formulation of d'Alembert's Principle—an extension of the Principle of Virtual work from statics to dynamics, which we will discuss in Section 4.4.

Example 14 (Equilibrium of the Pendulum) *Consider the pendulum in Figure 4.3, which is being acted by an external force $\vec{F}$ apart from the grabitational force $\vec{G}$. If we choose as coordinates $q_1 = x$ and $q_2 = y$, then, the constraint*

$$f: \qquad \sqrt{q_1^2 + q_2^2} - l = 0,$$

defines a circle of radius l centred at the origin of the configuration space—this circle is a manifold.

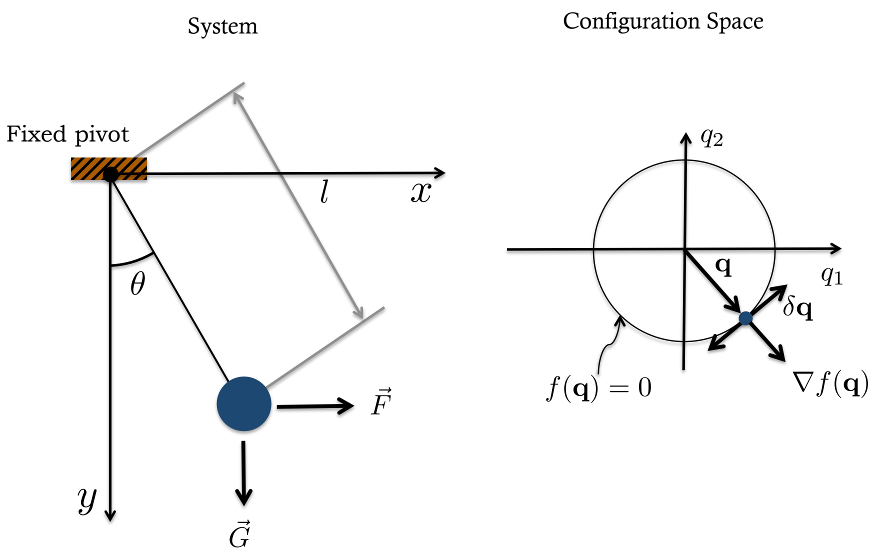


Figure 4.3: Pendulum with external force.

From (4.16), it follows that the virtual displacements in harmony with the constraints must satisfy,

$$\frac{\partial f}{\partial q_1}\delta q_1 + \frac{\partial f}{\partial q_2}\delta q_2 = 0,$$

which is equivalent to

$$\begin{bmatrix} \frac{\partial f}{\partial q_1} & \frac{\partial f}{\partial q_2} \end{bmatrix} \begin{bmatrix} \delta q_1 \\ \delta q_2 \end{bmatrix} = \frac{1}{2l} \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} \delta q_1 \\ \delta q_2 \end{bmatrix} = 0.$$

This last expression shows that the virtual displacements are perpendicular to the gradient of the constraint f as illustrated in Figure 4.3. This gradient points always in the radial *direction, which is the direction of the force of constraint (along the bar). Hence, the virtual displacements in harmony with the constraints are perpendicular to the force of constraint.*

Let us assume that the system is in equilibrium. Thus the virtual work of the acting forces must satisfy

$$\delta W = \vec{F} \cdot \delta\vec{r} + \vec{G} \cdot \delta\vec{r} = 0, \tag{4.19}$$

where $\vec{r}$ is the vector that describes the position of the pendulum. Let us define,

$$\mathbf{r} \triangleq \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \quad \mathbf{F} \triangleq \begin{bmatrix} F \\ 0 \end{bmatrix}, \quad \mathbf{G} \triangleq \begin{bmatrix} 0 \\ G \end{bmatrix}.$$

Using these, (4.19) becomes,

$$\delta W = F\,\delta q_1 + G\,\delta q_2 = 0, \tag{4.20}$$

which implies that the generalised force,

$$\mathbf{Q} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \triangleq \begin{bmatrix} F \\ G \end{bmatrix},$$

is perpendicular to the $\delta\mathbf{q}$. Since $\delta\mathbf{q}$ is tangential to the circle, $\mathbf{Q}$ must be radial for the system to be in equilibrium at any point on the circle. This translates into the condition that the resultant force $\vec{F_r} = \vec{F} + \vec{G}$ must be radial for the system to be in equilibrium.

Let us look at the same problem in a different way, and consider a generalised coordinate $\tilde{q} = \theta$, with leads to the coordinate transformations:

$$x = l\,\sin\tilde{q},$$
$$y = l\,\cos\tilde{q}.$$

Then,

$$\mathbf{r} = \begin{bmatrix} l\,\sin\tilde{q} \\ l\,\cos\tilde{q} \end{bmatrix}, \qquad \delta\mathbf{r} = \begin{bmatrix} l\,\cos\tilde{q} \\ -l\,\sin\tilde{q} \end{bmatrix} \delta\tilde{q}.$$

The virtual work (4.19), then becomes

$$\delta W = \mathbf{F}\,\delta\mathbf{r} + \mathbf{G}\,\delta\mathbf{r} = (F\,l\cos\tilde{q} - G\,l\sin\tilde{q})\,\delta\tilde{q} = 0, \tag{4.21}$$

which gives the generalised force

$$\tilde{Q} \triangleq (F\,l\cos\tilde{q} - G\,l\sin\tilde{q}). \tag{4.22}$$

Since $\delta\tilde{q}$ can be varied in any way, $\delta W = 0$ implies that

$$(F\,l\cos\tilde{q} - G\,l\sin\tilde{q}) = 0, \tag{4.23}$$

which gives the standard result: the resultant moment about the pivot must vanish. ■

4.4 d'Alembert's Principle

The french mathematician d'Alembert (1717-1785) re-arranged Newton's second law in the following form:

$$\vec{F} - m\vec{a} = \vec{0},$$

and defined the **force of inertia** $\vec{F}^I \triangleq -m\vec{a}$. By do ing this, Newtow's law becomes an equilibrium condition, namely the balance of external and inertial forces. This remarkably simple observation allowed d'Alembert to use the Principle to Virtual Work to study dynamic systems. That is, for a system of N particles, the 'equilibrium' condition results in

$$\delta W = \sum_{j=1}^{N} (\vec{F}_j + \vec{F}_j^c + \vec{F}_j^I) \cdot \delta \vec{r}_i = 0,$$

where $\vec{F}_j$ is the resultant of external forces. Using the corollary of the principle of virtual work, we can eliminate the forces of constraint since they do not do virtual work. This leads to our second principle of analytical mechanics:

Principle 2 (d'Alembert's Principle) *The motion of a mechanical system made of N particles is such that the virtual work of the sum of the acting forces $\vec{F}_j$ and the force of inertia $\vec{F}_j^I = -m\,\ddot{\vec{r}}_j$ is zero:*

$$\delta W = \sum_{j=1}^{N} (\vec{F}_j - m_j\,{}^{\mathcal{N}}\ddot{\vec{r}}_j) \cdot \delta \vec{r}_j = \vec{0},$$

where the accelerations are relative to an inertial frame $\mathcal{N}$.

Note that if the particles are free of constraints, then $\delta\vec{r}_j$ are independent, and therefore, d'Alembert's Principle reverts to Newton's second Law since each term in parenthesis must vanish. If the particles are constrained, d'Alembert's Principle provides a way to to deal with such problems. In fact the equations of motion of many mechanical systems can be derived directly from d'Alembert Principle. We will not pursue this path here, the interested student, can consult Egeland and Gravdahl (2002) and Kane and Levinson (1985). If you encounter a problem with a multi-body system, d'Alembert's Principle is a powerful tool to derive the equations of motion.

4.5 Kinetic and Potential Energy

Considered the motion in an inertial frame of particle acted by a force—see Figure 4.4.

The **power** supplied by the force to the particle is

$$P_W(t) = \vec{F}(t) \cdot \vec{v}(t) \qquad [\text{Watt} = \text{N m/s}], \tag{4.24}$$

The **work** done by the force in the time interval $[t_1, t_2]$ is defined as

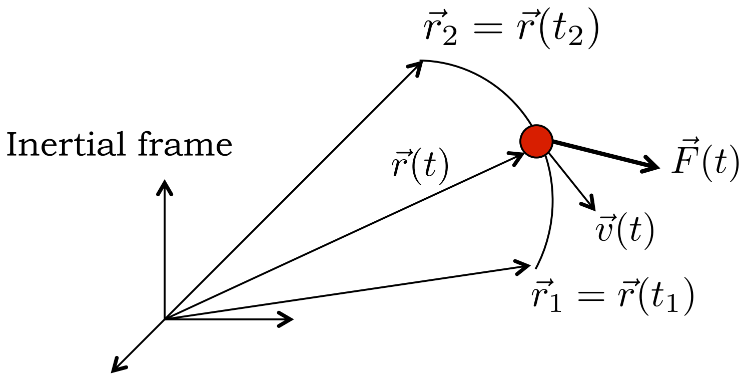


Figure 4.4: Particle acted by a force.

$$W_{12} = \int_{t_1}^{t_2} P_W(t)\ dt \qquad [\mathrm{J} = \mathrm{Watt\,s} = \mathrm{N\ m}] \tag{4.25}$$

The work can also be expressed as a line integral over the trajectory followed by the particle, that is,

$$W_{12} = \int_{t_1}^{t_2} \vec{F}(t) \cdot \frac{d\vec{r}(t)}{dt}\ dt = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F}(\vec{r}) \cdot d\vec{r}, \tag{4.26}$$

where $\vec{r}_1 = \vec{r}(t_1)$ and $\vec{r}_2 = \vec{r}(t_2)$. If the particle has mass m, we can use Newton's Second Law and express the work as

$$W_{12} = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F}(\vec{r}) \cdot d\vec{r} = m \int_{t_1}^{t_2} \frac{d\vec{v}(t)}{dt} \cdot \vec{v}(t)\ dt.$$

If we denote the speed as $v = |\vec{v}|$, then

$$v^2 = \vec{v} \cdot \vec{v},$$

and

$$\frac{d}{dt}v^2 = \frac{d}{dt}\vec{v} \cdot \vec{v} + \vec{v} \cdot \frac{d}{dt}\vec{v} = 2\frac{d}{dt}\vec{v} \cdot \vec{v} \quad \Leftrightarrow \quad \frac{1}{2}\frac{d}{dt}v^2 = \frac{d}{dt}\vec{v} \cdot \vec{v}.$$

Hence,

$$W_{12} = m \int_{t_1}^{t_2} \frac{1}{2}\ \frac{d}{dt}v^2(t)\ dt = \frac{1}{2}m \int_{v_1^2}^{v_2^2} dv^2 = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2, \tag{4.27}$$

where $v_1^2 = |\vec{v}(t_1)|^2$ and $v_2^2 = |\vec{v}(t_2)|^2$. Then we can define the **kinetic energy** as

$$\mathcal{T}(t) \triangleq \frac{1}{2}m|\vec{v}(t)|^2. \tag{4.28}$$

Substituting (4.28) into (4.27), we can see that in an inertial frame, the work done by the force is equal to the increment of kinetic energy:

$$W_{12} = \mathcal{T}_2 - \mathcal{T}_1. \tag{4.29}$$

The kinetic energy of a system of particles can be expressed as the sum of the individual kinetic energies:

$$\mathcal{T} = \frac{1}{2}\sum_{i=1}^{N} m_i \vec{v}_i \cdot \vec{v}_i \tag{4.30}$$

Let us consider the components of the position vector of the particle k in an inertial frame with basis $\{n\}$ and its derivative:

$$\mathbf{r}_k^n = [r_1, r_2, r_3]^\mathsf{T}, \quad {}^n\dot{\mathbf{r}}_k^n = [\dot{r}_1, \dot{r}_2, \dot{r}_3]^\mathsf{T}.$$

Then,

$$\mathcal{T} = \frac{1}{2}\sum_{k=1}^{3N} m_k \, \dot{r}_k^2, \tag{4.31}$$

where $\mathbf{r}_1^n = [r_1, r_2, r_3]^\mathsf{T}$, $\mathbf{r}_2^n = [r_4, r_5, r_6]^\mathsf{T}$, and so on. In terms of generalized coordinates,

$$\dot{r}_k = \sum_{i=1}^{n} \frac{\partial r_k}{\partial q_i}\,\dot{q}_i + \frac{\partial r_k}{\partial t}, \qquad k = 1, 2, \ldots, 3N, \tag{4.32}$$

and therefore,

$$\mathcal{T} = \frac{1}{2}\sum_{k=1}^{3N} m_k \left(\sum_{i=1}^{n} \frac{\partial r_k}{\partial q_i}\,\dot{q}_i + \frac{\partial r_k}{\partial t}\right)\left(\sum_{j=1}^{n} \frac{\partial r_k}{\partial q_j}\,\dot{q}_j + \frac{\partial r_k}{\partial t}\right). \tag{4.33}$$

Expanding and grouping terms, this leads to the following general representation (Greenwood, 2003):

$$\mathcal{T} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}(\mathbf{q}, t)\,\dot{q}_i\,\dot{q}_j + \sum_{i=1}^{n} a_i(\mathbf{q}, t)\,\dot{q}_i + \frac{1}{2}\sum_{k=1}^{3N} m_k \left(\frac{\partial r_k}{\partial t}\right)^2 \tag{4.34}$$

where

$$m_{ij}(\mathbf{q}, t) = \sum_{k=1}^{3N} m_k \frac{\partial r_k}{\partial q_i}\frac{\partial r_k}{\partial q_j}, \qquad i, j = 1, 2, \ldots, n, \tag{4.35}$$

$$a_i(\mathbf{q}, t) = \sum_{k=1}^{3N} m_k \frac{\partial r_k}{\partial q_i}\frac{\partial r_k}{\partial t}, \qquad i = 1, 2, \ldots, n. \tag{4.36}$$

Note that if the system is scleronomic, only the first term in (4.34) is present. The second term in (4.34), which is linear in the generalized velocities are called gyroscopic terms.

Let us assume that each of the N particles is acted upon by a force $\vec{F}_i$, and that there exists a potential function $\mathcal{V}(\mathbf{r}, t)$ (a scalar-valued function of a vector variable) such that

$$\mathbf{F}_i = -\frac{\partial \mathcal{V}}{\partial \mathbf{r}_i} \equiv -\left.\frac{\partial \mathcal{V}(\mathbf{r}, t)}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{r}_i}. \tag{4.37}$$

The virtual work

$$\delta W = \sum_{k=1}^{3N} \mathbf{F}_k \cdot \delta \mathbf{r}_k = -\sum_{k=1}^{3N} \frac{\partial \mathcal{V}}{\partial \mathbf{r}_k} \cdot \delta \mathbf{r}_k = -\delta \mathcal{V}. \tag{4.38}$$

Now we can use generalized coordinates, and consider $V(\mathbf{q}, t)$. Then,

$$\delta W = -\delta \mathcal{V} = \sum_{j=1}^{n} \frac{\partial \mathcal{V}}{\partial q_j} \delta q_j. \tag{4.39}$$

We note, however, that

$$\delta W = \sum_{j=1}^{n} Q_j \delta q_j, \tag{4.40}$$

where Q_j is a generalized force. Hence,

$$Q_j = -\frac{\partial \mathcal{V}}{\partial q_j} \equiv -\left.\frac{\partial \mathcal{V}(\mathbf{q}, t)}{\partial q}\right|_{q=q_j}. \tag{4.41}$$

For the particular case where $\mathcal{V}(\mathbf{q}, t)$ is independent of t, then $\mathcal{T} + \mathcal{V}$ are conserved and the system is called **conservative**. It further follows from the principle of virtual work and (4.39), that if the system is in equilibrium, then the potential energy is stationary; that is if $\bar{\mathbf{q}}$ is an equilibrium configuration, then

$$\delta \mathcal{V}(\bar{\mathbf{q}}) = 0 \quad \Leftrightarrow \quad \mathbf{0} = \left.\frac{\partial \mathcal{V}}{\partial \mathbf{q}}\right|_{\mathbf{q}=\bar{\mathbf{q}}}. \tag{4.42}$$

If the stationary point is a minimum, then the system is in a stable equilibrium point.

So the stable equilibrium points of mechanical systems are those for which the potential energy is minimum—think of the pendulum in a gravitational field example.

We will see that this concept can be used for control of mechanical systems, where the control forces are used to shape the potential energy of the closed-loop system such that it assumes a minimum at the desired equilibrium point. This is called energy-shaping control, and it is part of set of tools known as passivity-based control.

4.6 Kinetic and Potential Energy of a Rigid Body

For a rigid body, the kinetic energy can be defined as a volume integral:

$$\mathcal{T}_{RB} = \frac{1}{2}\int_{\mathcal{B}} \vec{v}_P \cdot \vec{v}_P \, dm, \tag{4.43}$$

where $\vec{v}_p$ is the velocity, in an inertial frame $\mathcal{N}$, of the element of mass dm. We can use the transport theorem and express the velocity of the particle in terms of that of the centre of mass and the angular velocity of the rigid body:

$$\vec{v}_P = \vec{v}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{P/C}. \tag{4.44}$$

Then,

$$\begin{aligned} \mathcal{T}_{RB} &= \frac{1}{2}\int_{\mathcal{B}} (\vec{v}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{P/C}) \cdot (\vec{v}_C + \vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{P/C}) \, dm, & (4.45)\\ &= \frac{1}{2} m \, \vec{v}_C \cdot \vec{v}_C + \frac{1}{2}\int_{\mathcal{B}} (\vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{P/C}) \cdot (\vec{\omega}_{\mathcal{B}/\mathcal{N}} \times \vec{r}_{P/C}) \, dm, & (4.46)\\ &= \frac{1}{2} m \, \vec{v}_C \cdot \vec{v}_C - \frac{1}{2}\vec{\omega}_{\mathcal{B}/\mathcal{N}} \cdot \int_{\mathcal{B}} \vec{r}_{P/C} \times \vec{r}_{P/C} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}} \, dm. & (4.47) \end{aligned}$$

Using the definition of the inertia tensor (3.22),

$$\vec{I}_C = -\frac{1}{2}\int_{\mathcal{B}} \vec{r}_{P/C} \times \vec{r}_{P/C} \times (\cdot) \, dm.$$

the **kinetic energy of the rigid body** can be expressed as

$$\mathcal{T}_{RB} = \frac{1}{2} m \, \vec{v}_C \cdot \vec{v}_C + \frac{1}{2} \, \vec{\omega}_{\mathcal{B}/\mathcal{N}} \cdot \vec{I}_C \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{4.48}$$

The latter shows that the total kinetic energy can be expressed as the kinetic energy of a particle of mass m concentrated at the centre of mass plus the kinetic of rotation of the body about its centre of mass.

We can express (4.48) in matrix form in body coordinates:

$$\mathcal{T}_{RB} = \frac{1}{2} \begin{bmatrix} \mathbf{v}_C^b \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} m\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_C^b \end{bmatrix} \begin{bmatrix} \mathbf{v}_C^b \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix}. \tag{4.49}$$

Alternatively, we can express it in inertial coordinates:

$$\mathcal{T}_{RB} = \frac{1}{2} \begin{bmatrix} \mathbf{v}_C^n \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} m\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_b^n(\mathbf{q}) \, \mathbf{I}_C^b \, (\mathbf{R}_b^n(\mathbf{q}))^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_C^n \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n \end{bmatrix}, \tag{4.50}$$

where we have made use of the fact that the angular momentum can be expressed as

$$\mathbf{L}_C^b = \mathbf{I}_C^b \, \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \quad \Leftrightarrow \quad \mathbf{L}_C^n = \mathbf{R}_b^n(\mathbf{q}) \, \mathbf{I}_C^b \, (\mathbf{R}_b^n(\mathbf{q}))^T \, \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^n.$$

Following a similar derivation as above, the kinetic energy can also be expressed relative to a general point B in the body:

$$\mathcal{T}_{RB} = \frac{1}{2} m \, \vec{v}_B \cdot \vec{v}_B - \vec{v}_B \cdot (m \, \vec{r}_{C/B} \times \vec{\omega}_{\mathcal{B}/\mathcal{N}}) + \frac{1}{2} \, \vec{\omega}_{\mathcal{B}/\mathcal{N}} \cdot \vec{I}_B \cdot \vec{\omega}_{\mathcal{B}/\mathcal{N}}. \tag{4.51}$$

The latter, can be expressed, for example, in body coordinates, which leads to

$$\mathcal{T}_{RB} = \frac{1}{2} \begin{bmatrix} \mathbf{v}_B^b \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix}^\mathsf{T} \begin{bmatrix} m\mathbf{I} & -m\mathbf{S}(\mathbf{r}_{C/B}^b) \\ m\mathbf{S}(\mathbf{r}_{C/B}^b) & \mathbf{I}_B^b \end{bmatrix} \begin{bmatrix} \mathbf{v}_B^b \\ \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}^b \end{bmatrix} \tag{4.52}$$

In the above, we have expressed the rotational kinetic energy in terms of the angular velocity vector. We can alternatively use as generalized coordinates the positions and Euler angles, namely $\mathbf{q} = [n, e, d, \phi, \theta, \psi]^T$, this would lead to the following form for the rotational component:

$$\mathcal{T}_{RB} = \frac{1}{2} \dot{\mathbf{q}}^\mathsf{T} \, \mathbf{J}(\mathbf{q}) \mathbf{M}_{RB}^b \mathbf{J}(\mathbf{q}) \dot{\mathbf{q}}, \tag{4.53}$$

where

$$\mathbf{J}(\mathbf{q}) = \begin{bmatrix} \mathbf{R}_b^n(\mathbf{q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}(\mathbf{q}) \end{bmatrix}, \tag{4.54}$$

and

$$\mathbf{R}_b^n(\mathbf{q}) = \begin{bmatrix} c_\psi c_\theta & -s_\psi c_\phi + c_\psi s_\theta s_\phi & s_\psi s_\phi + c_\psi c_\phi s_\theta \\ s_\psi c_\theta & c_\psi c_\phi + s_\phi s_\theta s_\psi & -c_\psi s_\phi + s_\psi c_\phi s_\theta \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{bmatrix},$$

and

$$\mathbf{T}(\mathbf{q}) = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & \frac{s_\phi}{c_\theta} & \frac{c_\phi}{c_\theta} \end{bmatrix}, \quad t_\theta \equiv \tan(\theta), \qquad \cos(\theta) \neq 0.$$

The potential energy of a rigid body observed in an inertial frame $\mathcal{N}$ with basis $\{n\}$ and point of reference N can be expressed as follows:

$$\mathcal{V}_{RB}(\mathbf{q}) = m(\mathbf{g}^n)^\mathsf{T} \, \mathbf{r}_{C/N}^n(\mathbf{q}), \tag{4.55}$$

where $\mathbf{r}_{C/N}^n(\mathbf{q})$ is the position of the centre of mass, which in the basis $\{n\}$ depends on the body orientation, and thus, on $\mathbf{q}$.

4.7 Lagrange Equations for Holonomic Systems

Let us start with a motivation example of a single particle and derive Lagrange's equations from Newton's Second Law; this derivation follows from Spong et al. (2006). Consider the vertical motion of a particle of mass m moving under the action of a force and the gravitational field. Then,

$$m \, \ddot{q} = Q - mg, \tag{4.56}$$

where q is the vertical displacement positive upwards from certain reference point, and Q is the vertical component of the acting external force positive upwards.

The left-hand side of (4.56) can be expressed as follows:

$$m\,\ddot{q} = \frac{d}{dt} m\,\dot{q} = \frac{d}{dt}\frac{\partial}{\partial \dot{q}}\frac{1}{2} m\,\dot{q}^2 = \frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q}}, \tag{4.57}$$

where T is the Kinetic energy. The gravity term can be expressed as

$$-\,mg = -\frac{\partial}{\partial q}(mgq) = -\frac{\partial \mathcal{V}}{\partial q}. \tag{4.58}$$

We can now form the Lagrangian:

$$\mathcal{L} = \mathcal{T} - \mathcal{V} = \frac{1}{2} m\dot{q}^2 - mgq. \tag{4.59}$$

Note that

$$\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{T}}{\partial \dot{q}}, \qquad \frac{\partial \mathcal{L}}{\partial q} = -\frac{\partial \mathcal{V}}{\partial q}. \tag{4.60}$$

Then, (4.56) can be alternatively expressed as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = Q. \tag{4.61}$$

The latter is the Lagrange's Euquation of motion. This equation can be generalised to mechanical systems other than a particle. The modeling approach then consists of formulating the Lagrangian and then simply take partial derivatives to obtain the equations of motion.

In the sequel, we will make a more formal derivation based on d'Alembert's Principle. Such derivation follows from Greenwood (2003).

Let us consider a system of N particles with positions relative to an inertial frame with basis $\{n\}$ described in Cartesian coordinates. Then d'Alemebert's Principle states that

$$\sum_{k=1}^{3N} (F_k - m_k \ddot{r}_k)\delta r_k = 0, \tag{4.62}$$

where $\mathbf{r}_1^n = [r_1, r_2, r_3]^\mathsf{T}$, $\mathbf{r}_2^n = [r_4, r_5, r_6]^\mathsf{T}$, and so on; and the forces—other than the forces of constrstaint—acting on the particles are $\mathbf{F}_1^n = [F_1, F_2, F_3]^\mathsf{T}$, $\mathbf{F}_2^n = [F_4, F_5, F_6]^\mathsf{T}$ and so on.

Let us now consider a set of generalized coordinates; hence,

$$\delta r_k = \sum_{i=1}^{n} \frac{\partial r_k}{\partial q_i}\,\delta q_i, \qquad k = 1, 2, \ldots, 3N. \tag{4.63}$$

Substituting in (4.62), we obtain

$$\sum_{k=1}^{3N}\sum_{i=1}^{n}\left(F_k \frac{\partial r_k}{\partial q_i} - m_k \ddot{r}_k \frac{\partial r_k}{\partial q_i}\right)\delta q_i = 0, \tag{4.64}$$

Since,

$$\dot{r}_k = \sum_{i=1}^{n} \frac{\partial r_k}{\partial q_i}\,\dot{q}_i + \frac{\partial r_k}{\partial t}, \tag{4.65}$$

it follows that

$$\frac{\partial \dot{r}_k}{\partial \dot{q}_i} = \frac{\partial r_k}{\partial q_i}. \tag{4.66}$$

Also,

$$\frac{d}{dt}\left(\frac{\partial r_k}{\partial q_i}\right) = \sum_{j=1}^{n} \frac{\partial^2 r_k}{\partial q_j \partial q_i}\dot{q}_i + \frac{\partial^2 r_k}{\partial t \partial q_i} = \frac{\partial \dot{r}_k}{\partial q_i}. \tag{4.67}$$

The kinetic energy of the system is

$$\mathcal{T} = \frac{1}{2}\sum_{k=1}^{3N} m_k\,\dot{r}^2, \tag{4.68}$$

and the **generalized conjugate momentum** is defined as

$$p_i \triangleq \frac{\partial \mathcal{T}}{\partial \dot{q}_i} = \sum_{k=1}^{3N} m_k\,\dot{r}_k\,\frac{\partial \dot{r}_k}{\partial \dot{q}_i} = \sum_{k=1}^{3N} m_k\,\dot{r}_k\,\frac{\partial r_k}{\partial q_i}, \tag{4.69}$$

where in the last step we have used (4.66).

Combining the above, we can express

$$\frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{q}_i}\right) = \sum_{k=1}^{3N} m_k\,\ddot{r}_k\,\frac{\partial r_k}{\partial q_i} + \sum_{k=1}^{3N} m_k\,\dot{r}\,\frac{\partial \dot{r}_k}{\partial q_i}. \tag{4.70}$$

On the other hand,

$$\frac{\partial \mathcal{T}}{\partial q_i} = \sum_{k=1}^{3N} m_k\,\dot{r}_k\,\frac{\partial \dot{r}_k}{\partial q_i}. \tag{4.71}$$

Therefore, (4.70) can be written as

$$\sum_{k=1}^{3N} m_k\,\ddot{r}_k\,\frac{\partial r_k}{\partial q_i} = \frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{T}}{\partial q_i}. \tag{4.72}$$

The right-hand side of the latter is the generalised inertia force Q_i^I. d'Alembert's principle can then be expressed as

$$\sum_{i=1}^{n}(Q_i + Q_i^I)\,\delta q_i = 0, \tag{4.73}$$

where the generalised forced Q_i is given in (4.13), namely,

$$Q_i = \sum_{i=1}^{n} F_k \frac{\partial r_k}{\partial q_i}. \tag{4.74}$$

Then, we obtain a form of d'Alembert's Principle called **Lagrange's Principle**:

$$\sum_{i=1}^{n}\left[\frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{q}_i}\right)-\frac{\partial \mathcal{T}}{\partial q_i}-Q_i\right]\delta q_i=0, \tag{4.75}$$

which is valid for both holonomic and non-holonomic systems due to the fact that d'Alembert's Principle is valid for both these classes of systems.

If the system is holonomic and q_i are generalized variables (independent), this implies that term in brackets in (4.75) must be zero. This leads to the **holonomic form of Lagrange's Equations**:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{q}_i}\right)-\frac{\partial \mathcal{T}}{\partial q_i}=Q_i, \qquad i=1,2,\ldots,n. \tag{4.76}$$

Let Q_i be separated into two components: a conservative component, which can be derived from a potential function, and a general component that cannot be obtained from a potential function, namely,

$$Q_i=-\frac{\partial \mathcal{V}}{\partial q_i}+Q_i'. \tag{4.77}$$

We can now define the **Lagrangian**:

$$\mathcal{L}(\mathbf{q},\dot{\mathbf{q}},t)\triangleq\mathcal{T}(\mathbf{q},\dot{\mathbf{q}},t)-\mathcal{V}(\mathbf{q},t). \tag{4.78}$$

and then the Lagrange Equations take the following form:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)-\frac{\partial \mathcal{L}}{\partial q_i}=Q_i', \qquad i=1,2,\ldots,n, \tag{4.79}$$

where the generalized forces Q_i' do not derive from a potential function.

4.8 Lagrange Equations for Nonholonomic Systems

If we consider systems with nonholonmic constraints, then the generalized coordinates are no longer independent; that is, we are forced to work with a surplus of variables subject to constraints in their differentials. The generalised forces of constraints F_i^c are as follows:

$$Q_i^c=\sum_{k=1}^{3N}F_k^c\frac{\partial r_k}{\partial q_i}, \tag{4.80}$$

where F_k^c are the cartesian coordinates of the forces of constraint in an inertial frame. The Principle of Virtual Work establishes that

$$\sum_{i=1}^{n} Q_i^c \, \delta q_i = 0, \tag{4.81}$$

provided that the virtual displacements satisfy the m nonholonomic constraints:

$$\sum_{i=1}^{n} G_{ji} \, \delta q_i = 0, \qquad j = 1, 2, \ldots, m. \tag{4.82}$$

We can now multiply by a **Lagrange multiplier** λ_i and sum over j:

$$\sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j G_{ji} \, \delta q_i = 0. \tag{4.83}$$

Substracting this last expression from (4.81), we obtain

$$\sum_{i=1}^{n} \left(Q_i^c - \sum_{j=1}^{m} \lambda_j G_{ji} \right) \delta q_i = 0. \tag{4.84}$$

We can choose the Lagrange multiplier such that the coefficient of each δq_i vanish; namely,

$$\lambda_j : \quad Q_i^c = \sum_{j=1}^{m} \lambda_j G_{ji}. \tag{4.85}$$

We can now add the forces of constraint to the Lagrange equations, which leads to the **nonholonomic form of Lagrange's Equations**:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{T}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{T}}{\partial q_i} = Q_i + \sum_{j=1}^{m} \lambda_j G_{ji}, \qquad i = 1, 2, \ldots, n. \tag{4.86}$$

If we can separate Q_i into a conservative component and a non-conservative components as in (4.77), then

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = Q_i' + \sum_{j=1}^{m} \lambda_j G_{ji}, \qquad i = 1, 2, \ldots, n. \tag{4.87}$$

The nonholonomic form of the Lagrangian equations (4.87) are also applicable to the case of holonomic systems with non independent variables q_i as we did in Example 1 of Chapter 1. In such cases,

$$G_{ik} = \frac{\partial f_k(\mathbf{q}, t)}{\partial q_i}. \tag{4.88}$$

Example 15 (Cylinder Rolling on a slope) *Consider a cylinder of mass m and radius r rolling down a slope of length ℓ and angle α relative to the horizontal—see Figure 4.1. Let us assume that the only force acting on the cylinder is due to gravity and that there is no friction along the slope and no slipping.*

The configuration of the system can be described in terms of x—the position of the point of contact along the slope, and θ—the angle of rotation about the axis of the cylinder. The system is subject to the following holonomic constraint:

$$\dot{x} = r\,\dot{\theta}. \tag{4.89}$$

Solution I - Modelling with an Independent Coordinate

The system moves in the plane, but since there is one constraint (4.89) which is holonomic, we can describe the configuration with a single generalised coordinate. Let us choose

$$q \triangleq \theta. \tag{4.90}$$

Integration of (4.89) leads to

$$x = r\,q. \tag{4.91}$$

The kinetic energy of the cylinder can be separated into a translational and a rotational component:

$$\mathcal{T} = \frac{1}{2}m\,\dot{x}^2 + \frac{1}{2}m\,r^2\,\dot{\theta}^2. \tag{4.92}$$

Using (4.91), we can express the kinetic energy as follows:

$$\mathcal{T}(\dot{q}) = \frac{1}{2}m\,r^2\,\dot{q}^2 + \frac{1}{2}m\,r^2\,\dot{q}^2 = m\,r^2\,\dot{q}^2. \tag{4.93}$$

The potential energy can be expressed as follows:

$$\mathcal{V} = mg(\ell - x)\sin\alpha = mg(\ell - rq)\sin\alpha. \tag{4.94}$$

Hence, the corresponding Lagrangian is

$$\mathcal{L} = \mathcal{T} - \mathcal{V} = m\,r^2\,\dot{q}^2 - mg(\ell - rq)\sin\alpha. \tag{4.95}$$

The partial derivatives are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 2mr^2\,\ddot{q}, \tag{4.96}$$

$$\frac{\partial L}{\partial q} = mgr\,\sin\alpha. \tag{4.97}$$

Since there is no external force,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \tag{4.98}$$

which leads to

$$\ddot{q} = \frac{g\sin\alpha}{2r}. \tag{4.99}$$

Note that the model is valid provided that $0 \leq q \leq \ell/r$.

Solution II - Modelling with Dependent Coordinates

Let us define the following dependent coordinates:

$$q_1 \triangleq x, \qquad q_2 \triangleq \theta, \tag{4.100}$$

with the constraint

$$f(q_1, q_2) = q_1 - r q_2 = 0. \tag{4.101}$$

We can now use the nonholonomic form of Lagrange equations:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = \lambda \frac{\partial f}{\partial q_i}, \quad i = 1, 2. \tag{4.102}$$

In terms of the chosen coordinates, the Lagrangian is

$$\mathcal{L} = \mathcal{T} - \mathcal{V} = \frac{1}{2} m \dot{q}_1^2 + \frac{1}{2} m r^2 \dot{q}_2^2 - mg(\ell - q_1)\sin\alpha. \tag{4.103}$$

Equation (4.102) *leads to*

$$m\ddot{q}_1 - \sin\alpha = \lambda, \tag{4.104}$$
$$mr^2\ddot{q}_2 = -r\,\lambda. \tag{4.105}$$

To compute λ*, we can take the second time derivative of the constraint* (4.101)

$$\ddot{q}_1 = r\ddot{q}_2. \tag{4.106}$$

Now we can use (4.104)-(4.105) *into* (4.106) *and we obtain:*

$$\lambda = -\frac{mg\,\sin\alpha}{2}. \tag{4.107}$$

This last expression together with (4.104)-(4.105) *completes the model.*

We can, for example, substitute (4.107) *into* (4.105) *and we obtain*

$$\ddot{q}_2 = \frac{g\sin\alpha}{2r}, \tag{4.108}$$

which is the same result we obtained in (4.99) *when we considered an independent coordinate.*

■

Example 16 (Rolling Vertical Disk) *Consider the vertical rolling disk shown in Figure 4.2. As configuration variables we chose*

$$q_1 = x_P, \qquad q_2 = y_P, \qquad q_3 = \psi, \qquad q_4 = \theta.$$

If the disc rolls without slipping, then the following constraints hold:

$$\dot{q}_1 - (r\dot{q}_4\,\cos q_3\,) = 0, \tag{4.109}$$
$$\dot{q}_2 - (r\dot{q}_4\,\sin q_3\,) = 0, \tag{4.110}$$

which are nonholonomic as discussed in Example 12. Consider also a torque τ_3, which acts vertically on the disc to produce a change in heading, and a torque τ_4, which acts to generate the rolling motion.

Since the system is nonholonomic, and we have two constraints, the nonholonomic Lagrange equations (4.87) reduce to

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = \tau_i + \lambda_1 G_{1i} + \lambda_2 G_{2i}, \qquad i = 1, 2, 3, 4. \tag{4.111}$$

From the Pfaffian constraints (4.109)-(4.110) and (4.4), it follows that

$$G_{11} = 1, \quad G_{12} = 0, \quad G_{13} = 0, \quad G_{14} = -r\cos q_3, \tag{4.112}$$

$$G_{21} = 0, \quad G_{22} = 1, \quad G_{23} = 0, \quad G_{24} = -r\sin q_3. \tag{4.113}$$

Since the disk is rolling vertically, there is no potential energy associated with the motion. Hence, the Lagrangian is made just of the kinetic energy:

$$\mathcal{L} = \frac{1}{2}m\,\dot{q}_1^2 + \frac{1}{2}m\,\dot{q}_2^2 + \frac{1}{2}J_z\,\dot{q}_3^2 + \frac{1}{2}J_y\,\dot{q}_4^2, \tag{4.114}$$

where J_z is the sagital moment of inertia (about the vertical axis through the centre of the disk) and J_y is the moment of inertia about the rolling axis.

Application of (4.111) leads to

$$m\,\ddot{q}_1 = \lambda_1 \tag{4.115}$$

$$m\,\ddot{q}_2 = \lambda_2 \tag{4.116}$$

$$J_z\,\ddot{q}_3 = \tau_3 \tag{4.117}$$

$$J_y\,\ddot{q}_4 = \tau_4 - \lambda_1 r\cos q_3 - \lambda_2 r\sin q_3. \tag{4.118}$$

We need to compute λ_i with the help of the constraints. From (4.115) and (4.116), it follows that

$$\lambda_1 = m\,\ddot{q}_1 = m\frac{d}{dt}\dot{q}_1 = m\frac{d}{dt}(r\dot{q}_4\,\cos q_3), \tag{4.119}$$

$$\lambda_2 = m\,\ddot{q}_2 = m\frac{d}{dt}\dot{q}_2 = m\frac{d}{dt}(r\dot{q}_4\,\sin q_3). \tag{4.120}$$

Taking the time derivatives and substituting into (4.118), leads to

$$J_z\,\ddot{q}_3 = \tau_3 \tag{4.121}$$

$$(J_y + mr^2)\,\ddot{q}_4 = \tau_4. \tag{4.122}$$

These equations together with the constraints form the sought mathematical model:

$$J_z\,\ddot{q}_3 = \tau_3 \tag{4.123}$$

$$(J_y + mr^2)\,\ddot{q}_4 = \tau_4, \tag{4.124}$$

$$\dot{q}_1 = r\dot{q}_4\,\cos q_3, \tag{4.125}$$

$$\dot{q}_2 = r\dot{q}_4\,\sin q_3. \tag{4.126}$$

■

4.9 Hamilton's Principle

Hamilton's Principle was formulated by Hamilton (1805-1865) and it is related to the Principle of Least Action previously established by Euler (1707-1783) and Lagrange (1736-1813) (Lanczos, 1970).

Hamilton's Principle: the configuration of a mechanical system described in terms of a set of generalised coordinates $q_1, q_2, \ldots, q_n$ evolves in such a way that the definite integral

$$I(q_1, q_2, \ldots, q_n) = \int_{t_1}^{t_2} L(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n)\, dt \tag{4.127}$$

is minimised, where L is the Lagrangian.

The integral (4.127) is called the **action**, and the principle shows the evolution of mechanical systems in nature is such that the action is minimised.

Euler solved this minimisation problem for conservative systems using standard differential calculus, but Lagrange, who also solved it, invented a new branch of mathematics called **Calculus of Variations**. When we try to find whether a function $f(x)$ has a extremum at a point $\bar{x}$, we use the differential $\tilde{x} = \bar{x} + dx$ to check if the differential of the function $df = f(\tilde{x}) - f(\bar{x})$, caused by dx is zero. If this holds true, we say that the function has a stationary point at $\bar{x}$, and we need a further condition on the second differentials to establish the nature of the stationary point (maximum, minimum, or saddle). The minimisation of the definite integral (4.127) is with respect to the functions $q_j(t)$. That is, for every set of functions $q_j(t)$ we may chose, the integral will gives us a number (because it is a definite integral), then we seek the functions $q_j(t)$ that give the smallest value of the integral. Lagrange devised the concept of variation δq, and expressed $\tilde{q} = \bar{q} + \delta q$. The variation δq is a function of our choice—a virtual displacement or mathematical experiment imposed on q_j at a given time but not triggered by a dt. If the variation $\delta I = I(\tilde{q}_1, \ldots, \tilde{q}_n) - I(\bar{q}_1, \ldots, \bar{q}_n)$ vanishes, then $\bar{q}_j$ makes I stationary. Using this, Hamilton's principle can be written as a **variational principle**:

$$\delta I(q_1, q_2, \ldots, q_n) = 0. \tag{4.128}$$

For conservative systems in which the generalised coordinates are not subject to constraints, (4.128) leads to the **Euler-Lagrange equations**:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \qquad j = 1, 2, \ldots, n. \tag{4.129}$$

The solutions $q_i(t)$ that satisfy (4.129) are those that make the action (4.127) stationary.

Hamilton's Principle is not something what we can use directly to solve problems, unlike d'Alembert's Principle. However, Hamilton's Principle leads to the Hamilton-Jacobi equation which is used for systems described in terms of partial differential equations. Also, Hamilton's Principle can be used to explain one of the fundamental properties of the Lagrangian formulation: **the invariance under transformations of generalised coordinates**. Following Gregory (2006), suppose that we have a set of generalised coordinates q_i and via a transformation, we obtain another set of generalized variables q'_i. This change of variables does not affect the value of the action provided that q_i and q'_i describe the same motion of the mechanical system:

$$\int_{t_1}^{t_2} L(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n)\, dt = \int_{t_1}^{t_2} L(q'_1, q'_2, \ldots, q'_n, \dot{q}'_1, \dot{q}'_2, \ldots, \dot{q}'_n)\, dt. \qquad (4.130)$$

In each case, Hamilton's Principle leads to the same Euler-Lagrange equations, which shows that the latter equations are invariant to changes of generalised coordinates.

4.10 Energy and Co-Energy

Consider the mass-spring system shown in Figure 4.5. The constitutive relations of the spring can be written as

$$e = k\,q, \qquad f = \dot{q}, \qquad (4.131)$$

where e denotes the force seen by the spring, and f denotes the velocity of the end of the spring.

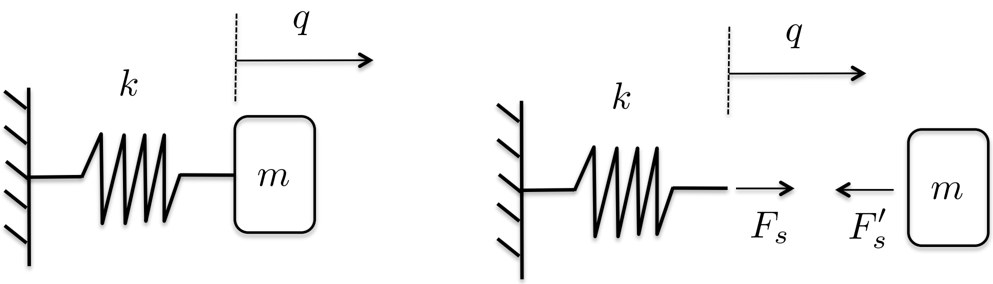


Figure 4.5: Mass-spring system.

Let us now consider the potential energy on the spring:

$$\begin{aligned}
\mathcal{V}(t) &= \mathcal{V}(0) + \int_0^t e f\, dt, \\
&= \mathcal{V}(0) + \int_{q_0}^{q} k\,q\, dq, \\
&= \mathcal{V}(0) + \frac{1}{2} k\, q^2 - \frac{1}{2} k\, q_0^2,
\end{aligned}$$

This is equivalent to

$$\mathcal{V}(q) = \frac{1}{2} k\, q^2, \qquad (4.132)$$

and that

$$e = \frac{\partial V}{\partial q} = k\, q,$$

which is the constitutive law of the spring.

Equation (4.132) shows that the energy stored in the spring is the area under the curve of its constitutive relation e *vs* q as illustrated in Figure 4.6. This figure also shows the so called **potential co-energy**, which satisfies the following relation:

$$V^*(e) = eq - V(q). \tag{4.133}$$

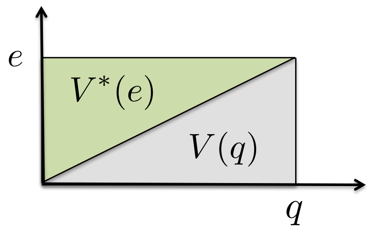


Figure 4.6: Potential energy and co-energy.

The co-energy can sometimes become useful to derive mathematical models. However, the capacity of the spring to do work is given by the energy and not the co-energy.

Consider now the mass or generalised inertia in Figure 4.5, which has constitutive relations:

$$p = I\, f, \qquad \dot{p} = e' = -e,$$

where f is the flow (velocity / angular velocity) and e' is the force seen by the mass. The kinetic energy stored in the inertia can be expressed as

$$\mathcal{T}(t) = \mathcal{T}(0) + \int_0^t f\, e'\, dt.$$

Using the constitutive relations, we can express the energy as

$$\mathcal{T}(t) = \mathcal{T}(0) + \int_{p_0}^{p} \frac{p}{I}\, dp.$$

$$\mathcal{T}(p) = \frac{1}{2I} p^2.$$

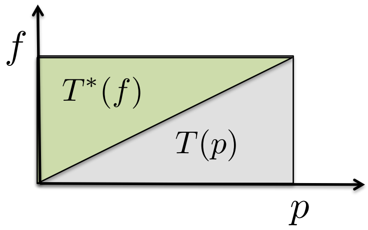


Figure 4.7: Kinetic energy and co-energy.

It follows from the derivation above, that the kinetic energy equals the area under the curve f-p of the constitutive relation as illustrated on the diagram on the left in Figure 4.7. From this figure, we can establish the following relationship:

$$\mathcal{T}^*(f) = fp - \mathcal{T}(p). \tag{4.134}$$

where $\mathcal{T}^*(f)$ is the so-called **co-kinetic energy**. Unless we consider motions close to the speed of light, the constitutive relations of inertias are linear for mechanical systems. Hence, the value of the kinetic energy and co-kinetic energy are the same:

$$\mathcal{T}(p) = \frac{1}{2I}p^2 = \frac{I}{2}f^2 = \mathcal{T}^*(f).$$

Because of this, many texts do not distinguish the kinetic energy from the co-kinetic energy, and often $\mathcal{T}^* = 1/2\,mv^2$ is called the kinetic energy.

The co-energy can sometimes become useful to derive mathematical models. However, the capacity of the inertia to do work is given by the energy and not the co-energy.

The relationships (4.133) and (4.134) follow a particular pattern, called the **Legendre Transformation**. The French mathematician Legendre (1752-1833) discovered a symmetric transformation in his work related to differential equations (Lanczos, 1970). Suppose that we have a function

$$F(u_1, u_2, \ldots, u_n),$$

and we introduce a new set of variables, called conjugate variables,

$$v_i \triangleq \frac{\partial F}{\partial u_i}, \quad i = 1, 2, \ldots, n, \tag{4.135}$$

which we assumed to be independent—that is, we assume that the determinant of the Hessian[2] of F is not zero.

[2]The Hessian of a scalar function $F(\mathbf{x})$ is a matrix denoted $\nabla^2 F$ or $\partial^2 F/\partial \mathbf{x}^2$ with entries $\nabla^2 F_{ij} = \partial^2 F/\partial x_i \partial x_j$.

We can now define a new function G:

$$G = \sum_{i=1}^{n} u_i\, v_i - F \tag{4.136}$$

The partial derivatives (4.135) result in a system of equations, which we can use to express each u_i as a function of $v_1, v_2, \ldots, v_n$. Using this we can write (4.136) as

$$G(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} u_i\, v_i - F(u_1, u_2, \ldots, u_n). \tag{4.137}$$

Expression (4.137) is known as the **Legendre Transformation**. This transformation is entirely symmetrical since we can reverse F and G in (4.137) and

$$u_i \triangleq \frac{\partial G}{\partial v_i}, \quad i = 1, 2, \ldots, n. \tag{4.138}$$

We can see that (4.133) and (4.134) satisfy (4.137). If we subtract the potential energy from (4.134), we obtain that

$$\mathcal{T}^*(f) - \mathcal{V}(q) = f p - \mathcal{T}(p) - \mathcal{V}(q) = p f - [\mathcal{T}(p) + \mathcal{V}(q)]. \tag{4.139}$$

The term on the left-hand side is the Lagrangian, and (4.139) shows that total energy and the Lagrangian satisfy the Legendre transformation.

Let us define the **Hamiltonian** as the total energy of the system:

$$\mathcal{H}(\mathbf{p}, \mathbf{q}, t) = \mathcal{T}(\mathbf{p}, \mathbf{q}, t) + \mathcal{V}(\mathbf{q}, t). \tag{4.140}$$

Then (4.139) can be be generalised to

$$\mathcal{H}(\mathbf{p}, \mathbf{q}, t) = \mathbf{p}^\top \dot{\mathbf{q}} - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t). \tag{4.141}$$

4.11 Hamilton's Equations

Hamilton's equations resemble a state-space model and have enabled significant contributions various branches of science like celestial mechanics, optics, quantum mechanics and control. Over the past 10 years, Hamilton's equations have attracted a significant attention from the control community, for they provide neat way of designing controllers for nonlinear physical systems by shaping the energy so it attains its minimum at the desired equilibrium point.

Let us consider the Legendre transformation (4.141). If we take a partial derivative on both sides of with respect of $\dot{\mathbf{q}}$, we obtain the **conjugate momentum**:

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} \qquad \equiv \qquad p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}. \tag{4.142}$$

If we take partial derivatives of (4.141),

$$\frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{q}}, \tag{4.143}$$

$$\frac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\frac{\partial \mathcal{L}}{\partial \mathbf{q}}, \tag{4.144}$$

$$\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \mathbf{p}. \tag{4.145}$$

From the latter and Euler-Lagrange equation,

$$\dot{\mathbf{p}} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}. \tag{4.146}$$

This last equation together with (4.143) form **Hamilton's canonical equations**:

$$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}(\mathbf{p}, \mathbf{q}, t)}{\partial \mathbf{q}} + \mathbf{Q}, \tag{4.147}$$

$$\dot{\mathbf{q}} = \frac{\partial \mathcal{H}(\mathbf{p}, \mathbf{q}, t)}{\partial \mathbf{p}}, \tag{4.148}$$

where we have added a vector of generalised forces $\mathbf{Q}$.

The following procedure can be used to obtain a Hamiltonian model:

1. Choose a set of generalised coordinates q_i and form the Lagrangian $\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})$,
2. Define the momenta $p_i = \partial \mathcal{L}/\partial \dot{q}_i$.
3. Write the Hamiltonian Using the Legendre transform: $\mathcal{H}(\mathbf{p}, \mathbf{q}) = \mathbf{p}^{\mathsf{T}}\dot{\mathbf{q}} - \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})$,
4. Use Hamilton's equations (4.147) and (4.148) to obtain the equations of motion.

For mechanical systems, step 3, above, can be done by simply relating the velocities $\dot{q}_i$ in the Lagrangian to p_i and changing the sign of the potential energy—this is consequence of the fact that the kinetic energy and kinetic co-energy have the same numerical value for non-relativistic mechanics.

For mechanical systems, the Hamiltonian and the conjugate momentum take the following general form:

$$\mathcal{H}(\mathbf{p}, \mathbf{q}, t) = \frac{1}{2}\,\mathbf{p}^\mathsf{T}\mathbf{M}^{-1}(\mathbf{q})\,\mathbf{p} + \mathcal{V}(\mathbf{q}, t), \qquad \mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \mathbf{M}(\mathbf{q})\dot{\mathbf{q}}. \tag{4.149}$$

Example 17 (2-DOF Mechanical System) *Consider the system depicted in Figure 4.8. We will follow the procedure above to obtain a Hamiltonian model.*

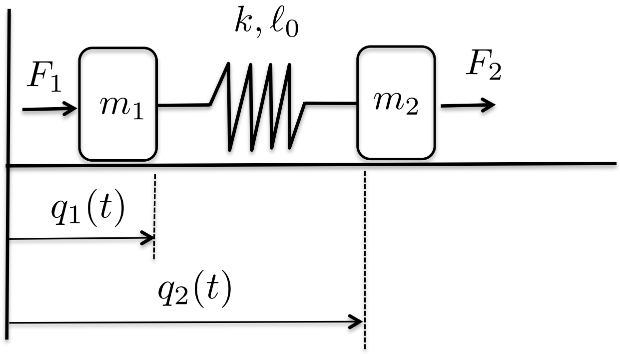


Figure 4.8: 2-DOF Mechanical System.

<u>*Step 1 - Lagrangian*</u>
We consider as generalised coordinates q_1 and q_2, which are the positions of the masses m_1 and m_2 respectively. Then we can form the following Lagrangian:

$$\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2}m_1\dot{q}_1^2 + \frac{1}{2}m_2\dot{q}_2^2 - \frac{1}{2}K(q_2 - q_1 - \ell_0)^2, \tag{4.150}$$

or alternatively we can use

$$L(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2}\dot{\mathbf{q}}^\mathsf{T}\mathbf{M}(\mathbf{q})\,\dot{\mathbf{q}} - \frac{1}{2}K(q_2 - q_1 - \ell_0)^2,$$

where in this case $\mathbf{M}(\mathbf{q})$ is independent of $\mathbf{q}$:

$$\mathbf{M} \triangleq \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}.$$

<u>*Step 2 - Momenta*</u>
From the Lagrangian, we can obtain the conjugate momenta:

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \mathbf{M}\dot{\mathbf{q}}, \tag{4.151}$$

where we have made use of the fact that

$$\frac{\partial \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^\mathsf{T})\mathbf{x} = 2\mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ *does not depend on* $\mathbf{x}$*, and the last equality holds if* $\mathbf{A}$ *is symmetric.*

In component form, (4.151) *gives*

$$p_1 = \frac{\partial \mathcal{L}}{\partial q_1} = m_1 \dot{q}_1, \quad p_2 = \frac{\partial \mathcal{L}}{\partial q_2} = m_2 \dot{q}_2.$$

Step 3 - Hamiltonian
The Legendre transformation leads to

$$\begin{aligned}\mathcal{H}(\mathbf{p}, \mathbf{q}) &= \mathbf{p}^\mathsf{T} \dot{\mathbf{q}} - \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}), \\ &= \mathbf{p}^\mathsf{T} \mathbf{M}^{-1} \mathbf{p} - \mathcal{L}(\mathbf{M}^{-1} \mathbf{p}, \mathbf{q}).\end{aligned}$$

The latter can be expressed as

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \mathbf{p}^\mathsf{T} \mathbf{M}^{-1} \mathbf{p} - \frac{1}{2} \mathbf{p}^\mathsf{T} \mathbf{M}^{-\mathsf{T}} \mathbf{M} \mathbf{M}^{-1} \mathbf{p} + \frac{1}{2} K (q_2 - q_1 - \ell_0)^2,$$

Since the mass matrix is always symmetric $\mathbf{M} = \mathbf{M}^\mathsf{T}$*, the latter becomes*

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \frac{1}{2} \mathbf{p}^\mathsf{T} \mathbf{M}^{-1} \mathbf{p} + \frac{1}{2} K (q_2 - q_1 - \ell_0)^2, \tag{4.152}$$

$$= \frac{1}{2m_1} p_1^2 + \frac{1}{2m_2} p_2^2 + \frac{1}{2} K (q_2 - q_1 - \ell_0)^2. \tag{4.153}$$

Note that the latter satisfies the general form (4.149)*.*

Step 4 - Hamilton Equations
Direct application of (4.147) *and* (4.148) *leads to*

$$\dot{p}_1 = K(q_2 - q_1 - \ell_0) + F_1, \tag{4.154}$$

$$\dot{p}_2 = -K(q_2 - q_1 - \ell_0) + F_2, \tag{4.155}$$

$$\dot{q}_1 = m_1^{-1} p_1, \tag{4.156}$$

$$\dot{q}_2 = m_2^{-1} p_2. \tag{4.157}$$

Does this model look familiar? It should. All the models we formulated in MCHA2000 were Hamiltonian models (with additional dissipative terms). ■

4.12 Energy Conservation

The time derivative of the Hamiltonian is

$$\frac{d\mathcal{H}}{dt} = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\right)^\mathsf{T} \dot{\mathbf{p}} + \left(\frac{\partial \mathcal{H}}{\partial \mathbf{q}}\right)^\mathsf{T} \dot{\mathbf{q}} + \frac{\partial \mathcal{H}}{\partial t}. \tag{4.158}$$

Using (4.147) and (4.148),

$$\frac{dH}{dt} = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\right)^\mathsf{T} \left[\mathbf{Q} - \frac{\partial \mathcal{H}}{\partial \mathbf{q}}\right] + \left(\frac{\partial \mathcal{H}}{\partial \mathbf{q}}\right)^\mathsf{T} \frac{\partial \mathcal{H}}{\partial \mathbf{p}} + \frac{\partial \mathcal{H}}{\partial t}, \tag{4.159}$$

$$= \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\right)^\mathsf{T} \mathbf{Q} - \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\right)^\mathsf{T} \frac{\partial \mathcal{H}}{\partial \mathbf{q}} + \left(\frac{\partial \mathcal{H}}{\partial \mathbf{q}}\right)^\mathsf{T} \frac{\partial \mathcal{H}}{\partial \mathbf{p}} + \frac{\partial \mathcal{H}}{\partial t}, \tag{4.160}$$

and finally,

$$\frac{d\mathcal{H}}{dt} = \dot{\mathbf{q}}^\mathsf{T}\,\mathbf{Q} + \frac{\partial \mathcal{H}}{\partial t}. \tag{4.161}$$

If the Hamiltonian is time independent, then

$$\frac{d\mathcal{H}}{dt} = \dot{\mathbf{q}}^\mathsf{T}\,\mathbf{Q}. \tag{4.162}$$

The latter has units of power, and it establishes that the rate at which the external forces injects or extract energy (right-hand side) is equal to the rate of change of the energy in the system (left-hand side).

A system for which (4.162) holds is called a **conservative system**. This name derives from the fact that when $\mathbf{Q} = \mathbf{0}$, the energy is conserved:

$$\frac{d\mathcal{H}}{dt} = 0. \tag{4.163}$$

This result is also known as the **Energy Theorem of Conservative Systems**:

$$\frac{d\mathcal{H}}{dt} = 0 \qquad \Leftrightarrow \qquad \mathcal{H}(\mathbf{p}_1, \mathbf{q}_1) = \mathcal{H}(\mathbf{p}_2, \mathbf{q}_2). \tag{4.164}$$

4.13 Port-Hamiltonian Systems

Port-Hamiltonian systems (PHS) are generalisations of Hamiltonian models of the form (4.147)-(4.148) in which an output variable is chosen such that the input-output variables form a port through which energy flows into the system. These models have gained a significant attention from the control community in the latter years, for they provide a basis for simple robust passivity-based (energy-based) control design. The basic idea is to design a controller so that the stability of the desired equilibrium point depends only on dissipativity properties of the system which arise from using the energy as a Lyapunov function (Ortega et al., 2002).

The Hamiltonian model (4.147)-(4.148) can be written in matrix form as follows

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \\ \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \mathbf{Q}. \tag{4.165}$$

If we define the vectors

$$\mathbf{x} \triangleq \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}, \qquad \mathbf{u} \triangleq \mathbf{Q}, \tag{4.166}$$

then the model (4.165) can be written as

$$\dot{\mathbf{x}} = \mathbf{J}\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}} + \mathbf{g}\,\mathbf{u}. \tag{4.167}$$

If, in particular, we choose the output variable

$$\mathbf{y} = \mathbf{g}^{\mathsf{T}} \frac{\partial \mathcal{H}}{\partial \mathbf{x}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{q}}, \tag{4.168}$$

then the input-output variables form a port through which energy flows into the system. The model (4.167)-(4.168) is called a **Port-controlled Hamiltonian System** (van der Schaft, 2000).

The derivative of the energy stored in the system is

$$\frac{d\mathcal{H}}{dt} = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \dot{\mathbf{x}} = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \mathbf{J} \frac{\partial \mathcal{H}}{\partial \mathbf{x}} + \left(\frac{\partial H}{\partial \mathbf{x}}\right)^{\mathsf{T}} \mathbf{g}\,\mathbf{u} = \mathbf{y}^{\mathsf{T}}(t')\mathbf{u}(t'), \tag{4.169}$$

where we have made use of the fact that $\mathbf{J}$ is skew-symmetric, and therefore its quadratic form ($\mathbf{z}^{\mathsf{T}}\mathbf{J}\mathbf{z}$) is zero. The latter can be integrated to

$$\mathcal{H}(t) = \mathcal{H}(0) + \int_0^t \mathbf{y}^{\mathsf{T}}(t')\mathbf{u}(t')\,dt', \tag{4.170}$$

which establishes that the total energy in the system equals the initial energy plus the energy supplied through the input-output port. This illustrated in Figure 4.9 using the bond-graph language. In the absence of input, that is $\mathbf{u} = \mathbf{0}$, the energy is constant, and therefore, non-increasing, which shows that the point $\mathbf{x}^*$, at which the energy is minimum is Lyapunov stable. Furthermore, since the potential energy is always bounded from below, we can assume, without loss of generality, that $\mathcal{H} \geq 0$, then it follows

$$\int_0^t \mathbf{y}^{\mathsf{T}}(t')\mathbf{u}(t')\,dt' \geq -\mathcal{H}(0), \tag{4.171}$$

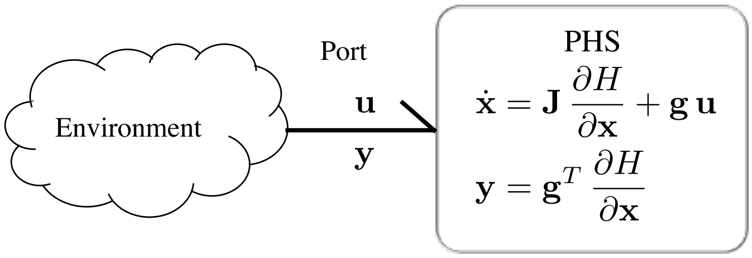
which satisfies the definition of a **passive system**, that is a system that cannot generate energy (Khalil, 2000; Brogliato et al., 2007). The importance of passive systems lies in the fact that the feedback interconnection of passive systems is passive, and under some observability conditions stable. Therefore, a passive system can be interconnected with a passive controller to give a passive and under some conditions stable closed-loop system—see, for example, Brogliato et al. (2007).

Let us consider an extension of the model (4.167), which is called a **Port-controlled Hamiltonian System with Dissipation (PCHD)** (van der Schaft, 2000):

$$\dot{\mathbf{x}} = [\mathbf{J}(\mathbf{x}) - \mathbf{R}(\mathbf{x})]\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}} + \mathbf{g}(\mathbf{x})\,\mathbf{u}, \tag{4.172}$$

$$\mathbf{y} = \mathbf{g}^{\mathsf{T}}(\mathbf{x})\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}}, \tag{4.173}$$

where the skew-symmetric matrix $\mathbf{J}(\mathbf{x}) = -\mathbf{J}^{\mathsf{T}}(\mathbf{x})$ describes the **interconnection** of the energy storing components in system and $\mathbf{R}(\mathbf{x}) \geq \mathbf{0}$ describes the **dissipation**.



$$\underbrace{H(t) - H(0)}_{\text{Energy increment}} = \underbrace{\int_0^t \mathbf{y}^T(t')\mathbf{u}(t')\,dt'}_{\text{Energy Supplied}}$$

Figure 4.9: The the input-output variables of a Port-Hamiltonian system (PHS) form a port through which energy flows into the system.

Following a similar procedure as above, we find for this system that

$$\underbrace{\frac{d\mathcal{H}}{dt}}_{\text{Power into the system}} = \left(\frac{\partial H}{\partial \mathbf{x}}\right)^{\mathsf{T}} \dot{\mathbf{x}} = \underbrace{\mathbf{y}^{\mathsf{T}}(t')\mathbf{u}(t')}_{\text{Supplied power}} - \underbrace{\left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \mathbf{R}(\mathbf{x})\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}}}_{\text{Dissipated power}}. \tag{4.174}$$

In the absence of input, that is $\mathbf{u} = \mathbf{0}$,

$$\frac{d\mathcal{H}}{dt} = -\left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \mathbf{R}(\mathbf{x})\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}} \leq \mathbf{0}, \tag{4.175}$$

which follows from the fact that $\mathbf{R}(\mathbf{x}) \geq \mathbf{0}$. This shows that the energy is non-increasing; and therefore, it shows that the point $\bar{\mathbf{x}}$, at which the energy is minimum is Lyapunov stable. Furthermore, consider the set

$$U = \left\{\mathbf{x} : \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^{\mathsf{T}} \mathbf{R}(\mathbf{x})\frac{\partial \mathcal{H}}{\partial \mathbf{x}} = \mathbf{0}\right\}.$$

If the largest invariant set in U under (4.172) with zero input is $\{\mathbf{x}^*\}$—the minimum of the energy—then it follows from the LaSalle's Invariance Principle [3] that $\mathbf{x}^*$ is a locally asymptotically stable equilibrium point.

A striking feature of the PCHD model (4.172)-(4.173) is that stability of the equilibrium point $\mathbf{x}^*$ at which the energy attain its minimum is guaranteed provided that

1. $\mathcal{H}(\mathbf{x})$ is bounded from below,
2. $\mathbf{J}(\mathbf{x})$ is skew-symmetric,
3. $\mathbf{R}(\mathbf{x})$ is positive semi definite.

[3] See Khalil (2000).

Therefore, if we can shape a model into this form and satisfy these conditions we can guarantee stability. This is a very attractive feature for control design, where we can try to shape the closed-loop system into a PCHD with no input.

In addition, (4.175) shows that

$$\mathcal{H}(t) - \mathcal{H}(0) \leq \int_0^t \mathbf{y}^\mathsf{T}(t')\mathbf{u}(t')\,dt', \tag{4.176}$$

which describes the passivity of the system with input $\mathbf{u}$ and output $\mathbf{y}$. This property can be used to assess stability of the interconnection of the system with other systems.

Example 18 (PHS Model of a 2-DOF Mechanical System) *Let us re-consider Example 17, but assume now that there is linear viscous friction between the masses and the floor with coefficients d_1 and d_2 respectively. In addition, assume that there is a damper between the two masses with a coefficient d_{12}.*

The damping forces can be expressed as

$$\mathbf{F}^d = -\mathbf{D}\,\dot{\mathbf{q}} \quad \Leftrightarrow \quad \begin{bmatrix} F_{d1} \\ F_{d2} \end{bmatrix} = -\begin{bmatrix} d_1 + d_{12} & -d_{12} \\ -d_{12} & d_2 + d_{12} \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix},$$

Using the Hamiltonian (4.152), the damping forces can alternatively be expressed as

$$\mathbf{F}^d = -\mathbf{D}\,\dot{\mathbf{q}} = -\mathbf{D}\frac{\partial \mathcal{H}}{\partial \mathbf{p}},$$

which shows that Hamilton equations can be written as

$$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} - \mathbf{D}\frac{\partial \mathcal{H}}{\partial \mathbf{p}} + \mathbf{F}, \tag{4.177}$$

$$\dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}. \tag{4.178}$$

We can alternatively write the latter as

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \left(\underbrace{\begin{bmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\mathbf{J}} - \underbrace{\begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{R}} \right) \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \\ \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \mathbf{F}. \tag{4.179}$$

Since $\mathbf{D} > 0$, it follows that $\mathbf{R} \geq 0$. In the absence of excitation forces, the equilibrium points $p_1 = p_2 = 0$ and $q_1, q_2 : q_2 - q_1 = \ell_0$ are Lyapunov stable.

In component form, the state-space equations (4.179) become

$$\dot{p}_1 = \quad K(q_2 - q_1 - \ell_0) - \frac{(d_1 + d_{12})}{m_1}p_1 + \frac{d_{12}}{m_2}p_2 + F_1, \tag{4.180}$$

$$\dot{p}_2 = -K(q_2 - q_1 - \ell_0) + \frac{d_{12}}{m_1}p_1 - \frac{(d_2 + d_{12})}{m_2}p_2 + F_2, \tag{4.181}$$

$$\dot{q}_1 = m_1^{-1}p_1, \tag{4.182}$$

$$\dot{q}_2 = m_2^{-1}p_2. \tag{4.183}$$

■

4.14 Control of Port-Hamiltonian Systems

As discussed in the previous section, port-controlled Hamiltonian systems (PCH) of the form (4.172)-(4.173) explicitly incorporate a function of the total energy stored in the system (Hamiltonian) and functions that describe interconnection and dissipation structure of the system. Conditions of symmetry, positiveness, and boundedness on the various functions lead to stability of the state at which the energy attains its minimum. This is a very attractive feature for control design: if we can design a controller such that the closed-loop system can be put into a PHS form and the desired equilibrium point is the point that minimises the closed-loop energy, then we can guarantee stability of the equilibrium point. Moreover, the control system will be stable even if there is model uncertainty (parameters or even model structure) provided that the closed loop PCH form is preserved. This gives robustness to model uncertainly.

The objective, it to achieve a **desired** closed-loop system of the form

$$\dot{\mathbf{x}} = [\mathbf{J}_d(\mathbf{x}) - \mathbf{R}_d(\mathbf{x})]\,\frac{\partial \mathcal{H}_d}{\partial \mathbf{x}}, \tag{4.184}$$

where $\mathbf{J}_d(\mathbf{x}) = -\mathbf{J}_d^{\mathsf{T}}(\mathbf{x})$, $\mathbf{R}_d(\mathbf{x}) \geq 0$, and the equilibrium point $\mathbf{x}^*$ minimises $\mathcal{H}_d(\mathbf{x})$.

The necessary and sufficient conditions for $\mathbf{x}^*$ to be a **strict local minimiser**[a] of $\mathcal{H}_d(\mathbf{x})$ are that the gradient vanishes and the Hessian is positive definite at $\mathbf{x}^*$ (Nocedal and Wright, 2006):

$$\frac{\partial \mathcal{H}_d}{\partial \mathbf{x}}[\mathbf{x}^*] = \mathbf{0}, \qquad \frac{\partial^2 \mathcal{H}_d}{\partial \mathbf{x}^2}[\mathbf{x}^*] > 0. \tag{4.185}$$

Furthermore, when $\mathcal{H}_d(\mathbf{x})$ is convex, that is

$$\frac{\partial^2 \mathcal{H}_d}{\partial \mathbf{x}^2} > 0, \quad \forall \mathbf{x}, \tag{4.186}$$

then any local $\mathbf{x}^*$ is a global minimiser. If $\mathcal{H}_d(\mathbf{x})$ is differentiable, then any stationary point $\mathbf{x}^*$ is a global minimiser of $\mathcal{H}_d(\mathbf{x})$.

[a] A point $\mathbf{x}^*$ is a strict local minimiser if there is a neighbourhood $\mathcal{N}$ of $\mathbf{x}^*$ such that $\mathcal{H}_d(\mathbf{x}^*) < \mathcal{H}_d(\mathbf{x})$ for all $\mathbf{x}$ in $\mathcal{N}$.

Ortega et al. (2002) provides a general methodology to design a feedback control law $\mathbf{u} = \boldsymbol{\beta}(\mathbf{x})$ that can renders the open-loop PCH system

$$\dot{\mathbf{x}} = [\mathbf{J}(\mathbf{x}) - \mathbf{R}(\mathbf{x})]\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}} + \mathbf{g}(\mathbf{x})\,\mathbf{u} \tag{4.187}$$

into the closed loop system (4.184). This technique is known as **Interconnection and Damping Assignment Passivity-based Control (IDA-PBC)**. In such design, the controller modifies the interconnection of the system ($\mathbf{J} \rightarrow \mathbf{J}_d$), assigns damping ($\mathbf{R} \rightarrow \mathbf{R}_d$), and the passivity-based control refers to the fact that the controller re-shapes the energy ($H \to H_d$) so that the desired equilibrium point $\mathbf{x}^*$ minimises H_d.

The IDA-PBC design reduces to finding the feedback control law $\mathbf{u} = \boldsymbol{\beta}(\mathbf{x})$ that forces a matching of the dynamics of the open loop system (4.187) to that of the desired closed-loop system (4.184), in which $\mathbf{x}^*$ is a stable equilibrium point minimises $\mathcal{H}_d$. That is, IDA-PBC seeks the control law that solves the following **Matching Problem**:

$$\boldsymbol{\beta}(\mathbf{x}): \qquad [\mathbf{J}_d(\mathbf{x}) - \mathbf{R}_d(\mathbf{x})]\,\frac{\partial \mathcal{H}_d}{\partial \mathbf{x}} = [\mathbf{J}(\mathbf{x}) - \mathbf{R}(\mathbf{x})]\,\frac{\partial \mathcal{H}}{\partial \mathbf{x}} + \mathbf{g}(\mathbf{x})\,\boldsymbol{\beta}(\mathbf{x}), \tag{4.188}$$

with the constraint that

$$\mathbf{x}^* = \arg\min \mathcal{H}_d(\mathbf{x}^*). \tag{4.189}$$

There are different ways of solving the matching problem (4.188)-(4.188) depending on the type of system we have and the desired closed-loop system functions, $\mathbf{J}_d(\mathbf{x})$, $\mathbf{R}_d(\mathbf{x})$, and $H_d(\mathbf{x})$.

For a large class of mechanical systems, the PCHD model takes the following form:

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{D}(\mathbf{p},\mathbf{q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \\ \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \end{bmatrix} + \begin{bmatrix} \mathbf{G}(\mathbf{p},\mathbf{q}) \\ \mathbf{0} \end{bmatrix} \mathbf{Q}, \tag{4.190}$$

where $\mathbf{p} \in \mathbb{R}^n$, $\mathbf{q} \in \mathbb{R}^n$, the generalised forces $\mathbf{Q} \in \mathbb{R}^m$, and

$$\mathcal{H}(\mathbf{p},\mathbf{q}) = \frac{1}{2}\mathbf{p}^\mathsf{T}\mathbf{M}^{-1}(\mathbf{q})\,\mathbf{p} + V(\mathbf{q}). \tag{4.191}$$

The system (4.190) is

- **Fully actuated** if $m = n$ and $\mathbf{G}(\mathbf{p},\mathbf{q})$ is full rank,
- **Over actuated** if $m > n$ and $\mathbf{G}(\mathbf{p},\mathbf{q})$ is full rank,
- **Under actuated** if $m < n$ or $m \geq n$ and $\mathbf{G}(\mathbf{p},\mathbf{q})$ is rank deficient.

The above conditions establish that for a fully- and over-actuated systems, there is freedom to impose a direct control action in all the momenta. These systems are the easiest to control.

One of the simplest matching problems to solve for mechanical systems of the form (4.190) is that where the system is fully actuated, and impose the desired potential energy with a minimum at the desired equilibrium point $\mathbf{x}^*$ and we inject damping. For this case, the desired closed-loop system dynamics takes the following form:

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{D}_d(\mathbf{p},\mathbf{q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} \frac{\partial H_d}{\partial \mathbf{p}} \\ \frac{\partial H_d}{\partial \mathbf{q}} \end{bmatrix}, \tag{4.192}$$

where

$$\mathcal{H}_d(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\mathbf{p}^{\mathsf{T}}\mathbf{M}^{-1}(\mathbf{q})\,\mathbf{p} + \mathcal{V}_d(\mathbf{q}). \tag{4.193}$$

Then the matching problem of (4.190) and (4.192) is automatically satisfied for $\dot{\mathbf{q}}$, and the matching for $\dot{\mathbf{p}}$ leads to

$$-\frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} - \mathbf{D}_d(\mathbf{p}, \mathbf{q})\frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} - \mathbf{D}(\mathbf{p}, \mathbf{q})\frac{\partial \mathcal{H}}{\partial \mathbf{p}} + \mathbf{G}(\mathbf{p}, \mathbf{q})\mathbf{Q}. \tag{4.194}$$

Hence, the controller that achieves **potential-energy shaping and damping assignment for a fully actuated system** (4.190) and desired Hamiltonian (4.193) is

$$\mathbf{Q} = \mathbf{G}^{-1}(\mathbf{p}, \mathbf{q})\left[\frac{\partial \mathcal{H}}{\partial \mathbf{q}} - \frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} + [\mathbf{D}(\mathbf{p}, \mathbf{q}) - \mathbf{D}_d(\mathbf{p}, \mathbf{q})]\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\right], \tag{4.195}$$

which reduces to

$$\mathbf{Q} = \mathbf{G}^{-1}(\mathbf{p}, \mathbf{q})\left[\frac{\partial \mathcal{V}}{\partial \mathbf{q}} - \frac{\partial \mathcal{V}_d}{\partial \mathbf{q}} + [\mathbf{D}(\mathbf{p}, \mathbf{q}) - \mathbf{D}_d(\mathbf{p}, \mathbf{q})]\,\mathbf{M}^{-1}(\mathbf{q})\,\mathbf{p}\right], \tag{4.196}$$

where $\frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}$ since we are not changing the kinetic energy.

Note that (4.196) can be applied to most robotic manipulators for operations in which they can be considered fully actuated.

Example 19 (Energy-based control of a fully-actuated mechanical system) *Let us consider the mechanical system shown in Figure (4.8), with the assumption that there is friction viscous between the masses and the floor with coefficients b_1 and b_2 respectively, and between the two masses with a coefficient b_{12}.*

As shown in Example 18 , the system can be modelled as

$$\begin{bmatrix}\dot{\mathbf{p}}\\ \dot{\mathbf{q}}\end{bmatrix} = \left(\begin{bmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{bmatrix} - \begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\right)\begin{bmatrix}\frac{\partial \mathcal{H}}{\partial \mathbf{p}}\\ \frac{\partial \mathcal{H}}{\partial \mathbf{q}}\end{bmatrix} + \begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}\begin{bmatrix}F_1\\ F_2\end{bmatrix}, \tag{4.197}$$

where

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \frac{1}{2}[p_1, p_2]\begin{bmatrix}m_1^{-1} & 0\\ 0 & m_2^{-1}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\end{bmatrix} + \frac{K}{2}(q_2 - q_1 - \ell_0)^2, \tag{4.198}$$

and

$$\mathbf{D} = \begin{bmatrix}d_1 + d_{12} & -d_{12}\\ -d_{12} & d_2 + d_{12}\end{bmatrix}.$$

We would like to stabilise the equilibrium point

$$\mathbf{p}^* = \begin{bmatrix}0\\ 0\end{bmatrix}, \quad \mathbf{q}^* = \begin{bmatrix}q_1^*\\ q_2^*\end{bmatrix},$$

and at the same time impose the closed-loop damping

$$\mathbf{D}_d = \begin{bmatrix} d_1^d & 0 \\ 0 & d_2^d \end{bmatrix}.$$

To do this, we can define the following desired Hamiltonian

$$\mathcal{H}_d(\mathbf{p}, \mathbf{q}) = \frac{1}{2}[p_1, p_2] \begin{bmatrix} m_1^{-1} & 0 \\ 0 & m_2^{-1} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \frac{K_1}{2}(q_1 - q_1^*)^2 + \frac{K_2}{2}(q_2 - q_2^*)^2, \tag{4.199}$$

where $K_1 > 0$ and $K_2 > 0$ are the control parameters. Note that this Hamiltonian has a minimum at the desired equilibrium point. Then we seek the following closed-loop system:

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{D}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} \frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} \\ \frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} \end{bmatrix}. \tag{4.200}$$

The matching of the top rows of (4.197) and (4.200) leads to the following control forces:

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = -\frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} - \mathbf{D}_d \frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} + \frac{\partial \mathcal{H}}{\partial \mathbf{q}} + \mathbf{D} \frac{\partial \mathcal{H}}{\partial \mathbf{p}}. \tag{4.201}$$

Since we are only shaping the potential energy,

$$\frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}},$$

and thus, the control forces can be expressed as

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \frac{\partial \mathcal{H}}{\partial \mathbf{q}} - \frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} + (\mathbf{D} - \mathbf{D}_d) \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \tag{4.202}$$

which is the specialisation of (4.196) to this problem.

Computing the gradients, we can finally express the forces as a function of the state and the desired equilibrium point:

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} -K(q_2 - q_1 - \ell_0) \\ K(q_2 - q_1 - \ell_0) \end{bmatrix} - \begin{bmatrix} K_1(q_1 - q_1^*) \\ K_2(q_2 - q_2^*) \end{bmatrix} + (\mathbf{D} - \mathbf{D}_d) \begin{bmatrix} m_1^{-1} & 0 \\ 0 & m_2^{-1} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}. \tag{4.203}$$

From the imposed desired Hamiltonian, we can see that the controller implements two springs, which force the system to the desired equilibrium point, this provides a physical interpretation of the controller. These two springs show on the second term on the right-hand side of the controller above. The first term on the right-hand side is used to cancel the forces of the spring acting between the masses, which is part of the open-loop system. ■

As we can see, the method is constructive and easy to apply: we propose the desired Hamiltonian and then solve the matching problem to find the stabilising controller.

A more general form than (4.192) for a desired closed-loop mechanical system is

$$\begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{J}(\mathbf{p}, \mathbf{q}) & -\mathbf{M}_d(\mathbf{q})\mathbf{M}^{-1}(\mathbf{q}) \\ \mathbf{M}^{-1}(\mathbf{q})\mathbf{M}_d(\mathbf{q}) & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{D}_d(\mathbf{p}, \mathbf{q}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} \frac{\partial \mathcal{H}_d}{\partial \mathbf{p}} \\ \frac{\partial \mathcal{H}_d}{\partial \mathbf{q}} \end{bmatrix}, \tag{4.204}$$

where

$$\mathcal{H}_d(\mathbf{p},\mathbf{q}) = \frac{1}{2}\mathbf{p}^{\mathsf{T}}(\mathbf{M}_d(\mathbf{q}))^{-1}\,\mathbf{p} + \mathcal{V}_d(\mathbf{q}). \tag{4.205}$$

This form allows one not only to shape the potential energy and to inject damping, but also to shape the kinetic energy. For fully-actuated systems, this gives more freedom to shape the closed-loop dynamic response. If the system is under-actuated, then the matching problem may not be solved algebraically in general. That is, one is restricted in choice of potential energy that can chosen, and for some problems, potential energy alone alone cannot stabilise the desired equilibrium point.

Example 20 (Potential-energy Shaping of Under-actuated System) *As an example of a restricted potential energy shaping, consider the system in Example 19, but assume that we only have F_1 as a control force. In this case the open-loop model (4.197) reduces to*

$$\begin{bmatrix}\dot{\mathbf{p}}\\ \dot{\mathbf{q}}\end{bmatrix} = \left(\begin{bmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{bmatrix} - \begin{bmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\right)\begin{bmatrix}\frac{\partial\mathcal{H}}{\partial\mathbf{p}}\\ \frac{\partial\mathcal{H}}{\partial\mathbf{q}}\end{bmatrix} + \begin{bmatrix}[1,0]^{\mathsf{T}}\\ \mathbf{0}\end{bmatrix} F_1. \tag{4.206}$$

For this system the matching can be solved algebraically for

$$\mathcal{H}_d(\mathbf{p},\mathbf{q}) = \frac{1}{2}[p_1,p_2]\begin{bmatrix}m_1^{-1} & 0\\ 0 & m_2^{-1}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\end{bmatrix} + \frac{1}{2}K_1(q_1-q_1*)^2 + \frac{1}{2}K(q_2-q_1-\ell_0)^2, \tag{4.207}$$

which has a minimum at

$$\mathbf{p}^* = \begin{bmatrix}0\\0\end{bmatrix}, \quad \mathbf{q}^* = \begin{bmatrix}q_1^*\\ q_1^*+\ell_0\end{bmatrix},$$

and

$$F_1 = -K_1(q_1 - q_1^*).$$

■

For more general cases of under-actuated systems, we can follow the IDA-PBC formulation presented in Ortega et al. (2002), which requires the solution of a partial differential equation (PDE). Indeed, if we consider the matching problem (4.188) and multiply on the left by the left-annihilator $\mathbf{g}^{\perp}(\mathbf{x})$, that is $\mathbf{g}^{\perp}(\mathbf{x})\mathbf{g}(\mathbf{x}) = \mathbf{0}$, we obtain

$$\mathbf{g}^{\perp}(\mathbf{x})[\mathbf{J}_d(\mathbf{x}) - \mathbf{R}_d(\mathbf{x})]\frac{\partial\mathcal{H}_d}{\partial\mathbf{x}} - \mathbf{g}^{\perp}(\mathbf{x})[\mathbf{J}(\mathbf{x}) - \mathbf{R}(\mathbf{x})]\frac{\partial\mathcal{H}}{\partial\mathbf{x}} = \mathbf{0}. \tag{4.208}$$

This is a PDE with unknowns $\mathcal{H}_d(\mathbf{x})$, $\mathbf{J}_d(\mathbf{x})$ and $\mathbf{R}_d(\mathbf{x})$. Here, we can, for example, choose $\mathbf{J}_d(\mathbf{x})$ and $\mathbf{R}_d(\mathbf{x})$ and solve for $H_d(\mathbf{x})$ and pick a solution that satisfies (4.189). Then, the control law is obtained from (4.188):

$$\boldsymbol{\beta}(\mathbf{x}) = [\mathbf{g}^{\mathsf{T}}(\mathbf{x})\mathbf{g}(\mathbf{x})]^{-1}\mathbf{g}^{\mathsf{T}}(\mathbf{x})\left([\mathbf{J}_d(\mathbf{x}) - \mathbf{D}_d(\mathbf{x})]\frac{\partial\mathcal{H}_d}{\partial\mathbf{x}} - [\mathbf{J}(\mathbf{x}) - \mathbf{D}(\mathbf{x})]\frac{\partial\mathcal{H}}{\partial\mathbf{x}}\right). \tag{4.209}$$

The PDE (4.208) is usually hard to solve. For mechanical systems, we can set the structure of H_d, which leads to a simpler PDE. For further details on IDA-PBC see the survey of Ortega and García-Canesco (2004).

4.15 Canonical Transformations and Optimal Control

We would like to finish this chapter, with a brief comment on the link between optimal control and the standard results of analytical mechanics. As mentioned at the beginning of Section 4.11, the Hamiltonian model (4.147)-(4.148) proved to be a significant step in the development of mechanics and then other sciences. Jacobi (1804-1851) called Hamilton equations **Canonical Equations** and developed the theory of canonical transformations, which, allows one to make a change of coordinates from $\mathbf{p}, \mathbf{q}$ to $\mathbf{p}', \mathbf{q}'$ whilst preserving the form of Hamilton's Equations, and solve the equations of motion without integrating them. Indeed, if the system is conservative, the original Hamiltonian $H(\mathbf{p}, \mathbf{q})$ can be transformed into one that equals one of the new configuration variables $K(\mathbf{p}', \mathbf{q}') = q'_k$, and if the system is rheonomic—that is, time-varying $\mathcal{H}(\mathbf{p}, \mathbf{q}, t)$—the transformation that leads to a new Hamiltonian that is identically zero, $\mathcal{K}(\mathbf{p}', \mathbf{q}', t) = 0$, in which case, Hamilton's equations take the form

$$\dot{\mathbf{q}}' = \frac{\partial \mathcal{K}(\mathbf{p}', \mathbf{q}', t)}{\partial \mathbf{p}'} = 0, \tag{4.210}$$

$$\dot{\mathbf{p}}' = -\frac{\partial \mathcal{K}(\mathbf{p}', \mathbf{q}', t)}{\partial \mathbf{q}'} = 0. \tag{4.211}$$

These lead to constant $\mathbf{p}'$ and $\mathbf{q}'$. In either case, the solution of the equations is not obtained by integrating (4.147)-(4.148) but by using the inverse transformation from $\mathbf{p}', \mathbf{q}'$ back to the original variables $\mathbf{p}, \mathbf{q}$.

All canonical transformations can be generated from a single scalar function $S(\mathbf{q}, t)$ called the generating function, which satisfies the following partial differential equation known as the **Hamilton-Jacobi Equation**:

$$\frac{dS(\mathbf{q}, t)}{dt} + \mathcal{H}(\mathbf{q}, \frac{\partial S(\mathbf{q}, t)}{\partial \mathbf{q}}, t) = \mathbf{0}. \tag{4.212}$$

The solution of this equation takes the form

$$S(\mathbf{q}, \alpha_1, \alpha_2, \ldots, \alpha_n, t), \tag{4.213}$$

where the α_i are constants of integration. The evolution of the system, and thus the solution of the motion problem, can then be expressed as

$$q_i = f_i(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n, t), \tag{4.214}$$

where the constants $\beta_i = \partial S / \partial \alpha_i$.

The Hamilton-Jacobi equation (4.212) was extended by Bellman (1957) in the context of optimal control theory. When specialised to this problem, equation (4.212) takes a form that is known as the **Hamilton-Jacobi-Bellman equation**. The solution of this equation, which can only be found for particular cases, gives the optimal cost to go from the current state to the final state of the system, and this leads to the optimal control law. When the optimal control problem is specialised to the case of linear systems with a quadratic cost (LQR), the Hamilton-Jacobi-Bellman equation takes a form known as the Riccati equation.

The unconstrained **optimal control problem** seeks to find the control law $\mathbf{u}^*(t)$, for the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t),$$

which minimises the following cost:

$$J(\mathbf{x}_0, t) = g(\mathbf{x}(t_f), t) + \int_{t_0}^{t_f} V(\mathbf{x}, \mathbf{u}, t)\, dt,$$

where g, V are positive definite functions and the system has initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$.

The optimal cost $J^*(\mathbf{x}, t)$ satisfies the following partial differential equation known as the Hamilton-Jacobi-Bellman equation:

$$\frac{dJ^*(\mathbf{x}, t)}{dt} + \min_{\mathbf{u}} \left\{ V(\mathbf{x}, \mathbf{u}, t) + \left[\frac{\partial J^*(\mathbf{x}, t)}{\partial \mathbf{x}} \right]^T \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\} = \mathbf{0},$$

with the boundary condition $J^*(\mathbf{x}(t_f), t_f) = g(\mathbf{x}(t_f), t)$. The optimal control law can then be expressed as function of the optimal cost and the initial state:

$$\mathbf{u}^*(t) = \boldsymbol{\beta}(J^*(\mathbf{x}(t), t), \mathbf{x}(t), t).$$

On reflection, this connection to optimal control should not come entirely as a surprise, for Hamilton's principle (4.127) establishes that the evolution of mechanical systems in nature is such that the action, which is the time integral of the Lagrangian is minimised. Optimal control theory, on the other hand, seeks the control law, $\mathbf{u}^*(t)$, that minimses a cost functional that represents deviations from a desired performance, and this functional is an integral.

For details about Jacobi's theory of canonical transformations and the Hamilton-Jacobi equation see Lanczos (1970) and references therein, and for an introduction to variational principles and optimal control see, for example, Naidu (2003) and Lee and Markus (1967).

Bibliography

Bellman, R.E. (1957). *Dynamic Programming.* Princeton University, NJ.

Brogliato, B., Lozano, R., Maschke, B., and Egeland, O. (2007). *Dissipative system analysis and control.* Springer-Verlag, London.

Egeland, O. and Gravdahl, J. (2002). *Modeling and Simulation for Automatic Control.* Marine Cybernetics, Trondheim.

Farrel, J. and Barth, M. (1999). *The Global Positioning System and Inertial Navigation.* McGraw-Hill.

Greenwood, D. (2003). *Advanced Dynamics.* Cambridge University Press.

Gregory, R.D. (ed.) (2006). *Classical Mechanics.* Cambridge University Press.

Hofmann-Wellenhof, B., Legat, K., and Weisser, M. (2003). *Naigation - Principles of Positioning and Guidance.* Springer Wien New York.

Kane, T. and Levinson, D. (1985). *Dynamics, Theory and Applications.* Mc-Graw-Hill Series in Mechanical Engineering. McGraw-Hill.

Khalil, H. (2000). *Nonlinear Systems.* Prentice-Hall, New Jersey.

Lanczos, C. (1970). *The Variational Principles of Mechanics.* Univ. of Toronto Press.

Lee, E. and Markus, L. (1967). *Foundations of optimal control theory.* SIAM series in applied mathematics. Wiley.

Marsden, J.E. and Ratiu, T.S. (2010). *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems.* Texts in Applied Mathematics. Springer.

Naidu, D.S. (2003). *Optimal Control Systems.* CRC Press.

Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization.* Springer.

Ortega, R. and García-Canesco, E. (2004). Interconnection and damping assignment passivity-based control: A survey. *European Journal of Control*, 10, 432–450.

Ortega, R., van der Schaft, A., Maschke, B., and Escobar, G. (2002). Interconnection and damping assignment passivity-based control of port-controlled hamiltonian systems. *Automatica*, 38(4), 585–596.

Perez, T., Mak, T., and T, A. (2007). Validation of a 4-dof manoeuvring model of a high-speed vehicle- passenger trimaran. In *9th International Conference on Fast Sea Transportation (FAST)*. Shangai, China.

Rao, A. (2006). *Dynamics of Particles and Rigid Bodies*. Cambridge University Press.

Rogers, R. (2007). *Applied Mathematics and Integrated Navigation Systems*. American Institute of Aeronautic and Astronautics AIAA.

Spong, M., Hutchinson, S., and Vidyasagar, M. (2006). *Robot Modeling and Control*. Wiley.

Stevens, B. and Lewis, F. (2003). *Aircraft Control and Simulation*. Wiley, second edition.

Strang, G. (1988). *Linear Algebra and Its Applications*. Brooks/Cole, 3 edition.

van der Schaft, A. (2000). *L2-gain and passivity techniques in nonlinear control*. Springer-Verlag, London.

Zipfel, P. (2007). *Modeling and Simulation of Aerospace Vehicle Dynamics*. American Institure of Aeronautics and Astronautics (AIAA).