ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

On What Riemannian Geometry Has to Say about Geometric Mechanics

Bachelor Thesis

Nicolás Bachmann Llauger

10th of July, 2023

Advisor: Prof. Dr. A. Cannas da Silva
Department of Mathematics, ETH Zürich

Abstract

This thesis is an elementary introduction to formulating classical mechanics in the language of Riemannian geometry. The four areas of focus are general mechanical systems, systems with holonomic constraints, rigid bodies and systems with non-holonomic constraints. The treatment of these four areas provides the necessary tools to solve an array of prominent problems in classical mechanics within the Riemannian framework.

Introduction

The study of the laws governing motion can trace its roots to Ancient Greece, with Aristotle's *Physics* constituting an early attempt. Also Archimedes and Hero can be highlighted as contributors to early mechanics.

The development of proper classical mechanics, however, began only over fifteen centuries later with the works of Galileo, Kepler and Descartes. Their developments paved the way for Newton's *Philosophiae Naturalis Principia Mathematica*, a seminal work that laid the foundation of what is now known as classical mechanics. After Newton, progressive reformulations and extensions of his theory managed to generalise it ever more, including Euler's study of rigid bodies, the Lagrangian reformulation using calculus of variations and the subsequent Hamiltonian reinterpretation of Lagrangian mechanics.

This thesis presents a formulation in the language of Riemannian geometry of Newtonian mechanics and Euler's expansion of the theory to rigid bodies. The beauty of this formulation lies in observing that all phenomena of classical mechanics can be understood to be objects of the theory of Riemannian geometry, wherein differential geometric concepts can be applied to describe the motion of systems of particles. Readers are assumed to have knowledge of linear algebra, fundamental differential equations, differential geometry and basic Riemannian geometry.

The thesis is divided in four chapters that have a similar structure: first the theory is developed and then it is applied to a particular or special case. The first chapter introduces the basic definitions to study mechanics using Riemannian geometry and solves the Kepler problem. The second chapter considers constraints imposed on the position of particles of a system and analyses the double pendulum as an instance of this phenomenon. Chapter three turns to rigid bodies and the Euler equations governing their motion in the absence of external forces. In chapter four constraints on the possible directions of motion are contemplated, another type of restriction that can be imposed on a system of particles, and the motion of an ice skate is studied

as an example thereof.
The present work relies heavily on chapter 5 of the book *An Introduction to Riemannian Geometry* by Leonor Godinho and José Natário. Accordingly, the notation tries to be consistent with that of the book.

Acknowledgements

I wish to extend my deepest gratitude to Prof. Dr. Ana Cannas da Silva for her support throughout this thesis. Her kindness is as profuse as her expertise, and she has regularly made time to answer my questions, always with helpful and valuable insights. I am very happy to have had Ana as my advisor, and I wish her great happiness.

Contents

Chapter 1

Mechanical Systems

In this chapter we introduce the first notions essential to studying the mechanics of a given system (of particles, for example) within the Riemannian framework. Once we have the necessary tools, we will apply them to an example: the Kepler problem, a special case of the two-body problem.

1.1 Basic Definitions

In order to start describing the motions (or mechanics) of a system of particles in a given ambient space (usually $\mathbb{R}^n$ for some $n > 0$), we need some form of description of whatever is guiding these motions. This is where **Newton's Second Law** comes in, which, for $x : (a, b) \subseteq \mathbb{R} \to \mathbb{R}^n$ the position of a particle in $\mathbb{R}^n$, stipulates that

$$m\ddot{x} = F(x, \dot{x}),$$

where F is an external force acting upon the particle. This is a second order ODE, and it is the starting point of the derivation of the equation of motion of the particle, i.e. the study of the mechanics of the particle.
We would therefore like to generalise Newton's Second Law and think about it using the language of Riemannian Geometry. To this end, we introduce the following definitions.

Definition 1.1 (Mechanical system) *A* ***mechanical system*** *is a triple* $(M, \langle \cdot, \cdot \rangle, \mathcal{F})$*, where:*

1. *M is a differentiable manifold, called the* ***configuration space****;*
2. *$\langle \cdot, \cdot \rangle$ is a Riemannian metric on M yielding the* ***mass operator*** *$\mu : TM \to T^*M$, defined by*

$$\mu(v)(w) = \langle v, w \rangle$$

for all $v, w \in T_pM$ and $p \in M$;

3. *$\mathcal{F} : TM \to T^*M$ is a differentiable map satisfying $\mathcal{F}(T_pM) \subseteq T_p^*M$ for all $p \in M$, called the **external force**.*

*A **motion** of the mechanical system is a solution $c : (a,b) \subseteq \mathbb{R} \to M$ of the **Newton equation***

$$\mu\left(\frac{D\dot{c}}{dt}\right) = \mathcal{F}(\dot{c}) \tag{1.1}$$

Remark 1.2 *In the language of Riemannian geometry, whenever the mass of a system needs to be considered, as in the Newton equation, the mass operator μ is used.*

Remark 1.3 *Note that for $(M, \langle\cdot,\cdot\rangle)$ a Riemannian manifold and $c : (a,b) \subseteq \mathbb{R} \to M$ a geodesic on the manifold, we have that*

$$\frac{D\dot{c}}{dt} = 0$$
$$\Leftrightarrow \mu\left(\frac{D\dot{c}}{dt}\right) = 0.$$

*Therefore, we can think of the geodesics of $(M, \langle\cdot,\cdot\rangle)$ as being the motions of the mechanical system which has configuration space M, mass operator $\langle\cdot,\cdot\rangle$ and vanishing external force $\mathcal{F} = 0$, i.e. $(M, \langle\cdot,\cdot\rangle, 0)$. And this is the mechanical system that describes a **free particle** on M.*

Definition 1.4 *Let $(M, \langle\cdot,\cdot\rangle, \mathcal{F})$ be a mechanical system. The external force $\mathcal{F}$ is said to be:*

1. ***positional** if $\mathcal{F}(v)$ depends only on $\pi(v)$, where $\pi : TM \to M$ is the natural projection;*
2. ***conservative** if there exists $U : M \to \mathbb{R}$ such that $\mathcal{F}(v) = -(dU)_{\pi(v)}$ for all $v \in TM$ (the function U is called a **potential energy**).*

*A mechanical system whose exterior force is conservative is called a **conservative mechanical system**.*

To conclude the section of basic definitions, we introduce the generalised notion of the kinetic energy. For a particle moving in $\mathbb{R}^n$, it is a function of its mass and its velocity in $\mathbb{R}^n$. The generalisation is a function on the tangent bundle, and, as in the Newton equation, the mass component comes into the equation via the mass operator μ. This generalised kinetic energy will prove essential in the computation of the equations of motion of mechanical systems.

Definition 1.5 *Let $(M, \langle\cdot,\cdot\rangle, \mathcal{F})$ be a mechanical system. The* ***kinetic energy*** *is the differentiable map $K : TM \to \mathbb{R}$ given by*

$$K(v) = \frac{1}{2}\langle v, v\rangle$$

for all $v \in TM$.

1.2 The Newton Equation in Local Coordinates

As is commonly the case with differentiable manifolds, it is convenient to work with local coordinates. In particular, we are interested in expressing the Newton equation in local coordinates, as it is the starting point for our derivation of the equations of motion.

Proposition 1.6 *Let $(M, \langle\cdot,\cdot\rangle, \mathcal{F})$ be a mechanical system. If $(x^1, ..., x^n)$ are local coordinates on M and $(x^1, ..., x^n, v^1, ..., v^n)$ are the local coordinates induced on TM, then*

$$\mu\left(\frac{D\dot{c}}{dt}(t)\right) = \sum_{i=1}^{n}\left[\frac{d}{dt}\left(\frac{\partial K}{\partial v^i}(x(t), \dot{x}(t))\right) - \frac{\partial K}{\partial x^i}(x(t), \dot{x}(t))\right] dx^i.$$

In particular, if $\mathcal{F} = -dU$ is conservative then the equations of motion are

$$-\frac{\partial U}{\partial x^i}(x(t)) = \frac{d}{dt}\left(\frac{\partial K}{\partial v^i}(x(t), \dot{x}(t))\right) - \frac{\partial K}{\partial x^i}(x(t), \dot{x}(t))$$

for all $i \in \{1, ..., n\}$.

Proof Let $p \in M$, φ a chart of M around p such that $x := \varphi(p) = (x^1, ..., x^n)$. For $v := (v^1, ..., v^n) \in T_pM$ $\left(\text{i.e. } v = \sum_{i=1}^n v^i \frac{\partial}{\partial \varphi^i} = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\right)$, for $g_{ij} := \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right\rangle$, the kinetic energy K at the element of TM whose local coordinates are given by (x, v) is

$$K = \frac{1}{2}\langle v, v\rangle = \frac{1}{2}\sum_{i,j=1}^{n}\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right\rangle v^i v^j = \frac{1}{2}\sum_{i,j=1}^{n} g_{ij} v^i v^j$$

$$\Rightarrow \forall i \in \{1, ..., n\} : \frac{\partial K}{\partial v^i} = \frac{\partial}{\partial v^i}\left(\frac{1}{2}\sum_{j,k=1}^{n} g_{jk} v^j v^k\right) = \sum_{j=1}^{n} g_{ij} v^j.$$

If $v = \dot{x} = \dot{x}(t)$, $x = x(t)$ and $i \in \{1, ..., n\}$ we therefore get

$$\frac{d}{dt}\left(\frac{\partial K}{\partial v^i}(x, \dot{x})\right) = \frac{d}{dt}\left(\sum_{j=1}^{n} g_{ij}(x)\dot{x}^j\right) = \sum_{j,k=1}^{n}\frac{\partial g_{ij}(x)}{\partial x^k}\dot{x}^j\dot{x}^k + \sum_{j=1}^{n} g_{ij}(x)\ddot{x}^j. \quad (1.2)$$

We further get

$$\frac{\partial K}{\partial x^i}(x, \dot{x}) = \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial g_{jk}(x)}{\partial x^i} \dot{x}^j \dot{x}^k. \tag{1.3}$$

By (1.2) and (1.3), we get

$$\frac{d}{dt}\left(\frac{\partial K}{\partial v^i}(x, \dot{x})\right) - \frac{\partial K}{\partial x^i}(x, \dot{x})$$
$$= \sum_{j,k=1}^{n} \left(\frac{\partial g_{ij}(x)}{\partial x^k} - \frac{1}{2}\frac{\partial g_{jk}(x)}{\partial x^i}\right) \dot{x}^j \dot{x}^k + \sum_{j=1}^{n} g_{ij}(x)\ddot{x}^j. \tag{1.4}$$

Now, we would like to compute $\sum_{j=1}^{n} g_{ij}(x)\left(\frac{D\dot{c}}{dt}\right)^j$ for $i, j \in \{1, ..., n\}$, where $\dot{c} = \dot{c}(t)$. Our goal is to show that $\sum_{j=1}^{n} g_{ij}(x)\left(\frac{D\dot{c}}{dt}\right)^j = (1.4)$. If we can show this, then the statement will follow, since for all $v \in T_pM$ $\left(v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\right)$:

$$\mu\left(\frac{D\dot{c}}{dt}\right)(v) = \left\langle \frac{D\dot{c}}{dt}, v \right\rangle = \sum_{i,j=1}^{n} g_{ij}\left(\frac{D\dot{c}}{dt}\right)^j v^i$$
$$= \sum_{i,j=1}^{n} g_{ij}\left(\frac{D\dot{c}}{dt}\right)^j dx^i(v) = \left(\sum_{i,j=1}^{n} g_{ij}\left(\frac{D\dot{c}}{dt}\right)^j dx^i\right)(v).$$

To this end, let us recall the explicit definition of the Christoffel symbols, which can be derived from their implicit definition:

$$\Gamma^j_{kl} = \frac{1}{2}\sum_{m=1}^{n} g^{jm}\left(\frac{\partial g_{lm}}{\partial x^k} + \frac{\partial g_{km}}{\partial x^l} - \frac{\partial g_{kl}}{\partial x^m}\right),$$

where g^{jm} for $j, m \in \{1, ..., n\}$ are the individual entries of the inverse matrix of the matrix whose individual entries are given by g_{ij} for $i, j \in \{1, ..., n\}$, meaning that $\sum_{j=1}^{n} g_{ij}g^{jm} = \delta_{im}$. Therefore,

$$\sum_{j=1}^{n} g_{ij}\Gamma^j_{kl} = \frac{1}{2}\sum_{j,m=1}^{n} g_{ij}g^{jm}\left(\frac{\partial g_{lm}}{\partial x^k} + \frac{\partial g_{km}}{\partial x^l} - \frac{\partial g_{kl}}{\partial x^m}\right)$$
$$= \frac{1}{2}\left(\frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^l} - \frac{\partial g_{kl}}{\partial x^i}\right). \tag{1.5}$$

We would like to rewrite $\frac{D\dot{c}}{dt}$ with the aid of the Christoffel symbols, which we can do by using the following lemma, which we will state without proving it, but noting what is needed for its proof.

Lemma 1.7 *Let M be a smooth n-dimensional manifold, let ∇ be a connection on TM. Let $(A_1, ..., A_n)$ be a moving frame on an open set $U \subseteq M$. For vector fields $X = \sum_{i=1}^n X^i A_i$ and $Y = \sum_{i=1}^n Y^i A_i$ on U,*

$$\nabla_X Y = \sum_{j=1}^{n} \left(X(Y^j) + \sum_{k,l=1}^{n} X^k Y^l \Gamma^j_{kl} \right) A_j,$$

where $\Gamma^j_{kl} \in C^\infty(U)$ for all $j, k, l \in \{1, ..., n\}$ are the Christoffel symbols of ∇ with respect to $(A_1, ..., A_n)$.

The above lemma can be proved by combining the two defining properties of a connection, together with the definition of the Christoffel symbols. It corresponds to Lemma 1.6 in [2], where its proof can be found.
To apply the above lemma to our case, we can take a smooth vector field that coincides with $\dot{c}$ in the domain of definition of $\dot{c}$. Also, we note that $\varphi \circ c = x$ and we identify c with x (instead of explicitly using the chart φ).

Remark 1.8 *In order to be fully rigorous, we would have to work with an analogous lemma to the one we are using. The analogous lemma is for connections along smooth maps between manifolds induced by connections on the manifolds. Further, we would need to identify the basis elements of the tangent spaces via the required composition with c, instead of identifying them directly.*

Using the lemma, we get

$$\frac{D\dot{c}}{dt} = \sum_{j=1}^{n} \left(\ddot{c}^j + \sum_{k,l=1}^{n} \dot{c}^k \dot{c}^l \Gamma^j_{kl} \right) \frac{\partial}{\partial x^j} = \sum_{j=1}^{n} \left(\ddot{x}^j + \sum_{k,l=1}^{n} \dot{x}^k \dot{x}^l \Gamma^j_{kl} \right) \frac{\partial}{\partial x^j} \tag{1.6}$$

$$\Rightarrow \left(\frac{D\dot{c}}{dt} \right)^j = \left(\ddot{x}^j + \sum_{k,l=1}^{n} \Gamma^j_{kl} \dot{x}^k \dot{x}^l \right). \tag{1.7}$$

Finally, by combining (1.5) and (1.7) (and writing $g_{ij} = g_{ij}(x)$) we get

$$\begin{aligned} \sum_{j=1}^{n} g_{ij} \left(\frac{D\dot{c}}{dt} \right)^j &= \sum_{j=1}^{n} g_{ij} \left(\ddot{x}^j + \sum_{k,l=1}^{n} \Gamma^j_{kl} \dot{x}^k \dot{x}^l \right) = \sum_{j=1}^{n} g_{ij} \ddot{x}^j + \sum_{j,k,l=1}^{n} g_{ij} \Gamma^j_{kl} \dot{x}^k \dot{x}^l \\ &= \sum_{j=1}^{n} g_{ij} \ddot{x}^j + \frac{1}{2} \sum_{k,l=1}^{n} \left(\frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^l} - \frac{\partial g_{kl}}{\partial x^i} \right) \dot{x}^k \dot{x}^l \\ &= \sum_{j=1}^{n} g_{ij} \ddot{x}^j + \sum_{k,l=1}^{n} \left(\frac{\partial g_{li}}{\partial x^k} - \frac{1}{2} \frac{\partial g_{kl}}{\partial x^i} \right) \dot{x}^k \dot{x}^l = (1.4). \quad \square \end{aligned}$$

1.3 The Kepler Problem

1.3.1 Solving the Kepler Problem

With the few definitions introduced and the expression of the Newton equation in local coordinates, we can already find a solution to the Kepler problem in the formalism of Riemannian geometry.

The **Kepler problem** (in appropriate units) consists in determining the motion of a particle of mass $m = 1$ in the central potential

$$U(r, \theta) = -\frac{1}{r}.$$

We will work our way to the solution by proving two intermediate claims.

Claim 1.9 *The equations of motion of the corresponding mechanical system can be integrated to*

$$r^2\dot{\theta} = p_\theta; \tag{1.8}$$

$$\frac{\dot{r}^2}{2} + \frac{p_\theta^2}{2r^2} - \frac{1}{r} = E, \tag{1.9}$$

where E and p_θ are integration constants.

Remark 1.10 *The constant p_θ corresponds to the angular momentum of the particle (observe that its mass $m = 1$), whereas the constant E is the total energy of the particle, as the sum of its kinetic and potential energies (*cf. (1.10)*). Therefore, the claim implies that both the angular momentum and the total energy of the system are conserved quantities.*

Proof In polar coordinates,

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases},$$

the Euclidean metric is

$$\langle \cdot, \cdot \rangle = dx \otimes dx + dy \otimes dy = dr \otimes dr + r^2 d\theta \otimes d\theta.$$

Therefore, the kinetic energy is

$$K(r, \theta, v^r, v^\theta) = \frac{1}{2}[(v^r)^2 + r^2(v^\theta)^2], \tag{1.10}$$

and we get the partial derivatives of K

$$\frac{\partial K}{\partial v^r} = v^r; \quad \frac{\partial K}{\partial v^\theta} = r^2 v^\theta; \quad \frac{\partial K}{\partial r} = r(v^\theta)^2; \quad \frac{\partial K}{\partial \theta} = 0.$$

By Proposition 1.6, the equations of motion are

$$\frac{d}{dt}(\dot{r}) - r\dot{\theta}^2 = -\frac{1}{r^2}; \tag{1.11}$$

$$\frac{d}{dt}(r^2\dot{\theta}) = 0. \tag{1.12}$$

By setting $p_\theta := r^2\dot{\theta}$ (which by equation (1.12) is constant as a function of time) we readily obtain the first desired equation (1.8). To obtain the other one, let us observe that (1.11) is equivalent to

$$\dot{r}d\dot{r} - \frac{p_\theta^2}{r^3}dr = -\frac{1}{r^2}dr,$$

which then integrates to

$$\frac{\dot{r}^2}{2} + \frac{p_\theta^2}{2r^2} = \frac{1}{r} + E,$$

where E is an integration constant. □

Claim 1.11 $u(t) = \frac{1}{r(t)}$ *satisfies the linear ODE*

$$\frac{d^2u}{d\theta^2} + u = \frac{1}{p_\theta^2}. \tag{1.13}$$

Proof Note that $\frac{d}{d\theta} = \frac{1}{\dot{\theta}}\frac{d}{dt}$. Together with (1.8) and (1.12), we get

$$\begin{aligned} \frac{d^2}{d\theta^2}(u) = \frac{d^2}{d\theta^2}\left(\frac{1}{r}\right) &= \frac{d}{d\theta}\left[\frac{1}{\dot{\theta}}\frac{d}{dt}\left(\frac{1}{r}\right)\right] \\ &= \frac{d}{d\theta}\left(-\frac{\dot{r}}{r^2\dot{\theta}}\right) = -\frac{1}{\dot{\theta}}\frac{d}{dt}\left(\frac{\dot{r}}{p_\theta}\right) = -\frac{\ddot{r}}{p_\theta\dot{\theta}} = -\frac{p_\theta\ddot{r}}{p_\theta^2\dot{\theta}}. \end{aligned} \tag{1.14}$$

Using equations (1.8), (1.11) and (1.14) we conclude.

$$\frac{d^2u}{d\theta^2} + u = \frac{-p_\theta\ddot{r}}{p_\theta^2\dot{\theta}} + \frac{1}{r} = \frac{-r^2\dot{\theta}\ddot{r}}{p_\theta^2\dot{\theta}} + \frac{r^3\dot{\theta}^3}{p_\theta^2\dot{\theta}} = \frac{-r^2(\ddot{r} - r\dot{\theta}^2)}{p_\theta^2} = \frac{-r^2\left(-\frac{1}{r^2}\right)}{p_\theta^2} = \frac{1}{p_\theta^2}. \quad \square$$

Claim 1.12 *Assume that the* ***pericenter*** *(the point in the particle's orbit closest to the center of attraction* $r = 0$*) occurs at* $\theta = 0$*. Then the equation of the particle's trajectory is*

$$r = \frac{p_\theta^2}{1 + \varepsilon\cos\theta}, \tag{1.15}$$

where

$$\varepsilon = \sqrt{1 + 2p_\theta^2 E}. \tag{1.16}$$

Remark 1.13 (1.15) *is the equation of a conic section with eccentricity ε in polar coordinates.*

Proof We first show that $u = \frac{1}{r}$ where r is given by (1.15) is a solution to equation (1.13):

$$\frac{d^2}{d\theta^2}\left(\frac{1+\varepsilon\cos\theta}{p_\theta^2}\right) + \frac{1+\varepsilon\cos\theta}{p_\theta^2} = -\frac{\varepsilon\cos\theta}{p_\theta^2} + \frac{1+\varepsilon\cos\theta}{p_\theta^2} = \frac{1}{p_\theta^2}.$$

That our solution u of equation (1.13) is unique follows from the Picard-Lindelöf Theorem, where the necessary conditions to apply the Theorem (especially the Lipschitz continuity) readily follow from the structure of (1.13) (note that we need the initial condition that the pericenter occurs at $r = 0$). We then proceed to derive the value of ε by inserting (1.15) into (1.9):

$$\begin{aligned} E = \frac{p_\theta^4\varepsilon^2\dot{\theta}^2\sin^2\theta}{2(1+\varepsilon\cos\theta)^4} + \frac{(1+\varepsilon\cos\theta)^2}{2p_\theta^2} - \frac{1+\varepsilon\cos\theta}{p_\theta^2} \\ = \frac{p_\theta^6\varepsilon^2\dot{\theta}^2\sin^2\theta + (1+\varepsilon\cos\theta)^5(-1+\varepsilon\cos\theta)}{2p_\theta^2(1+\varepsilon\cos\theta)^4}. \end{aligned}$$

The above equality must hold for all values of θ, and so it must hold in particular for $\theta = 0$:

$$\begin{aligned} E &= \frac{(\varepsilon+1)(\varepsilon-1)}{2p_\theta^2} \\ &\Leftrightarrow 1 + 2p_\theta^2 E = \varepsilon^2. \end{aligned}$$

We take the positive value of the square root for ε due to the initial condition that the pericenter occurs at $\theta = 0$, which implies that the value of r is minimal at $\theta = 0$. □

1.3.2 An Interesting Manifold

Having found a solution to the Kepler problem, we consider the manifold $(M, g) := (\mathbb{R}^2 \setminus \{(0,0)\}, \langle\cdot,\cdot\rangle)$, where

$$\langle\cdot,\cdot\rangle = \frac{1}{\sqrt{x^2+y^2}}(dx\otimes dx + dy\otimes dy) = \frac{1}{r}dr\otimes dr + rd\theta\otimes d\theta, \tag{1.17}$$

the final form being in polar coordinates. This manifold, as we will show, is isometric to the surface $\bar{M}$ of a cone with aperture $\frac{\pi}{3}$. If we then characterise its geodesics, we will find that they have a similar structure to the equations of motion in the Kepler problem.
Let

$$\bar{M} := \left\{\left(\frac{r}{2}\cos\theta, \frac{r}{2}\sin\theta, \frac{\sqrt{3}}{2}r\right) : r > 0, \theta \in [0, 2\pi)\right\};$$

$$\bar{g} := dx \otimes dx + dy \otimes dy + dz \otimes dz. \tag{1.18}$$

Let $f : M \to \bar{M}$ be defined by $f(r\cos\theta, r\sin\theta) = \sqrt{r}(cos\theta, sin\theta, \sqrt{3})$.

Claim 1.14 *f is an isometry from (M, g) to $(\bar{M}, \bar{g})$.*

Proof We must show that

1. f is an immersion.
2. f is a diffeomorphism.
3. $f^*\bar{g} = g$, where $f^*\bar{g}$ is the pull-back metric on M.

1. f is an immersion:
   For $p = (r\cos\theta, r\sin\theta) \in M$, the matrix form of df_p is

$$df_p = \begin{pmatrix} \frac{\cos\theta}{2\sqrt{r}} & -\sqrt{r}\sin\theta \\ \frac{\sin\theta}{2\sqrt{r}} & \sqrt{r}\cos\theta \\ \frac{\sqrt{3}}{2\sqrt{r}} & 0 \end{pmatrix}, \tag{1.19}$$

   which has rank 2, and so is injective.

2. f is a diffeomorphism:
   That f is bijective is quite straightforward: the injectivity is clear, for $q \in \bar{M}$ it holds that $q = \left(\frac{r}{2}\cos\theta, \frac{r}{2}\sin\theta, \frac{\sqrt{3}}{2}r\right)$ for some $r > 0$ and $\theta \in [0, 2\pi)$, and so, by taking $p = \left(\frac{r^2}{4}\cos\theta, \frac{r^2}{4}\sin\theta\right) \in M$, we get that $f(p) = q$. Also quite straightforwardly (or recalling how we showed that f is surjective), we see that $f^{-1} : \bar{M} \to M$ must be given by $f^{-1}\left(\frac{r}{2}\cos\theta, \frac{r}{2}\sin\theta, \frac{\sqrt{3}}{2}r\right) = \left(\frac{r^2}{4}\cos\theta, \frac{r^2}{4}\sin\theta\right)$. We then note that both f and f^{-1} are smooth as the composition of smooth functions in both variables, meaning that f is a diffeomorphism.

3. $f^*\bar{g} = g$, where $f^*\bar{g}$ is the pull-back metric on M:
   Let us recall that for $v, w \in T_pM$ for $p \in M$,

$$(f^*\bar{g})_p(v, w) = \bar{g}_{f(p)}(df_p(v), df_p(w)).$$

   Our goal is to show that this metric is equal to g on M:
   Let $v, w \in T_pM$ for $p \in M$, $\quad v = v_r\frac{\partial}{\partial r} + v_\theta\frac{\partial}{\partial\theta}, \quad w = w_r\frac{\partial}{\partial r} + w_\theta\frac{\partial}{\partial\theta}$. It

holds, by (1.17), (1.18) and (1.19):

$$\begin{aligned}
g_p(v,w) &= g_p\left(v_r\frac{\partial}{\partial r}+v_\theta\frac{\partial}{\partial\theta}, w_r\frac{\partial}{\partial r}+w_\theta\frac{\partial}{\partial\theta}\right)\\
&= v_r w_r g_p\left(\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right)+v_\theta w_\theta g_p\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right) = v_r w_r\frac{1}{r}+v_\theta w_\theta r\\
&= \left\langle v_r\frac{1}{2\sqrt{r}}\begin{pmatrix}\cos\theta\\ \sin\theta\\ \sqrt{3}\end{pmatrix}+v_\theta\sqrt{r}\begin{pmatrix}-\sin\theta\\ \cos\theta\\ 0\end{pmatrix},\right.\\
&\qquad \left. w_r\frac{1}{2\sqrt{r}}\begin{pmatrix}\cos\theta\\ \sin\theta\\ \sqrt{3}\end{pmatrix}+w_\theta\sqrt{r}\begin{pmatrix}-\sin\theta\\ \cos\theta\\ 0\end{pmatrix}\right\rangle\\
&= \bar{g}_{f(p)}(df_p(v),df_p(w)) \qquad \square
\end{aligned}$$

Now, let us characterise the geodesics of (M,g), i.e. the motions of $(M,g,0)$ (of course, in polar coordinates): The kinetic energy of the mechanical system is

$$\begin{aligned}
K(r,\theta,\dot r,\dot\theta) &= \frac{1}{2}\left(\frac{1}{r}\dot r^2+r\dot\theta^2\right)\\
&\Rightarrow \frac{\partial K}{\partial r}=\frac{1}{2}\left(-\frac{\dot r^2}{r^2}+\dot\theta^2\right);\quad \frac{\partial K}{\partial\theta}=0;\quad \frac{\partial K}{\partial\dot r}=\frac{\dot r}{r};\quad \frac{\partial K}{\partial\dot\theta}=r\dot\theta,
\end{aligned}$$

which, by Proposition (1.6) implies that the equations of motion are:

$$\frac{d}{dt}\left(\frac{\dot r}{r}\right)+\frac{1}{2}\left(\frac{\dot r^2}{r^2}-\dot\theta^2\right)=0$$

$$\Leftrightarrow \ddot r r-\frac{\dot r^2}{2}-\frac{r^4\dot\theta^2}{2r^2}=0; \tag{1.20}$$

$$\frac{d}{dt}(r\dot\theta)=0. \tag{1.21}$$

To conclude, we note the similar structure of equations (1.9) and (1.20), especially when substituting equation (1.8) into (1.9). We therefore ask whether there are solutions to the Kepler problem that are also geodesics of the surface of a cone with aperture $\frac{\pi}{3}$.

If such a solution were to exist, it would have to satisfy equations (1.12) and (1.21), implying that r would have to be constant, and therefore $\dot\theta$ would also have be constant (by equation (1.21)). Observing that r cannot be 0, and that $r = const. \Rightarrow 0 = \dot r = \ddot r$, equation (1.20) would imply that $\dot\theta = 0$. Equation (1.11) would then yield a contradiction, as we would have

$$0 = -\frac{1}{r^2}.$$

Therefore, there are no solutions to the Kepler problem that are also geodesics of M.

Chapter 2

Holonomic Constraints

Oftentimes, we have a system comprised of several particles that are not able to move freely in the whole configuration space, but are instead restricted to motions within a smaller subspace. This is the case of a pendulum or a particle that can only move along a surface in $\mathbb{R}^3$, for instance. Holonomic constraints are the notion we introduce to account for these restrictions. Further, we need an additional term in the Newton equation, an extra force. This is the so called reaction force, and it is responsible for enforcing the restrictions on the motion of a system. In the case of a pendulum, for instance, the reaction force corresponds to the tension force of the rod.
In this chapter, we present the necessary concepts to account for these constraints, and we study the example of the double pendulum, one of the most prominent chaotic systems.

2.1 Holonomic Constraints and Reaction Forces

Definition 2.1 (Holonomic constraint) *A **holonomic constraint** on a mechanical system $(M, \langle\cdot,\cdot\rangle, \mathcal{F})$ is a submanifold $N \subseteq M$ with $dim(N) < dim(M)$. A curve $c : (a,b) \subseteq \mathbb{R} \to M$ is said to be **compatible** with N if $c(t) \in N$ for all $t \in (a,b)$.*

The modification of the Newton equation (1.1) to account for the constraint is achieved by introducing the reaction force.

Definition 2.2 (Reaction force) *A **reaction force** on a mechanical system with holonomic constraint $(M, \langle\cdot,\cdot\rangle, \mathcal{F}, N)$ is a map $\mathcal{R} : TN \to T^*M$ satisfying $\mathcal{R}(T_pN) \subseteq T_p^*M$ for all $p \in M$ such that, for each $v \in TN$, there is a solution $c : (a,b) \subseteq \mathbb{R} \to N$ of the **generalised Newton equation***

$$\mu\left(\frac{D\dot{c}}{dt}\right) = (\mathcal{F} + \mathcal{R})(\dot{c})$$

with initial condition $\dot{c}(0) = v$.

The definition of a reaction force is by no means restrictive insofar as it allows for many different forces to be a reaction force on a mechanical system. However, we would like to find a particular reaction force with which to work. The choice is facilitated by the following definition and subsequent theorem.

Definition 2.3 *A reaction force on a mechanical system with holonomic constraint* $(M, \langle \cdot, \cdot \rangle, \mathcal{F}, N)$ *is said to be* ***perfect****, or to satisfy the* ***d'Alembert principle****, if*

$$\mu^{-1}(\mathcal{R}(v)) \in (T_p N)^{\perp}$$

for all $v \in T_p N$ *and* $p \in N$.

The variation of the kinetic energy of a solution of the generalised Newton equation is

$$\frac{dK}{dt} = \left\langle \frac{D\dot{c}}{dt}, \dot{c} \right\rangle = \mu\left(\frac{D\dot{c}}{dt}\right)(\dot{c}) = \mathcal{F}(\dot{c})(\dot{c}) + \mathcal{R}(\dot{c})(\dot{c}) = \mathcal{F}(\dot{c})(\dot{c}) + \left\langle \mu^{-1}(\mathcal{R}(\dot{c})), \dot{c} \right\rangle. \quad (2.1)$$

Since for any motion compatible with the constraint we have that $\dot{c}(t) \in T_{c(t)}N$ for all $t \in (a, b) \subseteq \mathbb{R}$, we have that the last term in the above equation will vanish if $\mathcal{R}$ is a perfect reaction force. In other words, a perfect reaction force neither contributes nor deducts kinetic energy to a particle whose motion is compatible with the holonomic constraint. In the case of a single pendulum, for instance, a perfect reaction force corresponds to a tension force that is radial along the rod. Since it is radial, there is no damping.

Theorem 2.4 *Given any mechanical system with holonomic constraint* $(M, \langle \cdot, \cdot \rangle, \mathcal{F})$*, there exists a unique reaction force* $R : TN \to T^*M$ *satisfying the d'Alembert principle.*

The proof can be found in Section 2 of Chapter 5 of [1].

2.2 The Double Pendulum

Having introduced the necessary concepts to account for the restriction on the position of the particles of a given system, we want to look at an example, the double pendulum, and compute its equations of motion. As stated in Definition 2.1, this restriction on the position of the particles will amount to defining an appropriate submanifold. It is noteworthy that we will be able to compute the equations of motion, however, without the need of explicitly computing the reaction force.

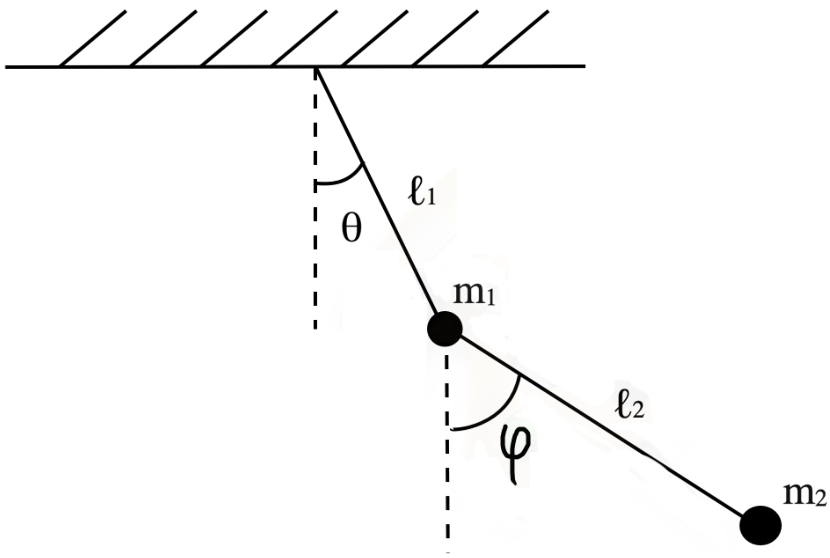

Figure 2.1: Double pendulum

2.2.1 Introduction and Equations of Motion

The **double pendulum** of lengths ℓ_1, ℓ_2 is the mechanical system defined by two particles of masses m_1, m_2 moving in $\mathbb{R}^2$ subject to a constant gravitational acceleration g and the holonomic constraint

$$N = \{(u_1, u_2) \in \mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2 : \|u_1\| = \ell_1, \|u_1 - u_2\| = \ell_2\}$$

We use the parametrisation $\psi : (-\pi, \pi) \times (-\pi, \pi) \to N$ of the holonomic constraint N given by

$$\psi(\theta, \varphi) = (\ell_1 \sin\theta, -\ell_1 \cos\theta, \ell_1 \sin\theta + \ell_2 \sin\varphi, -\ell_1 \cos\theta - \ell_2 \cos\varphi)$$

to write the equations of motion for the double pendulum. Set $p := \psi(\theta, \varphi)$. We want to use Proposition 1.6, and so we need to compute the kinetic energy K of the system. This total kinetic energy is the sum of the individual kinetic energies of both particles. To compute the kinetic energy of each particle, let us set $(x, y, z, w) := \psi(\theta, \varphi)$. Then the position of the first particle is given by (x, y) and that of the second by (z, w). Let $\frac{\partial}{\partial\theta}$ and $\frac{\partial}{\partial\varphi}$ be the basis vectors of T_pN induced by the local coordinates (θ, φ) on the tangent space. Further, let $v = v^\theta \frac{\partial}{\partial\theta} + v^\varphi \frac{\partial}{\partial\varphi} \in T_pN$ be the total velocity of the system. Then the velocities v^1 and v^2 of each particle are the orthogonal projections of v to the first and second $\mathbb{R}^2$-factors of $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$, respectively. The expressions of $\frac{\partial}{\partial\theta}$ and $\frac{\partial}{\partial\varphi}$ in terms of the standard basis of $\mathbb{R}^4$ $\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial w}\right\}$ are given by the columns of the matrix representation of the total derivative of ψ at (θ, φ):

$$D\psi(\theta,\varphi) = \left(\frac{\partial}{\partial\theta}(p), \frac{\partial}{\partial\varphi}(p)\right) = \begin{pmatrix} \ell_1\cos\theta & 0 \\ \ell_1\sin\theta & 0 \\ \ell_1\cos\theta & \ell_2\cos\varphi \\ \ell_1\sin\theta & \ell_2\sin\varphi \end{pmatrix}.$$

Therefore, the velocities v_1 and v_2 in terms of the standard basis of $\mathbb{R}^4$ are given by

$$v^1 = v^\theta\left(\ell_1\cos\theta\frac{\partial}{\partial x} + \ell_1\sin\theta\frac{\partial}{\partial y}\right);$$

$$v^2 = v^\theta\left(\ell_1\cos\theta\frac{\partial}{\partial z} + \ell_1\sin\theta\frac{\partial}{\partial w}\right) + v^\varphi\left(\ell_2\cos\varphi\frac{\partial}{\partial z} + \ell_2\sin\varphi\frac{\partial}{\partial w}\right).$$

Further, note that $\langle\cdot,\cdot\rangle$ on T_pN for any $p \in N$ is just the restriction to $T_pN \subseteq \mathbb{R}^4$ of the standard metric on $\mathbb{R}^4$. Since $\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial w}\right\}$ is an orthonormal basis, the individual kinetic energies K_1 and K_2 are

$$K_1 = \frac{1}{2}m_1\langle v^1, v^1\rangle = \frac{1}{2}m_1(v^\theta)^2\ell_1^2;$$

$$K_2 = \frac{1}{2}m_2\langle v^2, v^2\rangle = \frac{1}{2}m_2[(v^\theta)^2\ell_1^2 + (v^\varphi)^2\ell_2^2 + 2v^\theta v^\varphi\ell_1\ell_2\cos(\theta-\varphi)],$$

resulting in the total kinetic energy

$$\begin{aligned} K &= K_1 + K_2 \\ &= \frac{1}{2}[(m_1+m_2)(v^\theta)^2\ell_1^2 + m_2(v^\varphi)^2\ell_2^2] + m_2 v^\theta v^\varphi \ell_1\ell_2\cos(\theta-\varphi). \end{aligned}$$

Then the partial derivatives are

$$\frac{\partial K}{\partial\theta} = -m_2 v^\theta v^\varphi\ell_1\ell_2\sin(\theta-\varphi); \quad \frac{\partial K}{\partial\varphi} = m_2 v^\theta v^\varphi\ell_1\ell_2\sin(\theta-\varphi);$$

$$\frac{\partial K}{\partial v^\theta} = (m_1+m_2)v^\theta\ell_1^2 + m_2 v^\varphi\ell_1\ell_2\cos(\theta-\varphi);$$

$$\frac{\partial K}{\partial v^\varphi} = m_2 v^\varphi\ell_2^2 + m_2 v^\theta\ell_1\ell_2\cos(\theta-\varphi),$$

and the expressions that appear on the right hand side of Proposition 1.6 are, therefore, given by (where we write $\dot{\theta} = v^\theta$, $\dot{\varphi} = v^\varphi$ and $\ddot{\theta} = \frac{dv^\theta}{dt}$, $\ddot{\varphi} = \frac{dv^\varphi}{dt}$)

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial K}{\partial v^\theta}\right) - \frac{\partial K}{\partial \theta} &= (m_1+m_2)\ddot{\theta}\ell_1^2 + m_2\ddot{\varphi}\ell_1\ell_2\cos(\theta-\varphi) \\ &\quad - m_2\dot{\varphi}\ell_1\ell_2\sin(\theta-\varphi)(\dot{\theta}-\dot{\varphi}) + m_2\dot{\theta}\dot{\varphi}\ell_1\ell_2\sin(\theta-\varphi) \\ &= (m_1+m_2)\ddot{\theta}\ell_1^2 + m_2\ddot{\varphi}\ell_1\ell_2\cos(\theta-\varphi) + m_2\dot{\varphi}^2\ell_1\ell_2\sin(\theta-\varphi); \end{aligned} \tag{2.2}$$

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial K}{\partial v^\varphi}\right) - \frac{\partial K}{\partial \varphi} &= m_2\ddot{\varphi}\ell_2^2 + m_2\ddot{\theta}\ell_1\ell_2\cos(\theta-\varphi) \\ &\quad - m_2\dot{\theta}\ell_1\ell_2\sin(\theta-\varphi)(\dot{\theta}-\dot{\varphi}) - m_2\dot{\theta}\dot{\varphi}\ell_1\ell_2\sin(\theta-\varphi) \\ &= m_2\ddot{\varphi}\ell_2^2 + m_2\ddot{\theta}\ell_1\ell_2\cos(\theta-\varphi) - m_2\dot{\theta}^2\ell_1\ell_2\sin(\theta-\varphi). \end{aligned} \tag{2.3}$$

The potential energy $U : N \to \mathbb{R}$ of the system is given by

$$U(x,y,z,w) := m_1 g y + m_2 g w = -m_1 g \ell_1 \cos\theta - m_2 g(\ell_1\cos\theta + \ell_2\cos\varphi),$$

implying that the left hand side of the equations in Proposition 1.6 are given by

$$-\frac{\partial U}{\partial \theta} = -(m_1+m_2)g\ell_1\sin\theta; \tag{2.4}$$

$$-\frac{\partial U}{\partial \varphi} = -m_2 g\ell_2\sin\varphi. \tag{2.5}$$

By combining equations (2.4) and (2.2), (2.5) and (2.3), we get the equations of motion of the double pendulum

$$\begin{aligned} -\frac{\partial U}{\partial \theta} &= \frac{d}{dt}\left(\frac{\partial K}{\partial v^\theta}\right) - \frac{\partial K}{\partial \theta} \\ &\Leftrightarrow -(m_1+m_2)g\sin\theta = (m_1+m_2)\ddot{\theta}\ell_1 \\ &\qquad + m_2\ddot{\varphi}\ell_2\cos(\theta-\varphi) + m_2\dot{\varphi}^2\ell_2\sin(\theta-\varphi); \end{aligned} \tag{2.6}$$

$$\begin{aligned} -\frac{\partial U}{\partial \varphi} &= \frac{d}{dt}\left(\frac{\partial K}{\partial v^\varphi}\right) - \frac{\partial K}{\partial \varphi} \\ &\Leftrightarrow -g\sin\varphi = \ddot{\varphi}\ell_2 + \ddot{\theta}\ell_1\cos(\theta-\varphi) - \dot{\theta}^2\ell_1\sin(\theta-\varphi). \end{aligned} \tag{2.7}$$

2.2.2 Linearisation of the Equations of Motion

We conclude this example by linearising the equations of motion around $\theta = \varphi = 0$. We do this by the small-angle approximation $\sin\theta \approx \theta$ and $\cos\theta \approx 1 - \frac{\theta^2}{2} \approx 1$. We can impose the initial conditions $\dot{\theta}_0 := \dot{\theta}(0) = 0$ and $\dot{\varphi}_0 := \dot{\varphi}(0) = 0$, which imply $\dot{\theta}^2 \approx 0$ and $\dot{\varphi}^2 \approx 0$. We then get
(2.6) $\rightarrow$

$$-(m_1 + m_2)g\theta = (m_1 + m_2)\ddot{\theta}\ell_1 + m_2\ddot{\varphi}\ell_2; \tag{2.8}$$

(2.7) $\rightarrow$

$$-g\varphi = \ddot{\varphi}\ell_2 + \ddot{\theta}\ell_1. \tag{2.9}$$

The **normal modes** are the solutions that satisfy the relationship $\varphi = k\theta$ for $k \in \mathbb{R}$ a constant. Setting $\varphi = k\theta$ and simplifying some of the terms on both sides implies (assuming that $k\ell_2 \neq -\ell_1$) that

$$(2.9) \Leftrightarrow \ddot{\theta} = -\frac{gk}{k\ell_2 + \ell_1}\theta, \tag{2.10}$$

which, in turn, implies by substituting $\ddot{\theta}$ into equation (2.8) that

$$\begin{aligned}
-(m_1 + m_2)g\theta &= -[(m_1 + m_2)\ell_1 + m_2\ell_2 k]\frac{gk}{k\ell_2 + \ell_1}\theta \\
&\Leftrightarrow (m_1 + m_2)(k\ell_2 + \ell_1) = (m_1 + m_2)\ell_1 k + m_2\ell_2 k^2 \\
&\Leftrightarrow m_2\ell_2 k^2 + (m_1 + m_2)(\ell_1 - \ell_2)k - \ell_1(m_1 + m_2) = 0 \\
\Rightarrow k &= \frac{-(m_1 + m_2)(\ell_1 - \ell_2) \pm \sqrt{(m_1 + m_2)^2(\ell_1 - \ell_2)^2 + 4(m_1 + m_2)\ell_1\ell_2 m_2}}{2m_2\ell_2} \\
&= \frac{\ell_2 - \ell_1 \pm \sqrt{(\ell_2 - \ell_1)^2 + 4\ell_1\ell_2 M_2}}{2M_2\ell_2},
\end{aligned}$$

where $M_2 := \frac{m_2}{m_1 + m_2}$.

Remark 2.5 *In the expression for k, the discriminant in the square root will always be bigger than $(\ell_2 - \ell_1)^2$, implying that k will be positive by taking the sum $(+)$ in the numerator and negative with the subtraction $(-)$; this corresponds, respectively, to the oscillations of the individual pendulums being in phase or in opposite phase (reflected explicitly by the fact that $\varphi = k\theta$).*

Finally, we can find the period T of the normal modes as a function of k. By equation (2.10), we have (recall that we set $\dot{\theta}_0 = 0$)

$$\theta = \theta_0 \cos\omega t \quad \text{for } \theta_0 \in \mathbb{R} \text{ and } \omega = \sqrt{\frac{gk}{k\ell_2 + \ell_1}}$$

$$\Rightarrow T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{k\ell_2 + \ell_1}{gk}} = 2\pi\sqrt{\frac{\ell_2}{g} + \frac{\ell_1}{gk}}.$$

Claim 2.6 *T is well-defined, and the assumption made for equation* (2.10) *that $k\ell_2 \neq -\ell_1$ holds.*

Proof For the positive value k_+ of k (where we take the addition in the numerator), both quantities in the square root are positive, and so T is well-defined. Also, k_+ is positive, and so are ℓ_1 and ℓ_2, and so the second statement of the claim always holds.
For the negative value k_- of k (where we take the subtraction in the numerator), we have to show that

$$\frac{\ell_2}{g} + \frac{\ell_1}{gk_-} > 0 \Leftrightarrow \ell_2 > -\frac{\ell_1}{k_-}. \tag{2.11}$$

From this inequality will also follow the second statement of the claim. We prove the inequality (2.11):

$$\begin{aligned}\frac{\ell_1}{k_-} &= \frac{2M_2\ell_1\ell_2}{\ell_2 - \ell_1 - \sqrt{(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2}} \\ &= \frac{2M_2\ell_1\ell_2(\ell_2 - \ell_1 + \sqrt{(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2})}{(\ell_2-\ell_1)^2 - [(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2]} \\ &= -\frac{\ell_2 - \ell_1 + \sqrt{(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2}}{2},\end{aligned}$$

and, therefore,

$$\begin{aligned}\ell_2 > -\frac{\ell_1}{k_-} &\Leftrightarrow \ell_2 > \frac{\ell_2 - \ell_1 + \sqrt{(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2}}{2} \\ &\Leftrightarrow \ell_1 + \ell_2 > \sqrt{(\ell_2-\ell_1)^2 + 4\ell_1\ell_2 M_2} \Leftrightarrow 1 > M_2,\end{aligned}$$

which holds because, by definition, $M_2 = \frac{m_2}{m_1+m_2} \in (0, 1)$. □

Remark 2.7 *The result for T implies that the period of oscillation of a double pendulum is bigger if its pendulums oscillate in phase than if they oscillate in opposite phase.*

Chapter 3

Rigid Body

In this chapter, we turn our attention to rigid bodies and to several concepts that help characterise them. We often focus on properties of a freely moving rigid body, a system known as the Euler Top. We also present some results involving symmetries and observable effects, including the Poinsot Theorem.

3.1 From the Basics to the Geodesics of the Euler Top

3.1.1 Rigid Bodies and Angular Momentum

In its discrete form, a rigid body corresponds to a system of k particles of masses $m_1, ..., m_k$ connected by massless rods in such a way that their mutual distances remain constant. If we assume that one of the particles is fixed at the origin, the rigid body can be described by the holonomic constraint

$$N = \{(x_1, ..., x_k) \in \mathbb{R}^{3k} : x_1 = 0 \text{ and } \|x_i - x_j\| = d_{ij} \text{ for } 1 \leq i < j \leq k\}.$$

If at least three particles are not collinear, N is a manifold diffeomorphic to $O(3)$. To describe a motion in N, we observe that given a point $(\xi_1, ..., \xi_k) \in N : \forall(\bar{\xi}_1, ..., \bar{\xi}_k) \in N \; \exists! \; S \in O(3)$ such that $(\bar{\xi}_1, ..., \bar{\xi}_k) = (S\xi_1, ..., S\xi_k)$. Therefore, a motion in N can be characterised by a curve $S : (a, b) \subseteq \mathbb{R} \to O(3)$. We further observe that $O(3)$ has two diffeomorphic connected components, and that a motion necessarily has to happen in either of the connected components, meaning that we can choose $SO(3)$ instead of $O(3)$ as our configuration space. The generalisation of the discrete case to a continuum rigid body gives rise to the following definition.

Definition 3.1 (Rigid body with a fixed point) *A* ***rigid body with a fixed point*** *is any mechanical system of the form* $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle, \mathcal{F})$, *with*

$$\langle\langle V, W \rangle\rangle := \int_{\mathbb{R}^3} \langle V\xi, W\xi \rangle \, dm$$

*for all $V, W \in T_S SO(3)$ and all $S \in SO(3)$, where $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product on $\mathbb{R}^3$ and m (called the **mass distribution of the reference configuration**) is a positive finite measure on $\mathbb{R}^3$, not supported on any straight line through the origin, and satisfying $\int_{\mathbb{R}^3} \|\xi\|^2 \, dm < +\infty$.*

We are interested in studying the geodesics of $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle)$, i.e. the equations of motion of $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle, 0)$, a mechanical system known as the **Euler top**. We introduce the following quantity as is turns out to be very useful for their derivation.

Definition 3.2 (Angular momentum) *The **angular momentum** of a rigid body whose motion is described by $S : (a, b) \subseteq \mathbb{R} \to SO(3)$ is the vector*

$$p(t) := \int_{\mathbb{R}^3} [(S(t)\xi) \times (\dot{S}(t)\xi)] \, dm$$

(where $\times$ is the usual cross product on $\mathbb{R}^3$).

Theorem 3.3 *If $S : (a, b) \subseteq \mathbb{R} \to SO(3)$ is a geodesic of $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle)$, then p(t) is constant.*

The proof can be found in Section 3 of Chapter 5 of [1]. The proof in the book uses the following lemma without proving it, which we will do.

Lemma 3.4 *There exists a linear isomorphism $\Gamma : \mathfrak{so}(3) \to \mathbb{R}^3$ such that*

$$A\xi = \Gamma(A) \times \xi$$

for all $\xi \in \mathbb{R}^3$ and $A \in \mathfrak{so}(3)$. Moreover, $\Gamma([A, B]) = \Gamma(A) \times \Gamma(B)$ for all $A, B \in \mathfrak{so}(3)$ (i.e. Γ is a Lie algebra isomorphism between $\mathfrak{so}(3)$ and $(\mathbb{R}^3, \times)$).

Proof Let $A \in \mathfrak{so}(3)$, i.e. $A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$ for $a, b, c \in \mathbb{R}$.

Let $\Gamma : \mathfrak{so}(3) \to \mathbb{R}^3$ be given by $\Gamma(A) = \begin{pmatrix} -c \\ b \\ -a \end{pmatrix}$. Then

$$\begin{aligned} A\xi = A \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \xi_1 \begin{pmatrix} 0 \\ -a \\ -b \end{pmatrix} + \xi_2 \begin{pmatrix} a \\ 0 \\ -c \end{pmatrix} + \xi_3 \begin{pmatrix} b \\ c \\ 0 \end{pmatrix} \\ = \begin{pmatrix} -c \\ b \\ -a \end{pmatrix} \times \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \Gamma(A) \times \xi \end{aligned}$$

That Γ is both injective and surjective follows very easily from its definition. Its linearity is also pretty straightforward, since for any $\lambda \in \mathbb{R}$, for any

$A = \begin{pmatrix} 0 & a_A & b_A \\ -a_A & 0 & c_A \\ -b_A & -c_A & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & a_B & b_B \\ -a_B & 0 & c_B \\ -b_B & -c_B & 0 \end{pmatrix} \in \mathfrak{so}(3)$, we have that

$$\Gamma(\lambda A + B) = \lambda \begin{pmatrix} -c_A \\ b_A \\ -a_A \end{pmatrix} + \begin{pmatrix} -c_B \\ b_B \\ -a_B \end{pmatrix} = \lambda\Gamma(A) + \Gamma(B).$$

We finally need to check the compatibility of Γ with the commutator $[\cdot,\cdot]$ on $\mathfrak{so}(3)$ and the cross product $\times$ on $\mathbb{R}^3$:
With $A, B \in \mathfrak{so}(3)$ as above, we have that

$$\begin{aligned}\Gamma([A,B]) &= \Gamma(AB - BA) \\ &= \Gamma \begin{pmatrix} 0 & -b_A c_B + b_B c_A & a_A c_B - a_B c_A \\ -c_A b_B + c_B b_A & 0 & -a_A b_B + a_B b_A \\ a_B c_A - a_A c_B & -b_A a_B + b_B a_A & 0 \end{pmatrix} \\ &= \begin{pmatrix} -a_B b_A + a_A b_B \\ a_A c_B - a_B c_A \\ -c_A b_B + c_B b_A \end{pmatrix} = \begin{pmatrix} -c_A \\ b_A \\ -a_A \end{pmatrix} \times \begin{pmatrix} -c_B \\ b_B \\ -a_B \end{pmatrix} = \Gamma(A) \times \Gamma(B) \qquad \square\end{aligned}$$

Remark 3.5 *Given a curve $S : (a,b) \subseteq \mathbb{R} \to SO(3)$, we have $\dot{S} \in T_S SO(3)$ and so $\forall t \in (a,b) : \dot{S}(t) = S(t)A(t)$ for some $A(t) \in \mathfrak{so}(3)$. Therefore, we can define a function $\Omega : (a,b) \subseteq \mathbb{R} \to \mathbb{R}^3$ given by*

$$\Omega(t) = \Gamma(A(t)), \tag{3.1}$$

where Γ is the isomorphism of Lemma 3.4. This function then allows us to write $\forall t \in (a,b), \forall \xi \in \mathbb{R}^3$:

$$\dot{S}(t)\xi = S(t)A(t)\xi = S(t)(\Omega(t) \times \xi).$$

3.1.2 Moment of Inertia

Recall Definition 3.2 (angular momentum). By Remark 3.5, together with the fact that $S \in SO(3)$ preserves the cross product $\times$ in $\mathbb{R}^3$, we can now write

$$\begin{aligned} p = \int_{\mathbb{R}^3} [(S\xi) \times (\dot{S}\xi)]\, dm &= \int_{\mathbb{R}^3} [(S\xi) \times (SA\xi)]\, dm \\ &= \int_{\mathbb{R}^3} S[\xi \times (A\xi)]\, dm = \int_{\mathbb{R}^3} S[\xi \times (\Omega \times \xi)]\, dm \\ &= S \int_{\mathbb{R}^3} [\xi \times (\Omega \times \xi)]\, dm. \end{aligned} \tag{3.2}$$

This motivates defining the following operator.

Definition 3.6 (Moment of inertia tensor) *The linear operator $I : \mathbb{R}^3 \to \mathbb{R}^3$ defined as*

$$I(v) := \int_{\mathbb{R}^3} [\xi \times (v \times \xi)] \, dm$$

for all $v \in \mathbb{R}^3$ is called the rigid body's ***moment of inertia tensor****.*

Remark 3.7 *$I : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear operator, which means that it can also be expressed as a 3×3-matrix over $\mathbb{R}$. The preference for either form of I will depend on the particular case.*

Proposition 3.8 *The moment of inertia tensor of any rigid body is a symmetric positive definite linear operator, and the corresponding kinetic energy map $K : TSO(3) \to \mathbb{R}$ is given by*

$$K(V) = \frac{1}{2}\langle\langle V, V\rangle\rangle = \frac{1}{2}\langle\langle SA, SA\rangle\rangle = \frac{1}{2}\langle I\Omega, \Omega\rangle$$

for all $V \in T_S SO(3)$, $S \in SO(3)$ and $A \in \mathfrak{so}(3)$, where $V = SA$ and $\Omega = \Gamma(A)$.

The proof can be found in Section 3 of Chapter 5 of [1].

Proposition 3.9 *The matrix representation of the inertia tensor in the canonical basis of $\mathbb{R}^3$ is*

$$\begin{pmatrix} \int_{\mathbb{R}^3} (y^2 + z^2) \, dm & -\int_{\mathbb{R}^3} xy \, dm & -\int_{\mathbb{R}^3} xz \, dm \\ -\int_{\mathbb{R}^3} xy \, dm & \int_{\mathbb{R}^3} (x^2 + z^2) \, dm & -\int_{\mathbb{R}^3} yz \, dm \\ -\int_{\mathbb{R}^3} xz \, dm & -\int_{\mathbb{R}^3} yz \, dm & \int_{\mathbb{R}^3} (x^2 + y^2) \, dm \end{pmatrix}. \tag{3.3}$$

The proof can be found in Section 3 of Chapter 5 of [1].

Recall Proposition 3.8. Given a rigid body, the fact that its moment of inertia tensor is symmetric implies, by the Spectral Theorem, that it is diagonalisable by orthogonal matrices. Further, the fact that it is positive definite implies that its eigenvalues are positive. Therefore, we can define the following objects.

Definition 3.10 *Given a rigid body, there exist three positive numbers $I_1, I_2, I_3 \in \mathbb{R}^3$, called the* ***principal moments of inertia****, and an orthonormal basis of $\boldsymbol{R}^3$ $\{u_1, u_2, u_3\}$ that determines the so-called* ***principal axes****, such that $Iu_i = I_i u_i$ for $i \in \{1, 2, 3\}$.*

3.1.3 The Euler Equations

We can finally present the equations of motion of the Euler top $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle, 0)$, i.e. the geodesics of $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle)$. These are given by the so-called **Euler equations**.

Proposition 3.11 *The equations of motion of the Euler top* $(SO(3), \langle\langle \cdot, \cdot \rangle\rangle, 0)$ *are given by the so-called Euler equations*

$$I\dot{\Omega} = (I\Omega) \times \Omega. \tag{3.4}$$

Lemma 3.12 *We have* $p = SI\Omega$*, where* p *is the angular momentum of a rigid body,* $S : (a, b) \subseteq \mathbb{R} \to SO(3)$ *the curve describing its motion,* I *its inertia tensor and* Ω *is given by Remark 3.5.*

Proof The lemma follows from equation (3.2) and Definition 3.6. □

Proof (Euler Equations) By the above Lemma we have

$$p = SI\Omega.$$

Then, by Theorem 3.3, we have

$$\begin{aligned} 0 = \dot{p} = \dot{S}I\Omega + SI\dot{\Omega} = SAI\Omega + SI\dot{\Omega} \\ = S(AI\Omega + I\dot{\Omega}) = S(\Omega \times (I\Omega) + I\dot{\Omega}) \\ \Rightarrow 0 = \Omega \times (I\Omega) + I\dot{\Omega} \Leftrightarrow I\dot{\Omega} = (I\Omega) \times \Omega. \end{aligned}$$ □

Corollary 3.13 *Given a rigid body, let* I_1, I_2, I_3 *be the principal moments of inertia and* $\{u_1, u_2, u_3\}$ *be a basis of the principal axes. Further, let*

$$\Omega = \begin{pmatrix} \Omega^1 \\ \Omega^2 \\ \Omega^3 \end{pmatrix} \text{ and } \dot{\Omega} = \begin{pmatrix} \dot{\Omega}^1 \\ \dot{\Omega}^2 \\ \dot{\Omega}^3 \end{pmatrix}$$

be the representations of Ω *and* $\dot{\Omega}$ *in the basis* $\{u_1, u_2, u_3\}$*.*
Then, in the basis $\{u_1, u_2, u_3\}$ *of the principal axes, the Euler equations are*

$$\begin{cases} I_1\dot{\Omega}^1 = (I_2 - I_3)\Omega^2\Omega^3 \\ I_2\dot{\Omega}^2 = (I_3 - I_1)\Omega^3\Omega^1 \\ I_3\dot{\Omega}^3 = (I_1 - I_2)\Omega^1\Omega^2 \end{cases} . \tag{3.5}$$

3.2 The Poinsot Theorem

Having characterised the geodesics of the Euler top, we conclude the study of this mechanical system with a result that will imply the following phenomenon:
Given a freely moving rigid body with principal moments of inertia $I_1 > I_2 > I_3$, rotations of the rigid body about its first and third principal axes are stable, whereas rotations about its second principal axes are unstable. This is a well-known effect, called sometimes the Dzhanibekov effect or Tennis Racket Theorem.

Before proving the Poinsot Theorem, we make a detour into symmetries of general rigid bodies that will prove helpful to our main result. We take advantage of this necessary digression to show more properties of symmetries than are strictly necessary for our main purpose, as symmetries in themselves are an interesting aspect of the theory of rigid bodies.

Definition 3.14 (Symmetry) *A **symmetry** of a rigid body is an isometry $S \in O(3)$ which preserves the mass distribution (i.e. $m(SA) = m(A)$ for any measurable set $A \subseteq \mathbb{R}^3$).*

Proposition 3.15 *For $S \in O(3)$ a symmetry:*

1. *$SIS^T = I$, where I is the matrix representation of the inertia tensor;*
2. *if S is a reflection in a plane, then there exists a principal axis orthogonal to the reflection plane;*
3. *if S is a nontrivial rotation about an axis, then that axis is principal;*
4. *if moreover the rotation is not by π, then all axes orthogonal to the rotation axis are principal.*

Proof 1. $SIS^T = I$, where I is the matrix representation of the inertia tensor:
We will show that for all $(i,j) \in \{1,2,3\}^2$: $(SIS^T)_{ij} = I_{ij}$, which is equivalent to the proposition statement.
Let $\{e_i\}$ for $i \in \{1,2,3\}$ be the canonical basis of $\mathbb{R}^3$. We have, for $(i,j) \in \{1,2,3\}^2$ and $\langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{R}^3$:

$$\begin{aligned}(SIS^T)_{ij} &= \langle e_i, (SIS^T)e_j \rangle \\ &= \langle e_i, S\left(\int_{\mathbb{R}^3} [\xi \times ((S^T e_j) \times \xi)]\, dm\right)\rangle \\ &= \langle e_i, \int_{\mathbb{R}^3} S[\xi \times ((S^T e_j) \times \xi)]\, dm \rangle \end{aligned} \tag{3.6}$$

Now, recall that for A a 3×3 matrix over $\mathbb{R}$ and $x, y \in \mathbb{R}^3$, it holds:

$$(Ax) \times (Ay) = (detA)(A^{-1})^T (x \times y).$$

For $S \in O(3)$, the above identity simplifies to

$$(Sx) \times (Sy) = (\pm 1)(S^T)^T (x \times y) = \pm S(x \times y),$$

which applied to our computation yields

$$\begin{aligned} S[\xi \times ((S^T e_j) \times \xi)] &= \pm (S\xi) \times (S((S^T e_j) \times \xi)) \\ &= \pm (S\xi) \times (\pm (S(S^T e_j)) \times (S\xi)) = (S\xi) \times (e_j \times (S\xi)). \end{aligned}$$

Therefore, and using the assumption that S preserves the mass distribution, we get that

$$(3.6) = \langle e_i, \int_{\mathbb{R}^3} [(S\xi) \times (e_j \times (S\xi))] \, dm \rangle$$
$$= \langle e_i, \int_{\mathbb{R}^3} [\xi \times (e_j \times \xi)] \, dm \rangle = \langle e_i, I e_j \rangle = I_{ij}.$$

2. if S is a reflection in a plane then there exists a principal axis orthogonal to the reflection plane:
   Let S be a reflection on a plane P. Let $v \in \mathbb{R}^3 \setminus \{0\}$ such that $v \perp P \Leftrightarrow Sv = -v$. Then, by using the property we have already shown that $SIS^T = I$, we get for $u := Iv$:

   $$Su = SIv = SIS^T Sv = ISv = -Iv = -u$$
   $$\Rightarrow u \perp P \Leftrightarrow \exists \lambda \in \mathbb{R} \text{ such that } u = \lambda v.$$

   Finally, we have to show that $\lambda > 0$, but this holds because by Proposition 3.8 I is positive definite, i.e.

   $$0 < \langle Iv, v \rangle = \langle u, v \rangle = \lambda \langle v, v \rangle$$
   $$\Rightarrow \lambda > 0$$

3. if S is a nontrivial rotation about an axis then that axis is principal:
   Let S be a nontrivial rotation about an axis, let $v \in \mathbb{R}^3 \setminus \{0\}$ be a vector on that axis, which is equivalent to $Sv = v$ holding. Then, for $u := Iv$ and by using the property 1. that $SIS^T = I$ we get:

   $$Su = SIv = SIS^T Sv = ISv = Iv = u$$
   $$\Rightarrow u = Iv \text{ is also a vector on the rotation axis.}$$

   By an analogous reasoning to the proof of 2., we can then conclude that the rotation axis is principal.

4. if moreover the rotation is not by π then all axes orthogonal to the rotation axis are principal: Let $\{u_1, u_2, u_3\} \subseteq \mathbb{R}^3$ be an orthonormal basis of $\mathbb{R}^3$ of principal axes such that u_1 lies on the axis of rotation. Let $\{I_1, I_2, I_3\} \subseteq \mathbb{R}_{>0}$ be the corresponding principal moments of inertia (i.e. $Iu_i = I_i u_i$ for $i \in \{1,2,3\}$). Then, using the fact that $S^T \in O(3)$ preserves $\langle \cdot, \cdot \rangle$ and that $S^T u_1 = u_1$ since u_1 is on the axis of rotation, we have

   $$0 = \langle u_1, u_2 \rangle = \langle S^T u_1, S^T u_2 \rangle = \langle u_1, S^T u_2 \rangle \Rightarrow S^T u_2 \in span\{u_2, u_3\}.$$

Let $\alpha, \beta \in \mathbb{R}$ be such that $S^T u_2 = \alpha u_2 + \beta u_3$. Then, using the fact that S is not a rotation by $\pi \Rightarrow \beta \neq 0$ ($S^T u_2$ is not parallel to u_2), as well as the property 1. that $SIS^T = I \Leftrightarrow IS^T = S^T I$:

$$\begin{aligned} Iu_3 = I\left(\frac{S^T u_2 - \alpha u_2}{\beta}\right) &= \frac{1}{\beta}(IS^T u_2 - \alpha I u_2) = \frac{1}{\beta}(S^T I u_2 - \alpha I u_2) \\ &= \frac{1}{\beta}(S^T I_2 u_2 - \alpha I_2 u_2) = I_2\left(\frac{S^T u_2 - \alpha u_2}{\beta}\right) = I_2 u_3 \\ &\Rightarrow \{u_2, u_3\} \in Eig_{I_2}(I) \Rightarrow span\{u_2, u_3\} \in Eig_{I_2}(I), \end{aligned}$$

where $Eig_{I_2}(I)$ is the eigenspace of I to its eigenvalue I_2. Property 4. follows from the last implication. $\square$

We are now ready to prove the Poinsot Theorem.

Definition 3.16 (Inertia ellipsoid) *The* ***inertia ellipsoid*** *of a rigid body with moment of inertia tensor I is the set*

$$E = \left\{\xi \in \mathbb{R}^3 : \langle I\xi, \xi\rangle = 1\right\}. \tag{3.7}$$

Theorem 3.17 (Poinsot Theorem) *The inertia ellipsoid of a freely moving rigid body rolls without slipping on a fixed plane orthogonal to the angular momentum p.*

Proof (Poinsot Theorem) Let P be a plane orthogonal to p. For $u, v \in P$, we have that

$$\langle u - v, p\rangle = 0 \Leftrightarrow \langle u, p\rangle = \langle v, p\rangle$$

$$\Rightarrow \forall\lambda \in \mathbb{R} : P_\lambda := \{u \in \mathbb{R}^3 : \langle u, p\rangle = \lambda\} \text{ is a plane orthogonal to } p.$$

Let $S\xi$ be a point where E is tangent to P_λ for some $\lambda \in \mathbb{R}$. Then, and because of Lemma 3.12,

$$\langle SI\Omega, S\xi\rangle = \lambda; \tag{3.8}$$

$$\langle IS\xi, S\xi\rangle = 1. \tag{3.9}$$

We assume that $S \in SO(3)$ is a symmetry (i.e. it preserves the mass distribution), so that the assumptions of Proposition 3.15 are met, and thus, by statement 1. $SI = IS$

$$\Rightarrow \xi = \frac{1}{\lambda}\Omega$$

is a solution to equations (3.8) and (3.9) (assuming $\lambda \neq 0$). To prove that $S\frac{1}{\lambda}\Omega$ is indeed a point where E is tangent to P_λ (and not only a point where E and

P_λ intersect without being tangent), we use the fact that a normal vector of a plane tangential to E at $S\xi \in E$ is a constant multiple of the gradient of E at $S\xi$. In our case, this means that the gradient at $S\xi$ should be a multiple of $p = SI\Omega$, and, indeed,

$$\begin{aligned}\nabla\langle IS\frac{1}{\lambda}\Omega, S\frac{1}{\lambda}\Omega\rangle &= \frac{1}{\lambda^2}\nabla\langle IS\Omega, S\Omega\rangle = \frac{1}{\lambda^2}\nabla\langle I\omega,\omega\rangle \\ &= \frac{1}{\lambda^2}\nabla(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) = \frac{2}{\lambda^2}I\omega = \frac{2}{\lambda^2}IS\Omega = \frac{2}{\lambda^2}SI\Omega = \frac{2}{\lambda^2}p,\end{aligned}$$

where we have used the representation of I and ω in the principal axes basis, and we have again used Proposition 3.15 to commute I and S. So far, we know that $\xi = \frac{1}{\lambda}\Omega$ is the tangent point of E with the plane P_λ. To justify all the above results, we need to justify the assumption we made that $\lambda \neq 0$. A posteriori, this holds:

$$\begin{aligned}\lambda = \langle SI\Omega, S\xi\rangle = \langle SI\Omega, S\frac{1}{\lambda}\Omega\rangle = \frac{1}{\lambda}\langle I\Omega,\Omega\rangle = \frac{1}{\lambda}2K \\ \Rightarrow \lambda^2 = 2K \Rightarrow \lambda = \pm\sqrt{2K}.\end{aligned}$$

The above result implies that $S\xi = \pm\frac{1}{\sqrt{2K}}S\Omega = \pm\frac{1}{\sqrt{2K}}\omega$ are the two points where E is tangent to a plane orthogonal to p, and all the previous calculations where we assumed that $\lambda \neq 0$ hold.
To conclude the proof, we show that for these points the velocity is zero, which will then imply the statement of the theorem. For every given $\xi_0 \in \mathbb{R}^3$ and for $A \in \mathfrak{so}(3)$ such that $\dot{S}\xi_0 = SA\xi_0 = S(\Omega \times \xi_0)$ (*cf.* Remark 3.5)

$$\frac{d}{dt}(S\xi_0) = \dot{S}\xi_0 = SA\xi_0 = S(\Omega \times \xi_0).$$

For $\xi_0 = \pm\frac{1}{\sqrt{2K}}\Omega$ the above right hand side vanishes, and the Theorem follows. □

Remark 3.18 *In particular, the proof of Poinsot's Theorem implies that whenever the rigid body rotates about one of its principal axes, the angular velocity vector ω precesses about the principal axes. However, rotations about the first and last principal axes (the ones corresponding to I_1 and I_3 whenever $I_1 > I_2 > I_3$) are stable, whereas rotations about the second are not.*

3.3 Properties of a Rigid Body with Two Equal Principal Moments of Inertia

We know leave the particular case of the Euler top, and consider another type of special rigid body: one that satisfies $I_1 = I_2$. We want to provide an expression for its kinetic energy using local coordinates of $TSO(3)$. To this end, we introduce the so-called Euler angles.

Definition 3.19 (Euler angles) *The **Euler angles** correspond to the local coordinates* $(\theta, \varphi, \psi) : SO(3) \to (0, \pi) \times (0, 2\pi) \times (0, 2\pi)$ *associated with the local parametrisation* $S : (0, \pi) \times (0, 2\pi) \times (0, 2\pi) \to SO(3)$ *defined by*

$$S(\theta, \varphi, \psi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Proposition 3.20 *If* $I_1 = I_2$, *then the kinetic energy of a rigid body in the local coordinates* $(\theta, \varphi, \psi, v^\theta, v^\varphi, v^\psi)$ *of* $TSO(3)$ *is given by*

$$K = \frac{I_1}{2}((v^\theta)^2 + (v^\varphi)^2 \sin^2\theta) + \frac{I_3}{2}(v^\psi + v^\varphi \cos\theta)^2. \tag{3.10}$$

Proof Let $S : (a, b) \subseteq \mathbb{R} \to SO(3)$ be the curve characterising the motion of the rigid body. For $\Omega = \Omega(t)$ as in Remark 3.5, we know, by Proposition 3.8 that

$$K = \frac{1}{2}\langle I\Omega, \Omega\rangle. \tag{3.11}$$

If we set $\Omega^1, \Omega^2, \Omega^3 \in \mathbb{R}$ such that $\Omega = \begin{pmatrix} \Omega^1 \\ \Omega^2 \\ \Omega^3 \end{pmatrix}$ is the representation of Ω in the principal axes coordinate system, we have that

$$\langle I\Omega, \Omega\rangle = I_1(\Omega^1)^2 + I_2(\Omega^2)^2 + I_3(\Omega^3)^2. \tag{3.12}$$

Now, observe that if S is the curve characterising the motion of the rigid body with respect to the standard basis of $\mathbb{R}^3$, then performing a change of coordinates to determine the curve characterising the motion of the rigid body with respect to the principal axes coordinate system amounts to left-multiplication of S by another $\bar{S} \in SO(3)$, which results in a new curve $\bar{S}SI \subseteq \mathbb{R} \to SO(3)$. But since this curve is also a subset of $SO(3)$, we can just set $S = \bar{S}S$, and then we will be working in the principal axes coordinate system. What we have to do to prove the result, then, is to find $\Omega = \Gamma(A)$ for the $A \in \mathfrak{so}(3)$ such that $\dot{S} = SA$ holds, which we will then insert into equation (3.12).
By using the parametrisation of $S \in SO(3)$ given by the Euler angles in

Definition 3.19, we get that

$$
\begin{aligned}
&\dot{S}(\theta, \varphi, \psi) \\
&= \dot{\varphi} \begin{pmatrix} -\sin\varphi & -\cos\varphi & 0 \\ \cos\varphi & -\sin\varphi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&+ \dot{\theta} \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin\theta & -\cos\theta \\ 0 & \cos\theta & -\sin\theta \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&+ \dot{\psi} \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -\sin\psi & -\cos\psi & 0 \\ \cos\psi & -\sin\psi & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&=: \dot{\varphi} A_{\dot{\varphi}} B_{\dot{\varphi}} C_{\dot{\varphi}} + \dot{\theta} A_{\dot{\theta}} B_{\dot{\theta}} C_{\dot{\theta}} + \dot{\psi} A_{\dot{\psi}} B_{\dot{\psi}} C_{\dot{\psi}},
\end{aligned}
$$

and also

$$
\begin{aligned}
&S^{-1} = S^T \\
&= \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&=: A_S B_S C_S.
\end{aligned}
$$

Our goal is to compute

$$
A = S^{-1}\dot{S} = S^T\dot{S} = A_S B_S C_S (\dot{\varphi} A_{\dot{\varphi}} B_{\dot{\varphi}} C_{\dot{\varphi}} + \dot{\theta} A_{\dot{\theta}} B_{\dot{\theta}} C_{\dot{\theta}} + \dot{\psi} A_{\dot{\psi}} B_{\dot{\psi}} C_{\dot{\psi}}),
$$

which we will do in intermediate steps:

$$
C_S A_{\dot{\varphi}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow B_S C_S A_{\dot{\varphi}} B_{\dot{\varphi}} = \begin{pmatrix} 0 & -\cos\theta & \sin\theta \\ \cos\theta & 0 & 0 \\ -\sin\theta & 0 & 0 \end{pmatrix}
$$

$$
\Rightarrow A_S B_S C_S A_{\dot{\varphi}} B_{\dot{\varphi}} C_{\dot{\varphi}} = \begin{pmatrix} 0 & -\cos\theta & \sin\theta \\ \cos\theta & 0 & 0 \\ -\sin\theta & 0 & 0 \end{pmatrix};
$$

$$
C_S A_{\dot{\theta}} = Identity \Rightarrow B_S C_S A_{\dot{\theta}} B_{\dot{\theta}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}
$$

$$\Rightarrow A_S B_S C_S A_{\dot\varphi} B_{\dot\varphi} C_{\dot\varphi} = \begin{pmatrix} 0 & -\cos\theta & \sin\theta \\ \cos\theta & 0 & 0 \\ -\sin\theta & 0 & 0 \end{pmatrix};$$

$$C_S A_{\dot\theta} = Identity \Rightarrow B_S C_S A_{\dot\theta} B_{\dot\theta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\Rightarrow A_S B_S C_S A_{\dot\theta} B_{\dot\theta} C_{\dot\theta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix};$$

$$C_S A_{\dot\psi} = Identity \Rightarrow B_S C_S A_{\dot\psi} B_{\dot\psi} = Identity$$

$$\Rightarrow A_S B_S C_S A_{\dot\psi} B_{\dot\psi} C_{\dot\psi} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

The above results put together give us

$$\begin{aligned} A &= S^{-1}\dot S = S^T \dot S \\ &= \dot\varphi \begin{pmatrix} 0 & -\cos\theta & \sin\theta \\ \cos\theta & 0 & 0 \\ -\sin\theta & 0 & 0 \end{pmatrix} + \dot\theta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + \dot\psi \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\dot\varphi\cos\theta - \dot\psi & \dot\varphi\sin\theta \\ \dot\varphi\cos\theta + \dot\psi & 0 & -\dot\theta \\ -\dot\varphi\sin\theta & \dot\theta & 0 \end{pmatrix} \\ &\Rightarrow \Omega = \begin{pmatrix} \dot\theta \\ \dot\varphi\sin\theta \\ \dot\varphi\cos\theta + \dot\psi \end{pmatrix}. \end{aligned}$$

Inserting the above result into equation (3.12), and using the assumption that $I_1 = I_2$ we get

$$\langle I\Omega, \Omega\rangle = I_1(\dot\theta^2 + \dot\varphi^2 \sin^2\theta) + I_3(\dot\varphi\cos\theta + \dot\psi)^2,$$

which, together with equation (3.11) gives us the desired result. □

Let a rigid body's motion be described by the curve $S : (a, b) \subseteq \mathbb{R} \to SO(3)$. For $\Omega = \Omega(t)$ as in Remark 3.5, we can introduce the following object.

Definition 3.21 (Instantaneous angular velocity) *A rigid body's* ***instantaneous angular velocity*** *is the vector $\omega := S\Omega$, which determines the axis about which the rigid body rotates at any given time, with angular velocity $\|\Omega\|$.*

Remark 3.22 *By the above definition, Ω is the angular velocity as seen in the rigid body's rest frame, and so it determines the axis about which the rigid body rotates in its rest frame.*

We conclude our study of rigid bodies satisfying $I_1 = I_2$ with the following result about the variation of their angular velocity ω.

Proposition 3.23 *1. The angular velocity of a rigid body with $I_1 = I_2$ satisfies*

$$\dot{\omega} = \frac{1}{I_1} p \times \omega; \tag{3.13}$$

2. *if $I_1 = I_2 = I_3$, then the rigid body rotates about a fixed axis with constant angular speed (i.e. ω is constant).*

Proof Let $S : (a, b) \subseteq \mathbb{R} \to SO(3)$ be the curve characterising the motion of the rigid body. Set $\Omega = \Omega(t)$ as in Remark 3.5.

1. By the Euler equations in their general form (3.4) and their representation in the basis of the principal axes (3.5), and considering that in the basis of the principal axes

$$I^{-1} = \begin{pmatrix} \frac{1}{I_1} & 0 & 0 \\ 0 & \frac{1}{I_2} & 0 \\ 0 & 0 & \frac{1}{I_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{I_1} & 0 & 0 \\ 0 & \frac{1}{I_1} & 0 \\ 0 & 0 & \frac{1}{I_3} \end{pmatrix}, \tag{3.14}$$

we have that

$$\begin{aligned} \dot{\Omega} = I^{-1}((I\Omega) \times \Omega) &= I^{-1} \begin{pmatrix} (I_2 - I_3)\Omega^2\Omega^3 \\ (I_3 - I_1)\Omega^3\Omega^1 \\ 0 \end{pmatrix} \\ &= \frac{1}{I_1} \begin{pmatrix} (I_2 - I_3)\Omega^2\Omega^3 \\ (I_3 - I_1)\Omega^3\Omega^1 \\ 0 \end{pmatrix} = \frac{1}{I_1}((I\Omega) \times \Omega). \end{aligned} \tag{3.15}$$

Observe that $SI\Omega = p$ by Lemma 3.12. Together with the fact that $S \in SO(3)$ preserves the cross product $\times$ in $\mathbb{R}^3$, and by equation (3.15), we have that

$$S\dot{\Omega} = \frac{1}{I_1}((SI\Omega) \times (S\Omega)) = \frac{1}{I_1}(p \times \omega). \tag{3.16}$$

Also, we have that

$$\dot{S}\Omega = SA\Omega = S(\Omega \times \Omega) = S0 = 0. \tag{3.17}$$

Combining equations (3.16) and (3.17) we get

$$\dot{\omega} = \dot{S}\Omega + S\dot{\Omega} = \frac{1}{I_1}(p \times \omega).$$

2. If $I_1 = I_2 = I_3$, then

$$(I\Omega) \times \Omega = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \dot{\Omega} = I^{-1}((I\Omega) \times \Omega) = 0$$

$$\Rightarrow \dot{\omega} = \dot{S}\Omega + S\dot{\Omega} = 0 \quad \square$$

3.4 Fictitious Forces

To close this chapter, we derive the so-called **fictitious forces**, which appear to act on particles when viewed in a non-inertial frame of reference, as is the case on Earth's surface, for instance. These are the **centrifugal force** F_{Cent}, the **Coriolis force** F_{Cor} and the **Euler force** F_E. To this end, let us consider a rigid body whose motion is described by $S : \mathbb{R} \to SO(3)$, and a particle with mass m whose motion in the rigid body's rest frame is given by $\xi : \mathbb{R} \to \mathbb{R}^3$. For $\Omega = \Omega(t)$ as in Remark 3.5 (observe that, by Remark 3.22, Ω is also the rigid body's angular velocity in its rest frame), the fictitious forces are defined as

$$F_{Cent} = -m\Omega \times (\Omega \times \xi); \quad F_{Cor} = -2m\Omega \times \dot{\xi}; \quad F_E = -m\dot{\Omega} \times \xi.$$

For F the external force on the particle as seen in the rigid body's rest frame, the equation of motion of the particle is then

$$m\frac{d^2}{dt^2}(S\xi) = SF. \tag{3.18}$$

We then get:

$$\begin{aligned}
\frac{d^2}{dt^2}(S\xi) &= \frac{d}{dt}(\dot{S}\xi) + \frac{d}{dt}(S\dot{\xi}) = \frac{d}{dt}(S(\Omega \times \xi)) + \dot{S}\dot{\xi} + S\ddot{\xi} \\
&= \dot{S}(\Omega \times \xi) + S(\dot{\Omega} \times \xi) + S(\Omega \times \dot{\xi}) + S(\Omega \times \dot{\xi}) + S\ddot{\xi} \\
&= S(\Omega \times (\Omega \times \xi)) + S(\dot{\Omega} \times \xi) + 2S(\Omega \times \dot{\xi}) + S\ddot{\xi} \\
&\Rightarrow SF = m\frac{d^2}{dt^2}(S\xi) \\
&= mS[(\Omega \times (\Omega \times \xi)) + (\dot{\Omega} \times \xi) + 2(\Omega \times \dot{\xi}) + \ddot{\xi}] =: mSv \\
&\Rightarrow mv = mS^T Sv = F \\
&\Rightarrow m\ddot{\xi} = F - m(\Omega \times (\Omega \times \xi)) - 2m(\Omega \times \dot{\xi}) - m(\dot{\Omega} \times \xi),
\end{aligned}$$

and so we have derived the equation of motion of ξ in the rigid body's rest frame, which includes the fictitious forces.

Claim 3.24 *If the rigid body is a homogeneous sphere rotating freely (like the Earth, to a good degree of approximation, for instance) then the Euler force vanishes.*

Proof Recall that the Euler force is defined as $F_E = -m\dot{\Omega} \times \xi$ (*cf.* section 3.4). We will show that $\dot{\Omega} = 0$.
By Proposition 3.3, the matrix representation of the inertia tensor I of the rigid body in the canonical basis of $\mathbb{R}^3$ is

$$I = \begin{pmatrix} \int_{\mathbb{R}^3}(y^2+z^2)\,dm & -\int_{\mathbb{R}^3} xy\,dm & -\int_{\mathbb{R}^3} xz\,dm \\ -\int_{\mathbb{R}^3} xy\,dm & \int_{\mathbb{R}^3}(x^2+z^2)\,dm & -\int_{\mathbb{R}^3} yz\,dm \\ -\int_{\mathbb{R}^3} xz\,dm & -\int_{\mathbb{R}^3} yz\,dm & \int_{\mathbb{R}^3}(x^2+y^2)\,dm \end{pmatrix}. \tag{3.19}$$

If the rigid body is a homogeneous sphere rotating freely, then

$$\int_{\mathbb{R}^3} xy\,dm = \int_{\mathbb{R}^3} xz\,dm = \int_{\mathbb{R}^3} yz\,dm = 0,$$

due to symmetry. Further, and again due to symmetry,

$$\int_{\mathbb{R}^3}(y^2+z^2)\,dm = \int_{\mathbb{R}^3}(x^2+z^2)\,dm = \int_{\mathbb{R}^3}(x^2+y^2)\,dm =: \alpha \in \mathbb{R}.$$

Therefore,

$$I = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}. \tag{3.20}$$

And so we conclude that all principal moments of inertia are equal ($I_1 = I_2 = I_3 = \alpha$). By the second case of Proposition 3.23 and with $A \in \mathfrak{so}(3)$ such that $\dot{S}\xi = SA\xi = S(\Omega \times \xi)$ (*cf.* Remark 3.5) , we get that

$$0 = \dot{\omega} = \dot{S}\Omega + S\dot{\Omega} = SA\Omega + S\dot{\Omega} = S(\Omega \times \Omega) + S\dot{\Omega} = 0 + S\dot{\Omega}$$
$$(S \in S0(3) \Rightarrow S \neq 0) \Rightarrow \dot{\Omega} = 0. \quad \square$$

Chapter 4

Non-holonomic Constraints

In chapter 2, we studied systems of particles whose motions (and, therefore, whose positions) are restricted to a subspace of the configuration space. Non-holonomic constraints account for restrictions to the direction in which particles of a given system are allowed to move. That is, non-holonomic constraints impose certain conditions on the velocities of the particles of a given system. Examples of such systems are given by a wheel rolling without slipping on a plane or an ice skate that can only move forwards or backwards along the line on which the skate lies, or rotate about its middle point, for instance. Analogously to the case of holonomic constraints in chapter 2, we need an additional term in the Newton equation to impose the non-holonomic constraints on the system (which is known by the same name: reaction force). In the case of the ice skate, for example, the reaction force can be seen as a friction force preventing the skate from sliding sideways.
In the present chapter, we introduce the necessary concepts to consider non-holonomic constraints on a system. We also study the relationship between holonomic and non-holonomic constraints. Finally, we consider the example of a system with non-holonomic constraints given by an ice skate.

4.1 Non-holonomic Constraints and Reaction Forces

4.1.1 Differentiable Distributions and Non-holonomic Constraints

Definition 4.1 (Distribution) *A **distribution** Σ of dimension m on an n-dimensional differentiable manifold M is a choice of an m-dimensional subspace $\Sigma_p \subseteq T_pM$ for each $p \in M$. The distribution is said to be **differentiable** if for all $p \in M$ there exists a neighbourhood U containing p and differentiable vector fields $X_1, ..., X_m$ on U such that*

$$\Sigma_q = span\{(X_1)_q, ..., (X_m)_q\}$$

for all $q \in U$.

Proposition 4.2 *An m-dimensional distribution Σ on an n-dimensional differentiable manifold M is differentiable if and only if for all $p \in M$ there exist a neighbourhood U containing p and 1-forms $\omega^1, ..., \omega^{n-m} \in \Omega^1(U)$ such that*

$$\Sigma_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q$$

for all $q \in U$, where $\Omega^1(U)$ denotes the vector space of 1-forms on U.

We will need the following lemma.

Lemma 4.3 *Let M be a differentiable manifold of dimension n, let $p \in M$ be given, let U be a neighbourhood containing p. Let $\ell \in \{0, ..., n-1\}$, let $\tau^1, ..., \tau^{n-\ell} \in \Omega^1(U)$. Let $q \in U$ be arbitrary. If $\{(\tau^1)_q, ..., (\tau^{n-\ell})_q\}$ are linearly independent, then $dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-\ell})_q) = \ell$.*

Proof We prove the lemma by induction on the number of 1-forms, which is equivalent to backward induction on $\ell \in \{0, ..., n-1\}$:

1. Induction start, $\ell = n-1$: $\{(\tau^1)_q\}$ being linearly independent means that the 1-form is not 0, which implies that $dim(ker(\tau^1)_q) = n-1$.
2. Induction step: Assume the lemma holds for $\ell+1$, where $\ell \in \{0, ..., n-2\}$ is given. Then for ℓ we have that

$$dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-\ell})_q) = dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-(\ell+1)})_q) + dim(ker(\tau^{n-\ell})_q) - dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-(\ell+1)})_q + ker(\tau^{n-\ell})_q).$$

   By induction assumption $dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-(\ell+1)})_q) = \ell+1$. Further, since $\{(\tau^1)_q, ..., (\tau^{n-\ell})_q\}$ are linearly independent, in particular $(\tau^{n-\ell})_q \neq 0 \Rightarrow dim(ker(\tau^{n-\ell})_q) = n-1$. Therefore, the statement will follow if $dim(ker(\tau^1)_q \cap ... \cap ker(\tau^{n-(\ell+1)})_q + ker(\tau^{n-\ell})_q) = n$. To show this, let $v \in T_qM$. Since $dim(T_qM) = n$ and $dim(ker(\tau^{n-\ell})_q) = n-1$, we can choose $v_1 \in T_qM \setminus ker(\tau^{n-\ell})_q$. Then $T_qM = ker(\tau^{n-\ell})_q + span\{v_1\}$. Assume $(\tau^i)_q(v_1) \neq 0$ for some $i \in \{1, ..., n-(\ell+1)\}$. Then $ker(\tau^i)_q = ker(\tau^{n-\ell})_q$ and $(\tau^i)_q = \alpha(\tau^{n-\ell})_q$ for $\alpha := \frac{(\tau^i)_q(v_1)}{(\tau^{n-\ell})_q(v_1)}$, i.e. the set $\{(\tau^1)_q, ..., (\tau^{n-\ell})_q\}$ would not be linearly independent, a contradiction. Therefore $v_1 \in ker(\tau^1)_q \cap ... \cap ker(\tau^{n-(\ell+1)})_q$ and the statement follows. □

Proof (Proposition) $\underline{\Rightarrow}$: Let $p \in M$ be arbitrary. Assume that Σ is differentiable. Then there exists a neighbourhood U containing p and differentiable vector fields $X_1, ..., X_m$ on U such that

$$\Sigma_q = span\{(X_1)_q, ..., (X_m)_q\}$$

for all $q \in U$. Let $q \in U$. We show that there exist 1-forms $\omega^1, ..., \omega^{n-m} \in \Omega^1(U)$ such that

$$\Sigma_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q.$$

$\Sigma_q \subseteq T_qM$ is m-dimensional, and if a set of m vectors $\{(X_1)_q, ..., (X_m)_q\}$ spans an m-dimensional vector space, then it is a basis. This holds for all $q \in U$, and so we can complete $\{X_1, ..., X_m\}$ to a moving frame on U $\{X_1, ..., X_m, Y_1, ..., Y_{n-m}\}$, so that $\{(X_1)_q, ..., (X_m)_q, (Y_1)_q, ..., (Y_{n-m})_q\}$ is a basis of T_qM for all $q \in U$. Take the dual frame $\{\tau^1, ..., \tau^m, \omega^1, ..., \omega^{n-m}\}$ made up of n 1-forms on U such that they are linearly independent at each $q \in U$ and so that $\forall q \in U$:

$$\begin{cases} (\tau^i)_q((X_j)_q) = \delta^i_j \text{ for all } (i,j) \in \{1, ..., m\}^2 \\ (\omega^i)_q((Y_j)_q) = \delta^i_j \text{ for all } (i,j) \in \{1, ..., n-m\}^2 \\ (\tau^i)_q((Y_j)_q) = (\omega^j)_q((X_i)_q) = 0 \text{ for all } (i,j) \in \{1, ..., m\} \times \{1, ..., n-m\} \end{cases}$$

1. $\Sigma_q \subseteq ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q$:

$$\begin{aligned} v \in \Sigma_q \Leftrightarrow v = \sum_{i=1}^{m} \lambda_i (X_i)_q \text{ for some } \lambda_i \in \mathbb{R} \\ \Rightarrow (\omega^j)_q(v) = \sum_{i=1}^{m} \lambda_i (\omega^j)_q((X_i)_q) = 0 \text{ for all } j \in \{1, ..., n-m\} \\ \Rightarrow v \in ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q. \end{aligned}$$

2. $ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q \subseteq \Sigma_q$:

$$v \in T_qM \Leftrightarrow v = \sum_{i=1}^{m} \lambda_i (X_i)_q + \sum_{j=1}^{n-m} \mu_j (Y_j)_q \text{ for some } \lambda_i, \mu_j \in \mathbb{R}.$$

$$\begin{aligned} v \in ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q \Rightarrow 0 = (\omega^k)_q(v) \\ = \sum_{i=1}^{m} \lambda_i (\omega^k)_q((X_i)_q) + \sum_{j=1}^{n-m} \mu_j (\omega^k)_q((Y_j)_q) = \mu_k \\ \text{for all } k \in \{1, ..., n-m\} \\ \Rightarrow v = \sum_{i=1}^{m} \lambda_i (X_i)_q, \text{ i.e. } v \in \Sigma_q. \end{aligned}$$

$\underline{\Leftarrow}$: Again, let $p \in M$ be arbitrary. Assume there exist a neighbourhood U containing p and 1-forms $\omega^1, ..., \omega^{n-m} \in \Omega^1(U)$ such that

$$\Sigma_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q$$

for all $q \in U$. Let $q \in U$.

Claim 4.4 *The set of* 1*-forms* $\{(\omega^1)_q, ..., (\omega^{n-m})_q\}$ *is linearly independent.*

Proof We prove the claim by induction on the number of 1-forms, which is equivalent to backward induction on the dimension m of the distribution ($m \in \{1, ..., n-1\}$):

1. Induction start, $m = n-1$: $\{(\omega^1)_q\}$ is linearly independent if and only if $(\omega^1)_q \neq 0$, which holds because, by assumption, $n-1 = dim(\Sigma_q) = dim(ker(\omega^1)_q)$.
2. Induction step: Assume the claim holds for $m+1$, where $m \in \{1, ..., n-2\}$ is given. If the claim were not to hold, and because by assumption $\{(\omega^1)_q \cap ... \cap (\omega^{n-(m+1)})_q\}$ is linearly independent, we would have that $(\omega^{n-m})_q = \sum_{k=1}^{n-(m+1)} \lambda_k (\omega^k)_q$ for some $\lambda_k \in \mathbb{R}$, which would imply that $ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-(m+1)})_q$. However, by Lemma 4.3 we would have

$$\begin{aligned} m = dim(\Sigma_q) &= dim(ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q) \\ &= dim(ker(\omega^1)_q \cap ... \cap ker(\omega^{n-(m+1)})_q) = m+1, \end{aligned}$$

   a contradiction. □

Since $\{(\omega^1)_q, ..., (\omega^{n-m})_q\}$ is linearly independent for all $q \in U$, we can complete the set $\{\omega^1, ..., \omega^{n-m}\}$ to a local co-frame $\{\tau^1, ..., \tau^m, \omega^1, ..., \omega^{n-m}\}$. Let $\{X_1, ..., X_m, Y_1, ..., Y_{n-m}\}$ be the dual moving frame. By an analogous argument to the one used for the *only if* direction, we have that $\Sigma_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q \subseteq span\{(X_1)_q, ..., (X_m)_q\}$ for all $q \in U$. Finally, since both sub-spaces of T_qM are m-dimensional, equality holds. □

From this point on, we assume that all distributions are differentiable.

Definition 4.5 *A* ***non-holonomic constraint*** *on a mechanical system* $(M, \langle\cdot,\cdot\rangle, \mathcal{F})$ *is a distribution* Σ *on* M*. A curve* $c : (a, b) \subseteq \mathbb{R} \to M$ *is said to be* ***compatible*** *with* Σ *if* $\dot{c}(t) \in \Sigma_{c(t)}$ *for all* $t \in (a, b)$*.*

4.1.2 Reaction Forces

Similarly to the case of holonomic constraints, non-holonomic constraints are accounted for in the Newton equation (1.1) by adding the reaction force.
In a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$, any distribution Σ determines an **orthogonal distribution** $\Sigma^\perp$, given by

$$(\Sigma^\perp)_p = (\Sigma_p)^\perp \subseteq T_pM$$

for all $p \in M$. Therefore, we can consider the two orthogonal projections $^T : TM \to \Sigma$ and $^\perp : TM \to \Sigma^\perp$. The set of all external forces $\mathcal{F} : TM \to T^*M$ satisfying

$$\mathcal{F}(v) = \mathcal{F}(v^T)$$

for all $v \in TM$ is denoted by F_Σ.

Definition 4.6 *A **reaction force** on a mechanical system with non-holonomic constraints $(M, \langle\cdot,\cdot\rangle, \mathcal{F}, \Sigma)$ is a force $\mathcal{R} \in F_\Sigma$ such that the solutions of the **generalised Newton equation***

$$\mu\left(\frac{D\dot{c}}{dt}\right) = (\mathcal{F} + \mathcal{R})(\dot{c})$$

*with initial condition in Σ are compatible with Σ. The reaction force is said to be **perfect**, or to satisfy the **d'Alembert principle**, if*

$$\mu^{-1}(\mathcal{R}(v)) \in \Sigma_p^\perp$$

for all $v \in T_pM$ and for all $p \in M$.

Consider a solution to the generalised Newton equation. An analogous computation to (2.1) (and the subsequent explanation) for the variation of its kinetic energy, together with Definition 4.6, allows us to conclude that a perfect reaction force neither creates nor dissipates kinetic energy.
We have the following theorem about the existence and uniqueness of perfect reaction forces (analogously to the case of holonomic constraints, cf. chapter 2).

Theorem 4.7 *Given a mechanical system with non-holonomic constraints $(M, \langle\cdot,\cdot\rangle, \mathcal{F}, \Sigma)$, there exists a unique reaction force $\mathcal{R} \in F_\Sigma$ satisfying the d'Alembert principle.*

The proof can be found in Section 4 of Chapter 5 of [1].

4.2 Integrable distributions

Given the similarities in the objects we have defined for holonomic and non-holonomic constraints, a very reasonable question is what the relationship between these types of constraints might be. To answer, we must consider so-called integrable distributions.

Definition 4.8 *A **foliation** of dimension m on an n-dimensional differentiable manifold M is a family $\mathcal{F} = \{L_\alpha\}_{\alpha\in A}$ of subsets of M (called **leaves**) satisfying:*

1. $M = \cup_{\alpha\in A} L_\alpha$;
2. $L_\alpha \cap L_\beta = \varnothing$ *if* $\alpha \neq \beta$;
3. *each leaf L_α is pathwise connected, i.e. $\forall p, q \in L_\alpha \exists c : [0,1] \to L_\alpha$ a continuous curve such that $c(0) = p$ and $c(1) = q$;*

4. *for each point $p \in M$ there exists an open set U containing p and local coordinates $(x^1, ..., x^n) : U \to \mathbb{R}^n$ such that the connected components of the intersections of the leaves with U are the level sets of $(x^{m+1}, ..., x^n) : U \to \mathbb{R}^{n-m}$.*

Definition 4.9 *An m-dimensional distribution Σ on a differentiable manifold M is said to be* ***integrable*** *if there exists an m-dimensional foliation $\mathcal{F} = \{L_\alpha\}_{\alpha \in A}$ on M such that*

$$\Sigma_p = T_p(L_p)$$

for all $p \in M$, where L_p is the leaf containing p. The leaves of $\mathcal{F}$ are called the ***integral submanifolds*** *of the distribution.*

Integral distributions are particularly simple. For instance, the set of points $q \in M$ which are accessible from a given point $p \in M$ by a curve compatible with the distribution is simply the leaf L_p containing p. If the leaves are embedded submanifolds of M, then an integrable non-holonomic constraint is equivalent to a set of holonomic constraints on the mechanical system. This is the reason why integrable distributions are also known as **semi-holonomic constraints**, while non-integrable ones are known as **true non-holonomic constraints**. We are therefore interested in identifying integrable distributions. The Frobenius Theorem provides a necessary and sufficient condition.

Definition 4.10 *Let Σ be a distribution on a differentiable manifold M. A differentiable vector field X on M is said to be* ***compatible*** *with Σ if $X_p \in \Sigma_p$ for all $p \in M$.*

Definition 4.11 *A distribution Σ on a differentiable manifold M is said to be* ***involutive*** *if for any two differentiable vector fields X, Y on M that are compatible with Σ, $[X, Y]$ is also compatible with Σ.*

Theorem 4.12 (Frobenius) *A distribution Σ on a differentiable manifold M is integrable if and only if it is involutive.*

The proof can be found in Section 3 of Chapter 2 of [3].

Proposition 4.13 *An m-dimensional distribution Σ on an n-dimensional manifold M is integrable if and only if $\forall i \in \{1, ..., n-m\}$:*

$$d\omega^i \wedge \omega^1 \wedge ... \wedge \omega^{n-m} = 0$$

for all locally defined sets of differential forms $\{\omega^1, ..., \omega^{n-m}\}$ whose kernels determine Σ.

Proof Let $p \in M$ be arbitrary, let U be a neighbourhood containing p, let $q \in U$. In the proof of the *if* direction of Proposition 4.2, we argued that if

$$\Sigma_q = ker(\omega^1)_q \cap ... \cap ker(\omega^{n-m})_q,$$

then $\{(\omega^1)_q, ..., (\omega^{n-m})_q\}$ is a linearly independent set for all $q \in U$, and so we can complete $\{\omega^1, ..., \omega^{n-m}\}$ to a local coframe $\{\tau^1, ..., \tau^m \omega^1, ..., \omega^{n-m}\}$ on U. Let $\{X_1, ..., X_m Y_1, ..., Y_{n-m}\}$ be the moving frame dual to the coframe. Then

$$\Sigma_q = span\{(X_1)_q, ..., (X_m)_q\}$$

for all $q \in U$.
For a given $i \in \{1, ..., n-m\}$:

$$d\omega^i = \sum_{a,b=1}^{m} \lambda^i_{ab} \tau^a \wedge \tau^b + \sum_{a=1}^{m} \sum_{b=1}^{n-m} \mu^i_{ab} \tau^a \wedge \omega^b + \sum_{a,b=1}^{n-m} \nu^i_{ab} \omega^a \wedge \omega^b$$
$$\text{for some } \lambda^i_{ab}, \mu^i_{ab}, \nu^i_{ab} \in C^\infty(U).$$

This implies, on the one hand, that

$$d\omega^i \wedge \omega^1 \wedge ... \wedge \omega^{n-m} = \sum_{a,b=1}^{m} \lambda^i_{ab} \tau^a \wedge \tau^b \wedge \omega^1 \wedge ... \wedge \omega^{n-m};$$

and, on the other hand, that $\forall j, k \in \{1, ..., m\}$:

$$d\omega^i(X_j, X_k) = \sum_{a,b=1}^{m} \lambda^i_{ab} \tau^a \wedge \tau^b (X_j, X_k) + \sum_{a=1}^{m} \sum_{b=1}^{n-m} \mu^i_{ab} \tau^a \wedge \omega^b (X_j, X_k)$$
$$+ \sum_{a,b=1}^{n-m} \nu^i_{ab} \omega^a \wedge \omega^b (X_j, X_k) = \sum_{a,b=1}^{m} \lambda^i_{ab} \delta^a_j \delta^b_k = \lambda^i_{jk}.$$

Therefore, by combining the two obtained results, the statement of the proposition is equivalent to the statement Σ *is integrable if and only if* $d\omega^i(X_j, X_k) = 0$ *for all* $i \in \{1, ..., n-m\}$ *and for all* $j, k \in \{1, ..., m\}$.
Another expression for $d\omega^i(X_j, X_k)$ is

$$d\omega^i(X_j, X_k) = X_j \omega^i(X_k) - X_k \omega^i(X_j) - \omega^i([X_j, X_k]).$$

However, for all $i \in \{1, ..., n-m\}$ and for all $j \in \{1, ..., m\} : \omega^i(X_j) = 0$, and so

$$d\omega^i(X_j, X_k) = -\omega^i([X_j, X_k]) \Rightarrow [d\omega^i(X_j, X_k) = 0 \Leftrightarrow \omega^i([X_j, X_k]) = 0].$$

Now observe that, because of the bilinearity of $[\cdot, \cdot]$ and the fact that $\{(X_1)_q, ..., (X_m)_q\}$ is a basis of Σ_q, $\omega^i([X_j, X_k]) = 0$ for all $i \in \{1, ..., n-m\}$ and for all $j, k \in \{1, ..., m\}$ if and only if Σ is involutive.
Finally, the proposition follows by the Frobenius Theorem 4.12. □

4.3 An Ice Skate

We can model an ice skate in a simplified way as a line segment that can move either in the same direction of the line segment, either forwards or backwards, or rotate about its middle point (in particular the ice skate cannot slide sideways). To describe the position of the skate, we can use the configuration space $\mathbb{R}^2 \times \mathbb{S}^1$ and take the first two coordinates (x, y) to be the middle point of the line segment that represents the ice skate, and the third coordinate (θ) to be the angle between the line segment representing the ice skate and the x-axis. From now on we will say *ice skate* to mean *the line segment representing the ice skate*. The restrictions we impose on the possible motion of the ice skate are equivalent to a non-holonomic constraint on the system. That the motion of the ice skate can only be along the same direction of the ice skate means that $(\dot{x}, \dot{y})$ must be proportional to (x, y), which is itself proportional to $(\cos\theta, \sin\theta)$. Therefore, using the characterisation of distributions provided by Proposition 4.2, we must require that the motion of the ice skate be compatible with the distribution Σ defined on $\mathbb{R}^2 \times \mathbb{S}^1$ by the kernel of the 1-form

$$\omega = -\sin\theta dx + cos\theta dy.$$

To determine whether we are dealing with an integrable distribution, we use Proposition 4.13. We have

$$\begin{aligned} d\omega &= -\cos\theta d\theta \wedge d\theta - \sin\theta d\theta \wedge dy \\ &\Rightarrow d\omega \wedge \omega = -\cos^2\theta d\theta \wedge dx \wedge dy + \sin^2\theta d\theta \wedge dy \wedge dx \\ &\qquad = -dx \wedge dy \wedge d\theta \neq 0, \end{aligned}$$

and so the distribution is not integrable.
Proposition 4.13 makes use in its proof of the Frobenius Theorem 4.12. We can also show that Σ is not integrable without using the Frobenius Theorem by making the following observation:
Assume that a distribution $\tilde{\Sigma}$, m-dimensional, were integrable. Then the set of all points $q \in M$ which are accessible from a given point $p \in M$ by a curve compatible with $\tilde{\Sigma}$ is the leaf L_p containing p of the corresponding m-dimensional foliation $\mathcal{F}$ on M.
In the case of our present example, the ice skate, we will show that the whole configuration space $\mathbb{R}^2 \times \mathbb{S}^1$ is accessible from any given point $p \in M$ by a curve compatible with Σ, thereby implying that no such foliation exists, meaning that Σ is non-integrable.

Claim 4.14 *Let $p, q \in \mathbb{R}^2 \times \mathbb{S}^1$. Then there exists a piecewise differentiable curve $c : [0,1] \to \mathbb{R}^2 \times \mathbb{S}^1$ compatible with Σ such that $c(0) = p$ and $c(1) = q$.*

Proof In order to get from $p =: (x_1, y_1, \theta_1)$ to $q =: (x_2, y_2, \theta_2)$, we need to break our motion into three steps: first we need to align the ice skate with the vector $(x_2 - x_1, y_2 - y_1)$ (a rotation), then we need to get to (x_2, y_2) and finally we need to rotate the ice skate again to get to q. Set

$$\tilde{\theta} := \begin{cases} \arctan \frac{y_2 - y_1}{x_2 - x_1} & \text{if } x_1 \neq x_2 \\ sgn(y_2 - y_1)\frac{\pi}{2} & \text{if } x_1 = x_2 \end{cases}. \tag{4.1}$$

Define $c : [0,1] \to \mathbb{R}^2 \times \mathbb{S}^1$ as

$$c(t) := \begin{cases} (x_1, y_1, \theta_1 + 3t(\tilde{\theta} - \theta_1)) & \text{for } t \in [0, \frac{1}{3}] \\ (x_1 + (3t-1)(x_2 - x_1), y_1 + (3t-1)(y_2 - y_1), \tilde{\theta}) & \text{for } t \in [\frac{1}{3}, \frac{2}{3}] \\ (x_2, y_2, \tilde{\theta} + (3t-2)(\theta_2 - \tilde{\theta})) & \text{for } t \in [\frac{2}{3}, 1] \end{cases}.$$

c is clearly piecewise differentiable and $c(0) = p, c(1) = q$. Further,

$$\dot{c}(t) := \begin{cases} (0, 0, 3(\tilde{\theta} - \theta_1)) & \text{for } t \in (0, \frac{1}{3}) \\ (3(x_2 - x_1), 3(y_2 - y_1), 0) & \text{for } t \in (\frac{1}{3}, \frac{2}{3}) \\ (0, 0, 3(\theta_2 - \tilde{\theta})) & \text{for } t \in (\frac{2}{3}, 1) \end{cases},$$

which is in the kernel of $\omega = -\sin\theta dx + cos\theta dy$ for all t in either of the three intervals. In particular, for $t \in (\frac{1}{3}, \frac{2}{3})$ this holds by the definition of $\tilde{\theta}$ (4.1), since the definition implies that $y_2 - y_1 = \lambda_t \sin\tilde{\theta}$ and $x_2 - x_1 = \lambda_t \cos\tilde{\theta}$ for some $\lambda_t \in \mathbb{R}$. Therefore, c is compatible with Σ and the claim follows. □

We conclude by characterising the possible motions of the ice skate.

Claim 4.15 *Assume that the kinetic energy of the skate is*

$$K = \frac{M}{2}\left((v^x)^2 + (v^y)^2\right) + \frac{I}{2}(v^\theta)^2,$$

and that the reaction force is perfect. Then the motion of the ice skate is either a straight line or a circle, and it has constant speed.

We will need the following lemma.

Lemma 4.16 *Let V be an n-dimensional vector field over $\mathbb{R}$, let $f, g \in V^*$. If $ker(f) \subseteq ker(g)$, then $g = \lambda f$ for some $\lambda \in \mathbb{R}$.*

Proof If $g = 0$ the statement is trivial. Assume, therefore, that $g \neq 0$. If $f = 0$, then $V = ker(f) = ker(g) \Rightarrow g = 0$.
If $f \neq 0$, then $dim(im(f)) = 1$ (since $im(f) = \mathbb{R}$) $\Rightarrow dim(ker(f)) = n - 1$. Let $\{v_1, ..., v_{n-1}\}$ be a basis of $ker(f)$ (if $n = 1$, we take the empty set). For $v \notin ker(f) : \{v_1, ..., v_{n-1}\} \cup \{v\}$ is a basis of V. By setting $\lambda := \frac{g(v)}{f(v)}$ we have $g(u) = \lambda f(u)$ for all $u \in \{v_1, ..., v_{n-1}, v\}$ (i.e. for all elements of a basis of V) and so the lemma follows. □

Proof (Claim) By Proposition 1.6

$$\mu\left(\frac{D\dot{c}}{dt}\right) = M(\ddot{x}dx + \ddot{y}dy) + I\ddot{\theta}d\theta. \tag{4.2}$$

Observe that, since by assumption the reaction force $\mathcal{R}$ is perfect, for all $t \in (a,b) \subseteq \mathbb{R}$:

$$\mu^{-1}(\mathcal{R}(\dot{c}(t))) \in \Sigma^{\perp}_{c(t)} \Rightarrow \forall v \in \Sigma_{c(t)} : \mathcal{R}(\dot{c}(t))(v) = \langle \mu^{-1}(\mathcal{R}(\dot{c}(t))), v\rangle = 0$$
$$\Rightarrow \Sigma_{c(t)} \subseteq ker(\mathcal{R}(\dot{c}(t))).$$

Recalling that $\Sigma_{c(t)} = ker(\omega_{c(t)})$ and by using Lemma 4.16, we can therefore write

$$\mathcal{R}(\dot{c}(t)) = \alpha(t)\omega_{c(t)} \tag{4.3}$$

for all $t \in (a,b)$, where $\alpha : (a,b) \subseteq \mathbb{R} \to \mathbb{R}$. The motion of the ice skate must satisfy the generalised Newton equation in Definition 4.6. Since we have no external force besides the reaction force, the right hand side is given by equation (4.3) (recall the definition of ω in (4.3)). The left hand side is given by equation (4.2). Together, we obtain that the motion must satisfy (not-explicitly stating the dependency of t)

$$M\ddot{x}dx + M\ddot{y}dy + I\ddot{\theta}d\theta = -\alpha\sin\theta dx + \alpha\cos\theta dy.$$

The motion must also be compatible with the distribution Σ, i.e. (again leaving the dependency on t implicit, and writing Σ for $\Sigma_{c(t)}$ and $ker(\omega)$ for $ker(\omega)_{c(t)}$)

$$(\dot{x}, \dot{y}, \dot{\theta}) \in \Sigma = ker(\omega) \Leftrightarrow \dot{x}\sin\theta = \dot{y}\cos\theta.$$

We then obtain a system of ODEs that the motion of the ice skate must satisfy:

$$\begin{cases} M\ddot{x} = -\alpha\sin\theta \\ M\ddot{y} = \alpha\cos\theta \\ \ddot{\theta} = 0 \\ \dot{x}\sin\theta = \dot{y}\cos\theta \end{cases}. \tag{4.4}$$

Firstly, the third ODE of the system (4.4) implies that

$$\theta(t) = \theta_0 + v_0^{\theta}t \tag{4.5}$$

for some integration constants $\theta_0, v_0^{\theta} \in \mathbb{R}$. We make the following case distinction:

1. $v_0^\theta \neq 0$:
Differentiating the fourth ODE, taking into account that θ is given by equation (4.5), and substituting the values of $\ddot{x}$, $\ddot{y}$ and $\dot{y}$ into the ensuing equation according to the first, second and fourth ODEs respectively, we get

$$\frac{d}{dt}(\dot{x}\sin\theta) = \frac{d}{dt}(\dot{y}\cos\theta)$$
$$\Rightarrow \ddot{x}\sin\theta + v_0^\theta \dot{x}\cos\theta = \ddot{y}\cos\theta - v_0^\theta \dot{y}\sin\theta$$
$$\Leftrightarrow -\frac{\alpha}{M}\sin^2\theta + v_0^\theta \dot{x}\cos\theta = \frac{\alpha}{M}\cos^2\theta - v_0^\theta \dot{x}\tan\theta\sin\theta$$
$$\Leftrightarrow \frac{1}{\cos\theta}v_0^\theta \dot{x} = \frac{\alpha}{M} \Leftrightarrow \dot{x} = \frac{\alpha}{M v_0^\theta}\cos\theta. \quad (4.6)$$

Differentiating $\dot{x}$ and comparing the resulting equation to the first ODE in the system (4.4), we see that $\alpha(t) = \alpha = const. \in \mathbb{R}$. Finally, we get

$$x = x(t) = \frac{\alpha}{M(v_0^\theta)^2}\sin\theta + x_0 \quad (4.7)$$

for some integration constant $x_0 \in \mathbb{R}$. By substituting equation (4.6) into the fourth ODE of the system (4.4), we get

$$\dot{y} = \frac{\alpha}{M v_0^\theta}\sin\theta$$

and, consequently,

$$y = y(t) = -\frac{\alpha}{M(v_0^\theta)^2}\cos\theta + y_0 \quad (4.8)$$

for some integration constant $y_0 \in \mathbb{R}$.
Observe that, by equations (4.7) and (4.8)

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = \frac{|\alpha|}{M(v_0^\theta)^2},$$

implying that the motion of the ice skate is in a circle of radius $\frac{|\alpha|}{M(v_0^\theta)^2}$ centered at $(x_0, y_0) \in \mathbb{R}^2$. Further, the speed of the motion is

$$\sqrt{(\dot{x})^2 + (\dot{y})^2} = \frac{|\alpha|}{M v_0^\theta} = const. \in \mathbb{R}.$$

Remark 4.17 *In this case, the reaction force can be understood to be a friction force preventing the ice skate from sliding sideways, so that it stays in a circular trajectory.*

2. $v_0^\theta = 0$:
   In this case, equation (4.5) implies that $\theta = \theta_0 = const. \in \mathbb{R}$. By differentiating the fourth ODE in the system (4.4) and substituting $\ddot{x}$ and $\ddot{y}$ into the obtained equation according to the first and second ODEs in the system (4.4), we get

   $$\ddot{x}\sin\theta = \ddot{y}\cos\theta \Rightarrow -\frac{\alpha}{M}\sin^2\theta = \frac{\alpha}{M}\cos^2\theta \Leftrightarrow \alpha = 0.$$

   This readily implies, by the first and second ODEs in the system (4.4), that

   $$\ddot{x} = x_0 + v_0^x t \tag{4.9}$$

   $$\ddot{y} = y_0 + v_0^y t \tag{4.10}$$

   for some integration constants $x_0, y_0, v_0^x, v_0^y \in \mathbb{R}$. The motion of the ice skate is, therefore, along a straight line though $(x_0, y_0) \in \mathbb{R}^2$, with constant speed $\sqrt{(v_0^x)^2 + (v_0^y)^2}$.

Remark 4.18 *In this case, by the generalised Newton equation in Definition 4.6, we have that* $\frac{D\dot{c}}{dt} = 0 \Rightarrow \mathcal{R}(\dot{c}) = \mu\left(\frac{D\dot{c}}{dt}\right) = 0$*, and so the reaction force vanishes.*

. □

Bibliography

[1] Leonor Godinho and José Natário. *An Introduction to Riemannian Geometry*. Springer, 2014.

[2] Urs Lang. *Lecture Notes on Riemannian Geometry*. Based on the Courses Differential Geometry II Taught in the Spring Semesters of 2017 and 2020 at ETH Zürich.

[3] Shigeyuki Morita. *Geometry of Differential Forms*. American Mathematical Society, 2001.