Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques i Informàtica
Universitat de Barcelona

Trajectory interpolation for inertial measurement simulation

Autor: Àlvar Pineda Romero

Director: Dr. Àlex Haro Provinciale
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 29 de juny de 2017

Abstract

This work addresses the problem of trajectory interpolation for inertial measurement simulation. We have discarded the traditional interpolation procedures in order to implement more physically accurate techniques. Thus, we have adapted the work of the robotic community to our purposes, designing a quite innovative proceeding. Indeed, we have tested the proposed algorithm with a real 4 hours trajectory, obtaining satisfactory results.

Acknowledgements

I would like to express gratitude to my advisor Dr. Àlex Haro for his guidance and patience and also to Dr. Ismael Colomina, Chief Scientist of GeoNumerics, for having proposed this work and for his support. I would also like to thank my family and friends for their unconditional support, specially to my beloved mother.

Contents

Chapter 1

Introduction

1.1 Structure of the work

This work has been organized in three Chapters. This Chapter motivates and introduces the problem that we want to study: the trajectory interpolation. Chapter 2 provides the theoretical frame needed to derive the interpolation algorithms: we are going to mainly study and describe the Lie theory and also the Riemannian geometry theory. Finally, Chapter 3 focuses on the obtainment of different algorithms to interpolate a rigid body trajectory.
In addition, we also provide an Appendix with graphics comparing and analysing the studied algorithms.

1.2 Motivation

This work stems from a project of GeoNumerics. S.L., which I have been working on. GeoNumerics is a small company that focuses on the needs and technology of high precision kinematic surveying; sensor orientation and calibration and trajectory determination. The project aims to develop a simulator of an Inertial Measurement Unit (IMU). An IMU is an electronic device consisting of a set of accelerometers and a set of angular rate sensors that measure the linear acceleration and the angular velocity of a body. Thus, using the measures of an IMU, it is possible to obtain its position and velocity.
Indeed, this device provides a necessary complement to a Global Navigation Satellite System (GNSS) and therefore it is an essential tool for navigation and orientation with applications in many areas such as aviation (for manned planes and Unmanned Aerial Vehicles - UAV's -), terrestrial mobility (again, for manned and unmanned vehicles), cartography, indoor navigation, trajectories of satellites and space vehicles, robotics, navigation of submarines, etcetera.

The reason for developing a simulator is that one of the most important steps in the planning of a technological project is the choice of the sensors that should be used in that project. The type of the sensors that have to be used, the total number of sensors and also their quality are aspects that can have a great impact in the obtained results and also in the project costs. This selection process is not usually performed using real sensors because this would entail an increase in the costs. Alternatively, the realistic simulation of the sensor appears to be the best way to proceed.

In order to simulate the IMU measures corresponding to a trajectory of a rigid body, we need to know this trajectory. Thus, the steps to simulate these measures begin with a sparse trajectory provided by a client, obtained from an orientation application such as Google Maps, or even simulated. It is worth noting that by *trajectory* we are referring to a set of positions and orientations labelled by the time. These trajectories can also be referred as tPA trajectories. Subsequently, we have to interpolate this trajectory to obtain a frequency of thousands of Hz. This high frequency allow us to simulate vibrations and shocks. Finally, we compute the IMU measures with a lower frequency (hundreds of Hz).

Thus, this work focuses on the interpolation of the trajectory. Before explaining the techniques and procedures that we are going to follow we should recall some basic concepts about navigation.

1.3 Basic concepts of inertial navigation

First of all we define a set of navigation and orientation concepts. In this regard, a coordinate system can be defined [1] as a set of conventions and physical theories (or models) required to define at any time a triad of axes. In addition, a reference frame is a realisation of a coordinate system, i.e. a set of points determined by their position coordinates that locate and orient the coordinate system.
The usual reference frames are the following [1].

- i (*Operational Inertial Frame*): its origin is defined to be the center of the Earth; the z axis is parallel to the Earth rotation axis, the x axis points to the equinoctial colure and the y axis is orthogonal to z and x with the usual right-hand rule. Taking into account the precision of most of IMU's we can consider that this frame neither rotate nor is accelerated with respect a real inertial reference frame.

- e (ECEF, *Earth-centred Earth-fixed*): the Earth's center of mass is the origin of this reference frame and the z axis points to the North. The mean equatorial plane perpendicular to this axis forms the xy-plane and the xz-plane is generated by the mean meridian plane of Greenwich. Finally, the y axis is chosen following the right-hand rule.

- l (*Local-level*): this reference frame is defined with respect a certain object and therefore its origin is this object. There are different ways to define the axis orientation. Consider the Lned (Local North-East-Down): one axis points to the North, another one points to the East and the down direction is chosen following the right-hand rule.

- b (*Body*): this reference frame is also defined with respect a certain object and its origin is also this object. As with the local frame, there are different ways to define the axis as for instance, Bfrd (Body forward-right-down): one axis points forward, another one points to the right and the other points to the down direction, chosen by the right-hand rule. Notice that these directions can be arbitrary chosen.

Furthermore, the Earth's surface can be closely approximated by a rotation ellipsoid with flattened poles. This ellipsoid is described by the semimajor axis a and the semimenor axis b. Then, let P be a point with ECEF coordinates $P = (X, Y, Z)$, we can define its geodetic coordinates as follows. The latitude ϕ is the angle between the equatorial plane (x, y-plane) of the ellipsoid and the normal at the ellipsoid that passes through P; the longitude λ is the angle between the zero meridian (x axis) and the meridian plane of P; finally the position is parametrized using the height H of P above the ellipsoid. In addition, the following equations define the transformation from the geodetic coordinates (λ, ϕ, H) to the ECEF coordinates (X, Y, Z) [1]:

$$\begin{aligned} X &= (N(\phi) + h)\cos\phi\cos\lambda, \\ Y &= (N(\phi) + h)\cos\phi\sin\lambda, \\ Z &= ((1 - e^2)N(\phi) + h)\sin\phi, \end{aligned} \tag{1.1}$$

where

$$N(\phi) = \frac{a^2}{\sqrt{a^2\cos^2\phi + b^2\sin^2\phi}}, \qquad e = \frac{\sqrt{a^2 - b^2}}{a}.$$

The inverse transformation has to be computed numerically.

Moreover, the orientation of an object is described by the rotation between the local and the body frame and is also refered as *attitude*. This rotation can be parametrized in different ways: using Euler angles, quaternions, defining the principal axis of rotation, etcetera. The usual parametrization corresponds to the Tait-Bryan $z - y - x$ sequence of rotation and parametrizes the rotation in terms of the Euler angles heading (ψ), pitch (θ) and roll (γ). Thus, the rotation from Bfrd to Lned can be expressed as [1]

$$R_b^l = R_{b_2}^l(\psi) R_{b_1}^{b_2}(\theta) R_b^{b_1}(\gamma) = R_z(\psi) R_y(\theta) R_x(\gamma),$$

$$\begin{aligned} R_b^l &= \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\psi & \sin\gamma\sin\theta\cos\psi - \cos\gamma\sin\psi & \cos\gamma\sin\theta\cos\psi + \sin\gamma\sin\psi \\ \cos\theta\sin\psi & \sin\gamma\sin\theta\sin\psi + \cos\gamma\cos\psi & \cos\gamma\sin\theta\sin\psi - \sin\gamma\cos\psi \\ -\sin\theta & \sin\gamma\cos\theta & \cos\gamma\cos\theta \end{pmatrix}. \end{aligned} \tag{1.2}$$

Thus, the Tait-Bryan rotation $z-y-x$ can be decomposed as a heading rotation with respect the z axis, a pitch rotation with respect y and a roll rotation with respect x.
We can also express this rotation using the unit quaternions $q = (q_0, q_1, q_2, q_3)$ [1]:

$$R_l^b = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_0^2 & 2(q_1q_2 + q_3q_0) & 2(q_1q_3 - q_2q_0) \\ 2(q_1q_2 - q_3q_0) & -q_1^2 + q_2^2 - q_3^2 + q_0^2 & 2(q_2q_3 + q_1q_0) \\ 2(q_1q_3 + q_2q_0) & 2(q_2q_3 - q_1q_0) & -q_1^2 - q_2^2 + q_3^2 + q_0^2 \end{pmatrix}. \tag{1.3}$$

Recall that unit quaternions satisfy

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1.$$

Sometimes, the attitude may also refer to the rotation R_b^e. Indeed, this matrix can be obtained from R_b^l using:

$$R_b^e = R_l^e R_b^l, \tag{1.4}$$

where R_l^e is a rotation that depends on the position. Let (λ, ϕ, H) be the position expressed using geodetic coordinates, then [1]:

$$R_l^e = \begin{pmatrix} -\sin\phi\cos\lambda & -\sin\lambda & -\cos\phi\cos\lambda \\ -\sin\phi\sin\lambda & \cos\lambda & -\cos\phi\sin\lambda \\ \cos\phi & 0 & -\sin\phi \end{pmatrix}. \tag{1.5}$$

1.4 Interpolation considerations

To interpolate a tPA trajectory we have to study the interpolation of the position and the interpolation of the attitude.
The interpolation of the position in $\mathbb{R}^3$ has been widely discussed in the literature. One of the most common techniques are the well known cubic splines, but Bézier curves or even the simple linear interpolation can also be implemented. On the other hand, the interpolation of the attitude has not been so extensively studied.
In the previous section we have described two usual parametrizations of the rotation. Indeed, in the context of navigation, the attitude interpolation has been usually analysed as an interpolation in $\mathbb{R}^n$ and therefore the usual procedure is to interpolate the heading-pitch-roll angles ($n = 3$) or the quaternions ($n = 4$) using the above described methods.
However, these interpolations may produce an attitude trajectory with no physical sense. An additional reason to avoid these parametrizations are their singularities. From Eq. (1.2) we observe that for $\theta = \pm\frac{\pi}{2}$, we can not obtain the other angles and from Eq. (1.3) we get that q and $-q$ define the same rotation and this may produce some wrong interpretations for small rotations. In addition, the additive operation for quaternions is not defined as usual and this leads to some complexities to adapt the algorithms.
Nevertheless, as we are going to explain in Chapter 3, a tPA trajectory can be though as a set of points in $SE(3)$ and therefore we can study the interpolation of both position and attitude in $SE(3)$. In this regard, in recent decades, professionals of robotics and computer graphics have developed physically accurate interpolations based on the Lie group interpretation of $SE(3)$. Thus, we want to adapt these algorithms to a trajectory interpolation. This proceeding is quite innovative in the context of navigation.

Chapter 2

Theoretical framework

This chapter describe the theoretical framework needed to derive the interpolation algorithms described in Chapter 3. At the beginning of each section we are going to state the main bibliographic sources that have been used. Other references can be punctually cited to indicate that they have been used for a specific definition or statement. The theoretical content of these references has been complemented with our own results.

2.1 Differentiable manifolds and vector fields

In this section we are going to introduce the differentiable manifolds, the tangent space, the vector fields and other related concepts needed to define the Lie groups and the associated Lie algebras. The main reference for this section has been [2].

We begin with the definition of a differentiable manifold.

Definition 2.1. ***Differentiable manifold.*** *A differentiable manifold of dimension n is a set M and a family of injective mappings $\widetilde{x}_\alpha : U_\alpha \subset \mathbb{R}^n \to M$ of open sets U_α of $\mathbb{R}^n$ into M such that:*

i) $\bigcup_\alpha \widetilde{x}_\alpha(U_\alpha) = M$.

ii) For any pair α, β, with $\widetilde{x}_\alpha(U_\alpha) \cap \widetilde{x}_\beta(U_\beta) = W \neq \emptyset$, the sets $\widetilde{x}_\alpha^{-1}(W)$ and $\widetilde{x}_\beta^{-1}(W)$ are open sets in $\mathbb{R}^n$ and the mappings $\widetilde{x}_\beta^{-1} \circ \widetilde{x}_\alpha$ are differentiable.

The pair $(U_\alpha, \widetilde{x})$ with $p \in \widetilde{x}_\alpha(U_\alpha)$ is called a parametrization (or system of coordinates) of M at p. $\widetilde{x}_\alpha(U_\alpha)$ is called a coordinate neighbourhood at p.
A family $\{U_\alpha, \widetilde{x}_\alpha\}$ satisfying the conditions of this definition is called a differentiable structure on M. Moreover, it is usual to include a third condition in the above definition:

iii) The family $\{U_\alpha, \widetilde{x}_\alpha\}$ is maximal relative to the conditions (i) and (ii).

Nevertheless, given a differentiable structure on M we can take the union of all the parametrizations that, together with any of the parametrization of the provided differentiable structure, satisfy condition (ii). Thus, we can complete the given differentiable structure to a maximal one and therefore condition (iii) is a consequence of (i) and (ii).

Remark 2.2. A differentiable structure on a set M induces a natural topology on M. We can define $A \subset M$ to be an open set in M if and only if $\widetilde{x}_\alpha^{-1}(A \cap \widetilde{x}_\alpha(U_\alpha))$ is an open set in $\mathbb{R}^n$, for all α. Thus it is easy to show that M and the empty set are open sets, that a union of open sets is an open set and finally that the finite intersection of open sets is again an open set. Observe that $\widetilde{x}_\alpha(U_\alpha)$ are open sets and therefore the mappings $\widetilde{x}_\alpha$ are continuous.

Now we are going to introduce the differential mappings between differentiable manifolds.

Definition 2.3. *Let N and M be two differentiable manifolds of dimension n and m, respectively. A mapping $f : N \to M$ is differentiable at $p \in N$ if given a parametrization $\widetilde{y} : V \subset \mathbb{R}^m \to M$ at $f(p)$ there exists a parametrization $\widetilde{x} : U \subset \mathbb{R}^n \to N$ at p such that $f(\widetilde{x}(U)) \subset \widetilde{y}(V)$ and the mapping*

$$\widetilde{y}^{-1} \circ f \circ \widetilde{x} : U \subset \mathbb{R}^n \to \mathbb{R}^m, \tag{2.1}$$

is differentiable at $\widetilde{x}^{-1}(p)$. This mapping is called the expression of f in the parametrizations $\widetilde{x}$ and $\widetilde{y}$. Moreover, f is differentiable on an open set of N if it is differentiable at all the points of this open set.

From condition (ii) of Definition 2.1, the definition of a differentiable mapping between manifolds is independent of the choice of the parametrizations.

Moreover, the tangent space at a point of a manifold has many different definitions. We are going to use the differentiable curves to define it.

Definition 2.4. ***Curve on a manifold.*** *Let M be a differentiable manifold. Let $I = (-\varepsilon, \varepsilon) \subset \mathbb{R}$ be an open set. A differentiable function*

$$\alpha : I \to M,$$

is called a (differentiable) curve in M. In addition, α is said to be a curve through a point $p \in M$ if $\alpha(0) = p$.

Definition 2.5. ***Tangent vector to a curve.*** *Let M be a differentiable manifold and $\alpha : I \to M$ a curve through p. Let $\mathcal{D}$ be the set of functions on M that are differentiable at p. The tangent vector to the curve α at $t = 0$ is a function $\alpha'(0) : \mathcal{D} \to \mathbb{R}$ given by*

$$\alpha'(0)f = \left.\frac{d(f \circ \alpha)}{dt}\right|_{t=0},$$

for all $f \in \mathcal{D}$.

Thus, a tangent vector at p is the tangent vector at $t = 0$ of some curve $\alpha : I \to M$ with $\alpha(0) = p$. The set of all tangent vectors to M at p can be denoted as T_pM.

Now let M be an n dimensional differentiable manifold and choose a parametrization $\widetilde{x} : U \to M$ at $p = \widetilde{x}(0) \in M$. We can express every function $f \in \mathcal{D}$ and every curve α in this parametrization by

$$f \circ \widetilde{x}(q) = f(x_1, \ldots, x_n), \qquad q = (x_1, \ldots, x_n) \in U;$$
$$\widetilde{x}^{-1} \circ \alpha(t) = (x_1(t), \ldots, x_n(t)).$$

Therefore we get

$$\begin{aligned}\alpha'(0)f &= \left.\frac{d}{dt}(f \circ \alpha)\right|_{t=0} = \left.\frac{d}{dt} f(x_1(t), \ldots, x_n(t))\right|_{t=0} \\ &= \sum_{i=1}^{n} x_i'(0)\left(\frac{\partial f}{\partial x_i}\right)_0 = \left(\sum_{i=1}^{n} x_i'(0)\left(\frac{\partial}{\partial x_i}\right)_0\right) f. \end{aligned} \tag{2.2}$$

Since this is true for all $f \in \mathcal{D}$:

$$\alpha'(0) = \sum_{i=1}^{n} x_i'(0)\left(\frac{\partial}{\partial x_i}\right)_0. \tag{2.3}$$

Thus, a tangent vector to the curve α at p depends only on the derivative of α in a coordinate system.
In addition, from Eq. (2.3) it follows that the set T_pM forms a vector space of dimension n and that the choice of a parametrization $\widetilde{x} : U \to M$ determines an associated basis in T_pM:

$$\left\{ \left(\frac{\partial}{\partial x_1}\right)_0, \ldots, \left(\frac{\partial}{\partial x_n}\right)_0 \right\}.$$

Observe that $\left(\frac{\partial}{\partial x_i}\right)_0$ is the tangent vector at p of the curve $x_i \mapsto \widetilde{x}(0,\dots,0,x_i,0,\dots,0)$. Furthermore, it is clear that the linear structure in T_pM does not depend on the chosen parametrization. The vector space T_pM is called the tangent space of M at p.

Once we have defined the tangent space we can introduce the differential of a function between two manifolds.

Proposition 2.6. *Let N and M be differentiable manifolds of dimension n and m respectively. Let $f : N \to M$ be a differentiable mapping. For every $p \in N$ and for each $v \in T_pN$, we choose a differentiable curve $\alpha : (-\varepsilon, \varepsilon) \to N$ with $\alpha(0) = p$ and $\alpha'(0) = v$. We define another differentiable curve β by $\beta = f \circ \alpha$. Then, the mapping $d_pf : T_pN \to T_{f(p)}M$ given by $d_pf(v) = \beta'(0)$ is a linear mapping that does not depend on the choice of α. d_pf is called the differential of f at p.*

Proof. We choose the parametrization $\widetilde{x} : U \to N$ and $\widetilde{y} : V \to M$ at p and $f(p)$, respectively. Now we express f in these parametrization:

$$\widetilde{y}^{-1} \circ f \circ \widetilde{x}(q) = (y_1(x_1,\dots,x_n),\dots,y_m(x_1,\dots,x_n)),$$

where

$$q = (x_1,\dots,x_n) \in U, \qquad (y_1,\dots,y_m) \in V.$$

On the other hand, we can express α in the parametrization $\widetilde{x}$:

$$\widetilde{x}^{-1} \circ \alpha(t) = (x_1(t),\dots,x_n(t)).$$

Thus, from Eq. (2.3) we get that the expression of $\beta'(0)$ with respect to the basis $\left\{\left(\frac{\partial}{\partial y_i}\right)_0\right\}_{i=1,\dots,m}$ of $T_{f(p)}M$, associated to the parametrization $\widetilde{y}$ is given by:

$$\beta'(0) = \left(\sum_{i=1}^{n} \frac{\partial y_1}{\partial x_i} x_i'(0), \dots, \sum_{i=1}^{n} \frac{\partial y_m}{\partial x_i} x_i'(0)\right).$$

Since $x_i'(0)$ is determined by v, $\beta'(0)$ does not depend on the choice of α.
Now let $\left(\frac{\partial y}{\partial x}\right)$ be the $m \times n$ matrix with elements $\left(\frac{\partial y}{\partial j}\right)_{ij} = \frac{\partial y_i}{\partial x_j}$ $(i = 1,\dots,m,\ j = 1,\dots,n)$ and let $x'(0)$ be the vector with elements $x'(0)_j = x_j'(0)$ $(j = 1,\dots,n)$. Then we get:

$$d_pf(v) = \beta'(0) = \left(\frac{\partial y}{\partial x}\right) x'(0). \tag{2.4}$$

Hence, d_pf is a linear mapping between T_pN and $T_{f(p)}M$ with matrix $\left(\frac{\partial y}{\partial x}\right)$ in the associated bases obtained from the chosen parametrizations. □

Let M be an n dimensional manifold. We define the set TM by:

$$TM = \{(p,v); p \in M, v \in T_pM\}. \tag{2.5}$$

TM is called the tangent bundle. The following proposition shows that we can provide TM with a $2n$ dimensional differential structure.

Proposition 2.7. *Let M be an n dimensional manifold. TM can be provided with a differential structure of dimension $2n$.*

Proof. Let $\{U_\alpha, \widetilde{x}_\alpha\}$ be a maximal differentiable structure on M. We denote by $(x_1^\alpha,\dots,x_n^\alpha)$ the coordinates of U_α and by $\{\frac{\partial}{\partial x_1^\alpha},\dots,\frac{\partial}{\partial x_n^\alpha}\}$ the associated bases to the tangent spaces of $\widetilde{x}_\alpha(U_\alpha)$. Then for each α we define $\widetilde{y}_\alpha : U_\alpha \times \mathbb{R}^n \to TM$ by

$$\widetilde{y}_\alpha(x_1^\alpha,\dots,x_n^\alpha,u_1,\dots,u_n) = \left(\widetilde{x}_\alpha(x_1^\alpha,\dots,x_n^\alpha), \sum_{i=1}^{n} u_i \frac{\partial}{\partial x_i^\alpha}\right).$$

We are going to prove that $\{U_\alpha \times \mathbb{R}^n, \widetilde{y}_\alpha\}$ is a differentiable structure on TM. First of all, we have $\cup_\alpha \widetilde{x}_\alpha(U_\alpha) = M$ and $(d\widetilde{x}_\alpha)_q(\mathbb{R}^n) = T_{\widetilde{x}_\alpha(q)}M$ for all $q \in U_\alpha$. Thus we get the condition (i) of Definition 2.1:

$$\bigcup_\alpha \widetilde{y}_\alpha(U_\alpha \times \mathbb{R}^n) = TM.$$

We have to show that condition (ii) is also verified. Let

$$(p, v) \in \widetilde{y}_\alpha(U_\alpha \times \mathbb{R}^n) \cap \widetilde{y}_\beta(U_\beta \times \mathbb{R}^n).$$

Then:

$$\begin{cases} (p, v) = (\widetilde{x}_\alpha(q_\alpha), d\widetilde{x}_\alpha(v_\alpha)) \\ (p, v) = (\widetilde{x}_\beta(q_\beta), d\widetilde{x}_\beta(v_\beta)), \end{cases}$$

where $q_\alpha \in U_\alpha$, $q_\beta \in U_\beta$, $v_\alpha, v_\beta \in \mathbb{R}^n$. Thus we have:

$$\widetilde{y}_\beta^{-1} \circ \widetilde{y}_\alpha(q_\alpha, v_\alpha) = \widetilde{y}_\beta^{-1}(\widetilde{x}_\alpha(q_\alpha), d\widetilde{x}_\alpha(v_\alpha)) = ((\widetilde{x}_\beta^{-1} \circ \widetilde{x}_\alpha)(q_\alpha), d(\widetilde{x}_\beta^{-1} \circ \widetilde{x}_\alpha)(v_\alpha)).$$

Since $\widetilde{x}_\beta^{-1} \circ \widetilde{x}_\alpha$ and $d(\widetilde{x}_\beta^{-1} \circ \widetilde{x}_\alpha)$ are differentiable, we get that $\widetilde{y}_\beta^{-1} \circ \widetilde{y}_\alpha$ is differentiable, showing the desired condition. □

Remember now the definition of a vector field.

Definition 2.8. ***Vector field.*** *A vector field on a differentiable manifold M is a mapping $X : M \to TM$ such that $\pi \circ X = id_M$, where π is the projection $\pi : TM \to M$ defined by $\pi(p, v) = p$. The vector field is differentiable if the mapping $X : M \to TM$ is differentiable.*
We are going to denote the space of all vector fields on M by $Vect(M)$.

Thus, the vector field assigns a tangent vector to each point $p \in M$. We are going to denote this vector by $X_p = X(p)$.
Moreover, a vector field can be thought as a mapping $X : \mathcal{D}(M) \to \mathcal{F}(M)$ (where $\mathcal{D}(M)$ are the set of differentiable functions on M and $\mathcal{F}(M)$ is the set of functions on M) defined by $X(f)(p) = (f \circ \gamma)'(0)$, where γ is a curve through $p \in M$ such that $\gamma'(0) = X_p$. Thus, an analogue definition is $X(f)(p) = (d_p f)(X_p)$.
Now we choose a parametrization $\widetilde{x} : U \subset \mathbb{R}^n \to M$. Then we can express $X(p)$ as

$$X(p) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x_i},$$

where $a_i : U \to \mathbb{R}$, and $\{\frac{\partial}{\partial x_i}\}_{i=1,\dots,n}$ is the basis associated to $\widetilde{x}$. A vector field X is differentiable if and only if the functions a_i are also differentiable for any parametrization (indeed, if they are differentiable for some parametrization they will be differentiable for any parametrization). Therefore:

$$X(f)(p) = \sum_{i=1}^n a_i(p) \frac{\partial f}{\partial x_i}(p), \tag{2.6}$$

where by abuse of notation, f denotes the expression of f in the chosen parametrization $\widetilde{x}$. Thus, X is differentiable if and only if $X(f) \in \mathcal{D}(M)$ for all $f \in \mathcal{D}(M)$ and therefore $X : \mathcal{D}(M) \to \mathcal{D}(M)$.

We are going to finish this section with the definition of the integral curves and the vector fields along a curve.

Definition 2.9. ***Integral curves.*** *Let X be a vector field on a manifold M. An integral curve (with respect to X) through $m \in M$ is a curve $\gamma : J \subset \mathbb{R} \to M$ such that $\gamma(0) = m$ and γ is a solution to the following differential equation*

$$\frac{d\gamma}{dt}(t) = X_{\gamma(t)},$$

for all $t \in J$.

Therefore, the fundamental theorem of ordinary differential equations implies that for any vector field X on M there exists $\varepsilon > 0$ and an integral curve $\gamma : (-\varepsilon, \varepsilon) \to M$ through $m \in M$. In addition, γ is locally unique.

Similarly, we can introduce the vector fields along a curve.

Definition 2.10. ***Vector field along a curve.*** *A vector field V along a curve $\gamma : I \to M$ is a differentiable mapping that associates to every $t \in I$ a tangent vector $V(t) \in T_{\gamma(t)}M$. In addition, V is differentiable if for any differentiable function f on M, the function $t \mapsto V(t)f$ is a differentiable function on I.*

2.2 Lie groups and Lie algebras

After all the definitions provided in the previous section, it is time to introduce the Lie groups and the Lie algebras. In order to do this we are going to define also the adjoint representations, that are going to be useful in Chapter 3. The main reference for this section has been [3].

Definition 2.11. ***Lie group.*** *A group G is said to be a Lie group if G is a manifold such that the multiplication map and the inversion map are both smooth maps.*

The group multiplication,

$$\begin{aligned} \cdot_G : G \times G &\to G \\ (g, h) &\mapsto g \cdot_G h, \end{aligned}$$

will be simply noted as gh instead of $g \cdot_G h$ $(g, h \in G)$. On the other hand, the inversion map can be noted as:

$$\begin{aligned} ^{-1} : G &\to G \\ g &\mapsto g^{-1}, \end{aligned}$$

where $gg^{-1} = g^{-1}g = e$, and e is the identity of the group G.

Example 2.12. The group of $n \times n$ non-singular and real valued matrices, $GL(n, \mathbb{R})$, is an example of a Lie group. First we have that $GL(n, \mathbb{R})$ is subset of $\mathbb{R}^{n^2}$. In addition, from the continuity of the determinant we get that $\det^{-1}(0)$ is a closed set of $\mathbb{R}^{n^2}$ and therefore $GL(n, \mathbb{R})$ is an open set of $\mathbb{R}^{n^2}$. Thus $GL(n, \mathbb{R})$ is a differentiable manifold.
In addition, the matrix multiplication is clearly smooth and so is the inversion map (by the Cramer's rule, the inverse A^{-1} of a matrix $A \in GL(n, \mathbb{R})$ is a smooth function of the elements of A).

Some subgroups H of $GL(n, \mathbb{R})$ are also Lie groups. In order to stablish when H is a Lie group we first provide the following definitions [4].

Definition 2.13. ***Matrix convergence.*** *Let A_m be a sequence of $n \times n$ real matrices. We say that A_m converges to a matrix A if each entry of A_m converges to the corresponding entry of A.*

Definition 2.14. ***Matrix Lie group.*** *A matrix Lie group is any subgroup H of $GL(n, \mathbb{R})$ with the following property: if A_m is any sequence of matrices in H and A_m converges to some matrix A, then either $A \in H$, or A is not invertible.*

Indeed, Definition 2.14 is equivalent to saying that a matrix Lie group is a closed subgroup of $GL(n, \mathbb{R})$.
The following Theorem shows that any closed subgroup of a Lie group is also a Lie group and therefore it implies that any matrix Lie group is a Lie group. We are not going to prove this Theorem due to its extension (see [4] for a detailed prove).

Theorem 2.15. ***Closed subgroup Theorem.*** *Suppose G is a Lie group and $H \subset G$ is a subgroup that is also a closed subset. Then H is an embedded Lie subgroup.*

Subsequently we will introduce the Lie algebras. Before that, it is convenient to recall more definitions that are going to be needed not only to define the Lie algebra associated to a Lie group but also in further sections. We begin with the group actions and the group representations.

Definition 2.16. ***Group actions.*** *An action of a group G on a set X is a mapping*

$$a : G \times X \to X,$$

satisfying:

i) $a(e, x) = x$ for all $x \in X$.

ii) $a(g_1, a(g_2, x)) = a(g_1 g_2, x)$ for all $x \in X$ and all $g_1, g_2 \in G$.

The action is said to be smooth if the map a is differentiable.

Definition 2.17. ***Group representations.*** *A group representation is a linear action of a group G on a vector space X. Thus, a representation is an action which fulfils:*

iii) $a(g, \alpha x_1 + \beta x_2) = \alpha a(g, x_1) + \beta a(g, x_2)$ for all $g \in G$, all $x_1, x_2 \in X$ and for all scalars α, β.

Indeed, we are going to consider smooth actions of a Lie group G on itself. Simple examples of these actions are the left and right actions:

$$L_{g_1}(g) \equiv L(g_1, g) = g_1 g; \qquad R_{g_1}(g) \equiv R(g_1, g) = g g_1^{-1}.$$

Notice that we have to use the inverse of g_1 in the definition of the right action to cover non commutative cases and therefore so that property (ii) of definition 2.16 is satisfied.
Combining the two actions defined above we can introduce the adjoint action:

$$\overline{Ad}(g, h) = g h g^{-1}.$$

If we fix $g \in G$, we obtain a mapping $\overline{Ad}_g : G \to G$ such that $\overline{Ad}_g(h) = \overline{Ad}(g, h)$. Now we differentiate this mapping at $e \in G$:

$$d_e \overline{Ad}_g : T_e G \to T_e G, \tag{2.7}$$

This result leads to the definition of the adjoint representation Ad as follows.

Definition 2.18. ***Adjoint representation.*** *Let G be a Lie group, the adjoint representation of G is a map:*

$$Ad : G \times T_e G \to T_e G,$$

defined by

$$Ad(g, Y) := d_e \overline{Ad}_g(Y). \tag{2.8}$$

Indeed, $Ad(g, -) = d_e \overline{Ad}_g : T_e G \to T_e G$. Hence, we are going to use the notation $Ad(g)$ instead of $Ad(g, -)$ and also $Ad(g)Y$ instead of $Ad(g, Y)$.

We can verify that the adjoint representation is indeed a group representation:

i) Since $\overline{Ad}_e$ is the identity map on G, it is clear that for all $X \in T_e G$, $Ad(e)X = X$.

ii) For all $X \in T_e G$ and all $g_1, g_2 \in G$:

$$\begin{aligned} Ad(g_1)(Ad(g_2)X) &= ((d_e \overline{Ad}_{g_1}) \circ (d_e \overline{Ad}_{g_2}))(X) = d_e(\overline{Ad}_{g_1} \circ \overline{Ad}_{g_2})(X) \\ &= d_e \overline{Ad}_{g_1 g_2}(X) = Ad(g_1 g_2)X. \end{aligned}$$

iii) Since the differential of a mapping between two differentiable manifolds is linear (see Proposition 2.6) the third condition is also clear: for all $g \in G$, all $X_1, X_2 \in T_eG$ and for all $\alpha, \beta \in \mathbb{R}$

$$Ad(g)(\alpha X_1 + \beta X_2) = d_e\overline{Ad}_g(\alpha X_1 + \beta X_2) = \alpha d_e\overline{Ad}_g X_1 + \beta d_e\overline{Ad}_g X_2 = \alpha Ad(g)X_1 + \beta Ad(g)X_2.$$

Example 2.19. The adjoint representation for a matrix Lie group can be easily computed. Let G be a matrix Lie group. For all $g, h \in G$, the adjoint action is the following product of matrices: $\overline{Ad}(g,h) = ghg^{-1}$. Now we want to differentiate $\overline{Ad}_g : G \to G$ at $e \in G$, where $\overline{Ad}_g(h) = \overline{Ad}(g,h)$. Thus, we assume h to be close to the identity so that $h \simeq e + tY + O(t^2)$, where $Y \in T_eG$. By differentiating $\overline{Ad}_g$ with respect to h in the direction of Y we get:

$$Ad(g)Y = \lim_{\varepsilon \to 0} \frac{g(e+\varepsilon Y)g^{-1} - g(e)g^{-1}}{\varepsilon} = gYg^{-1}. \tag{2.9}$$

Moreover, the following Lemma shows that $Ad(g) \in Aut(T_eG)$, the group of all automorphism of T_eG.

Lemma 2.20. *For all $g \in G$, the map $Ad(g) : T_eG \to T_eG$ is an invertible linear map.*

Proof. We have previously proved that $Ad(g) \circ Ad(h) = Ad(gh)$ for all $g, h \in G$. Thus, the inverse map of $Ad(g)$ is $Ad(g^{-1})$ since $Ad(g) \circ Ad(g^{-1})X = Ad(e)X = X$ for all $X \in T_eG$ and similarly, $Ad(g^{-1}) \circ Ad(g)X = X$ for all $X \in T_eG$. □

Thus, we have also proved that Ad defines a group homomorphism:

$$\begin{aligned} Ad : G &\to Aut(T_eG) \\ g &\mapsto Ad(g). \end{aligned}$$

Indeed, this is a morphism of Lie groups since $Aut(T_eG)$ is isomorph to $GL(n,\mathbb{R})$. Furthermore, the tangent space to $GL(n,\mathbb{R})$ at the identity $Id_{n\times n} \in GL(n,\mathbb{R})$ can be easily obtained. Let $\gamma : (-\varepsilon, \varepsilon) \to GL(n,\mathbb{R})$ be a curve such that $\gamma(0) = Id_{n\times n}$ and $\gamma'(0) \in T_eGL(n,\mathbb{R})$. Then $\gamma'(0) \in gl(n,\mathbb{R})$, the space of all real $n \times n$ matrices. Thus, the differential of Ad at the identity is the following map

$$d_eAd : T_eG \to End(T_eG), \tag{2.10}$$

where $End(T_eG)$ are all the endomorphisms of T_eG. We can now define the following mapping.

Definition 2.21. *Let G be a group, then we define the mapping*

$$ad : T_eG \to End(T_eG),$$

by

$$ad = d_eAd.$$

This map is usually called the adjoint representation associated to T_eG.

Thus, for all $X \in T_eG$:

$$ad(X) : T_eG \to T_eG,$$

is an endomorphism of T_eG and $ad(X)(Y) \in T_eG$. We are going to use the notations $ad(X)(Y)$, $ad(X)Y$ and $ad(X,Y)$ indistinctly.

Example 2.22. Again, we can compute $ad(X,Y)$ for the matrix Lie groups. Let G be a matrix Lie group. From Eq. (2.9) we get that $Ad(g,Y) = gYg^{-1}$ for any $g \in G$, $Y \in T_eG$. In order to compute d_eAd we proceed as follows: for each $Y \in T_eG$, we are going to differentiate the map $g \mapsto Ad(g)Y = gYg^{-1}$ at $e \in G$. Thus, we assume that g is close to the identity so that $g \simeq e + tX + O(t^2)$, where $X \in T_eG$. Moreover, $(e+tX)^{-1} = e - tX + O(t^2)$. Then:

$$\begin{aligned} ad(X,Y) &= \lim_{\varepsilon \to 0} \frac{(e+\varepsilon X)\,Y\left(e - \varepsilon X + O(\varepsilon^2)\right) - Y}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{\varepsilon XY - \varepsilon YX + O(\varepsilon^2)}{\varepsilon} = XY - YX. \end{aligned} \tag{2.11}$$

Thus, for the matrix Lie groups, ad is the usual commutator of matrices.

The following Proposition shows an interesting property of ad. The proof of this Proposition is beyond the scope of this work due to its complexity and therefore we are not going to prove it. See [3] for the proof.

Proposition 2.23. *Let G be a Lie group. For all $X, Y \in T_eG$, the map ad fulfils:*

$$ad(ad(X,Y)) = ad(X) \circ ad(Y) - ad(Y) \circ ad(X), \tag{2.12}$$

where $\circ$ denotes the composition of functions.

On the other hand, this proposition can be easily proved for the matrix Lie groups. From Eq. (2.11) it follows that

$$ad(ad(X,Y))(Z) = ad(X)(ad(Y)(Z)) - ad(Y)(ad(X)(Z)),$$

for any $X, Y, Z \in T_eG$. And consequently $ad(ad(X,Y)) = ad(X) \circ ad(Y) - ad(Y) \circ ad(X)$.

At this point we have all the ingredients to define the Lie algebras associated to the Lie groups. First we define the concept of Lie algebra.

Definition 2.24. ***Lie algebra.*** *Let $\mathfrak{g}$ be a vector space. Then, $\mathfrak{g}$ is said to be a Lie algebra if there exists a bilinear map $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called Lie bracket, that fulfils the following properties:*

i) ***Antisymmetry****: $[X,Y] = -[Y,X]$ for all $X, Y \in \mathfrak{g}$.*

ii) ***Jacobi identity:*** *$[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.*

Lemma 2.25. *The first axiom of the above definition is equivalent to the condition $[X,X] = 0$ for all $X \in \mathfrak{g}$.*

Proof. Let $X, Y \in \mathfrak{g}$, from $[X,Y] = -[Y,X]$ we get $[X,X] = -[X,X]$ by taking $Y = X$, and then $[X,X] = 0$.
On the other hand, from $[X,X] = 0$ we can take $[X+Y, X+Y] = 0$. Then:

$$\begin{aligned} 0 &= [X+Y, X+Y] = [X, X+Y] + [Y, X+Y] \\ &= [X,X] + [X,Y] + [Y,X] + [Y,Y] = [X,Y] + [Y,X]. \end{aligned}$$

Finally $[X,Y] = -[Y,X]$. □

Notice that this Lemma is true since we are studying vector spaces over $\mathbb{R}$. On the other hand, the Lemma is not valid for vector spaces over a field $\mathbb{K}$ with characteristic 2.

The following proposition shows that $\mathfrak{g} := T_eG$ is a Lie algebra with Lie bracket $[-,-]$ defined by $[X,Y] = ad(X,Y)$.

Proposition 2.26. *Let G be a Lie group, let $\mathfrak{g} := T_eG$ and let $[-,-]$ be defined by $[X,Y] = ad(X,Y)$. Then the pair $(\mathfrak{g}, [-,-])$ is a Lie algebra.*

Proof. First of all, T_eG is a $\mathbb{R}$-vector space by definition and we can easily check that $[-,-]$ is bilinear. Let $X, Y, Z \in \mathfrak{g}$ and $\alpha, \beta \in \mathbb{R}$., Then:

$$[\alpha X + \beta Y, Z] = ad(\alpha X + \beta Y)(Z) = \alpha ad(X)(Z) + \beta ad(Y)(Z) = \alpha[X,Z] + \beta[Y,Z].$$

The second equality follows from the linearity of the differential of Ad since:

$$ad(\alpha X + \beta Y) = d_e Ad(\alpha X + \beta Y) = \alpha d_e Ad(X) + \beta d_e(Y).$$

In addition, since any endomorphism of vector spaces is linear we have:

$$[X, \alpha Y + \beta Z] = ad(X)(\alpha Y + \beta Z) = \alpha ad(X)(Y) + \beta ad(X)(Z) = \alpha ad(X) + \beta ad(Y) = \alpha[X,Y] + \beta[X,Z].$$

Furthermore, we are going to show that $[X, X] = 0$ (proving the antisymmetry by Lemma 2.25). From the definition of a tangent space, an element $X \in \mathfrak{g}$ can be written as $\gamma'(0)$ for some $\gamma : (-\varepsilon, \varepsilon) \to G$ such that $\gamma(0) = e$. Let $Y = \rho'(0) \in \mathfrak{g}$ another element. We can now express $[X, Y]$ in terms of γ and ρ. For each $t \in (-\varepsilon, \varepsilon)$ and for each $s \in (-\varepsilon, \varepsilon)$, $\overline{Ad}(\gamma(t))(\rho(s)) = \gamma(t)\rho(s)\gamma(t)^{-1}$. We now differentiate it respect to s and we evaluate it at $s = 0$ to get $Ad(\gamma(t))(Y)$:

$$Ad(\gamma(t))Y = \left(\frac{d}{ds}\gamma(t)\rho(s)\gamma(t)^{-1}\right)|_{s=0}.$$

By differentiating respect to t and evaluating it at $t = 0$ we obtain $ad(X)Y$:

$$ad(X)(Y) = \left(\frac{d}{dt}\left(\frac{d}{ds}\gamma(t)\rho(s)\gamma(t)^{-1}\right)|_{s=0}\right)|_{t=0}.$$

Thus, if we take $Y = X$:

$$\begin{aligned}[X, X] = ad(X)(X) &= \left(\frac{d}{dt}\left(\frac{d}{ds}\gamma(t)\gamma(s)\gamma(t)^{-1}\right)|_{s=0}\right)|_{t=0} \\ &= \left(\frac{d}{dt}\left(\frac{d}{ds}\gamma(s)\right)|_{s=0}\right)|_{t=0} = \left(\frac{d}{dt}X\right)|_{t=0} = 0. \end{aligned} \tag{2.13}$$

We have used that $\gamma(t)\gamma(s) = \gamma(s)\gamma(t)$ for all s, t. This property is going to be proved in the lemma 2.32.
From Proposition 2.23 we have:

$$(ad(X) \circ ad(Y) - ad(Y) \circ ad(X))(Z) = ad(ad(X, Y))(Z),$$

for all $X, Y, Z \in \mathfrak{g}$. This equation is equivalent to

$$ad(X)(ad(Y)(Z)) - ad(Y)(ad(X)(Z)) = ad(ad(X, Y))(Z),$$

and therefore

$$[X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z].$$

Then, using the antisymmetry we obtain the Jacobi identity. □

The Lie algebra of this Proposition is called the Lie algebra associated to a Lie group.

Example 2.27. We have previously shown that the tangent space to $GL(n, \mathbb{R})$ at the identity corresponds to the space of $n \times n$ real matrices $gl(n, \mathbb{R})$. Thus, from Eq. (2.11), $gl(n, \mathbb{R})$ with the commutator of matrices as the Lie bracket form the Lie algebra associated to $GL(n, \mathbb{R})$.
Indeed, for any matrix Lie group G, $T_{Id}G$ and the commutator of matrices form the associated Lie algebra.

Another example of Lie algebra will be derived subsequently: the Lie algebra of the vector fields [2]. As we have stated before, a vector field X can be thought as a mapping $X : \mathcal{D}(M) \to \mathcal{F}(M)$ and it is differentiable if and only if $X : \mathcal{D}(M) \to \mathcal{D}(M)$. Thus, let X and Y be differentiable vector fields on M and $f \in \mathcal{D}(M)$, we can consider the functions $X(Yf)$ and $Y(Xf)$ and the following Lemma.

Lemma 2.28. *Let X and Y be differentiable vector fields on a differentiable manifold M. Then there exists a unique vector field Z such that, for all $f \in \mathcal{D}(M)$, $Z(f) = (XY - YX)(f)$.*

Proof. First, assume the existence of this vector field Z. Let $p \in M$ and let $\widetilde{x} : U \to M$ be a parametrization at p. The expressions for X and Y in this parametrization is $X = \sum_i a_i \frac{\partial}{\partial x_i}$ and $Y = \sum_i = b_i \frac{\partial}{\partial x_i}$, respectively. Then for all $f \in \mathcal{D}(M)$:

$$XY(f) = X\left(\sum_{j=1}^{n} b_j \frac{\partial f}{\partial x_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_i \frac{\partial b_j}{\partial x_i}\frac{\partial f}{\partial x_j} + a_i b_j \frac{\partial^2 f}{\partial x_i \partial x_j}\right),$$

$$YX(f) = Y\left(\sum_{j=1}^{n} a_j \frac{\partial f}{\partial x_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(b_i \frac{\partial a_j}{\partial x_i}\frac{\partial f}{\partial x_j} + a_j b_i \frac{\partial^2 f}{\partial x_i \partial x_j}\right).$$

Since f satisfies the symmetry of second derivatives we get that the expression of Z in the parametrization $\widetilde{x}$ is

$$Z(f) = XY(f) - YX(f) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_i\frac{\partial b_j}{\partial x_i} - b_i\frac{\partial a_j}{\partial x_i}\right)\frac{\partial f}{\partial x_j}, \tag{2.14}$$

which shows the uniqueness of Z. To prove the existence, we define Z_α in each coordinate neighbourhood $\widetilde{x}_\alpha(U_\alpha)$ of a differentiable structure $\{(U_\alpha, \widetilde{x}_\alpha)\}$ on M by Eq. (2.14). By uniqueness, $Z_\alpha = Z_\beta$ on $\widetilde{x}_\alpha(U_\alpha) \cap \widetilde{x}_\beta(U_\beta) \neq \emptyset$, and therefore we can define Z over all M. □

The vector field Z defined by this Lemma is called the bracket $[X, Y] = XY - YX$ of X and Y. Clearly, Z is differentiable. To show that the vector space $Vect(M)$ of vector fields on M with this bracket form a Lie algebra we can show the following Proposition.

Proposition 2.29. *Let X, Y and Z be differentiable vector fields on M, $a, b \in \mathbb{R}$ and $f, g \in \mathcal{D}(M)$, then*

a) ***Bilinearity:*** $[aX + bY, Z] = a[X, Z] + b[Y, Z]$, $[Z, aX + bY] = a[Z, X] + b[Z, Y]$.

b) ***Antisymmetry:*** $[X, Y] = -[Y, X]$.

c) ***Jacobi identity:*** $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$.

d) $[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X$.

Proof. From Eq. (2.14) the properties (a) and (b) are immediate. Now we prove (c) using (a):

$$\begin{aligned}
[[X,Y],Z] &= [XY - YX, Z] = [XY, Z] - [YX, Z] = XYZ - ZXY - YXZ + ZYX,\\
[[Y,Z],X] &= [YZ - ZY, X] = [YZ, X] - [ZY, X] = YZX - XYZ - ZYX + XZY,\\
[[Z,X],Y] &= [ZX - XZ, Y] = [ZX, Y] - [XZ, Y] = ZXY - YZX - XZY + YXZ.
\end{aligned}$$

Adding these three equations we obtain the property (c). Finally we show (d):

$$\begin{aligned}
[fX, gY] &= fX(gY) - gY(fX) = fgXY + fX(g)Y - gfYX - gY(f)X\\
&= fg[X,Y] + fX(g)Y - gY(f)X.
\end{aligned}$$

□

Thus, from properties (a), (b) and (c) it follows that $Vect(M)$ with the bracket defined above form a Lie algebra. The property (d) is not required for this purpose, but it will be used in the following Chapter 3.

2.3 The exponential map

Once we have properly defined the Lie groups and the associated Lie algebras, we can introduce the exponential map that maps the Lie algebra to the corresponding Lie group. This map is going to be crucial in the trajectory interpolation. In order to define it we provide some additional definitions.
As for the previous section, the main reference has been [3].

The first step is defining an action of a Lie group G on $\mathcal{D}(G)$. Let $f \in \mathcal{D}(G)$ and $g, h \in G$, we can define a new function $g \cdot f$ by:

$$(g \cdot f)(h) = f(g^{-1}h). \tag{2.15}$$

On the other hand, we are going also to define the action of G on $Vect(G)$. Let $X \in Vect(G)$ be a vector field on G, $g \in G$ and $f \in \mathcal{D}(G)$:

$$(g \cdot X)(f) = g \cdot (X(g^{-1} \cdot f)). \tag{2.16}$$

Definition 2.30. *A vector field is said to be left invariant if $g \cdot X = X$ for all $g \in G$. We can denote by $Vect^L(G)$ the subspace of all left invariant vector fields. Clearly, $Vect^L(G) \subset Vect(G)$.*

Now we are going to find the expression of X_h, where $X \in Vect^L(G)$ and $h \in G$. Let $g, h \in G$ and $f \in \mathcal{D}(G)$. From Eq. (2.15) we have that $g \cdot f = f \circ L_{g^{-1}}$, where L_g is the left action. Thus, the differential of $g \cdot f$, $d_h(g \cdot f) : T_hG \to \mathbb{R}$ can be expressed as $d_h(g \cdot f) = d_{g^{-1}h} f \circ d_h L_{g^{-1}}$. Then:

$$\begin{aligned} ((g \cdot X)(f))(h) &= (g \cdot (X(g^{-1} \cdot f)))(h) = (X(g^{-1} \cdot f))(g^{-1}h) \\ &= (d_{g^{-1}h}(g^{-1} \cdot f))(X_{g^{-1}h}) \\ &= (d_h f) \circ (d_{g^{-1}h} L_g)(X_{g^{-1}h}). \end{aligned} \tag{2.17}$$

Since X is left invariant we get:

$$((g \cdot X)(f))(h) = (X(f))(h) = (d_h f)(X_h). \tag{2.18}$$

From Eq. (2.17) and Eq. (2.18) we obtain:

$$(d_h f)(X_h) = (d_h f) \circ (d_{g^{-1}h} L_g)(X_{g^{-1}h}),$$

for all $f \in \mathcal{D}(G)$. Consequently, for all $g, h \in G$:

$$X_h = (d_{g^{-1}h} L_g)(X_{g^{-1}h}). \tag{2.19}$$

Indeed, it is usual to define the left invariant vector fields as the vector fields that fulfil Eq. (2.19) [5].

The following lemma is crucial to define the exponential map.

Lemma 2.31. *Let G be a Lie group. The map $X \mapsto X_e$ defines an isomorphism of vector spaces $Vect^L(G) \to \mathfrak{g}$.*

Proof. Let X be a left invariant vector field and let $g, h \in G$. From Eq. (2.19) and considering $h = g$ we obtain $X_g = (d_e L_g)(X_e)$ and therefore X is uniquely defined by X_e and the map $X \mapsto X_e$ is injective.
Now we are going to prove that it is exhaustive. Let $X \in \mathfrak{g}$ and consider the vector field $\overline{X}$ on G defined by $\overline{X}_g = (d_e L_g)(X)$. By this definition, $\overline{X}_e = X$. We need to show that $\overline{X} \in Vect^L(G)$. On the one hand:

$$(\overline{X}(f))(h) = (d_h f)(d_e L_h)(X) \tag{2.20}$$

On the other hand, using the result from Eq.(2.17) we get:

$$((g \cdot \overline{X})(f))(h) = (d_h f) \circ (d_{g^{-1}h} L_g)(\overline{X}_{g^{-1}h}) = (d_h f) \circ (d_{g^{-1}h} L_g)(d_e L_{g^{-1}h})(X). \tag{2.21}$$

Then, we need to show that $d_e L_h = (d_{g^{-1}h} L_g)(d_e L_{g^{-1}h})$. This result can be obtained by differentiating the equality $L_h = L_g L_{g^{-1}h}$ at e. Therefore, we have proved the equality of Eq. (2.20) and Eq. (2.21) i.e., $\overline{X} \in Vect^L(G)$. $\square$

This lemma implies that each element $X \in \mathfrak{g}$ of a Lie algebra defines a unique left invariant vector field, that we are going to denote $\nu(X)$, fulfilling $\nu(X)_e = X$.
Furthermore, as we have previously stated, associated to $\nu(X)$ there is an unique integral curve $\varphi_X : J \subseteq \mathbb{R} \to G$ through e such that:

$$\left.\frac{d}{dt}\varphi_X\right|_{t=0} = \nu(X)_{\varphi(0)} = \nu(X)_e = X.$$

Lemma 2.32. *The integral curve $\varphi_X : J \to G$ satisfies that for any $s, t \in J$, $\varphi_X(s+t) = \varphi_X(s)\varphi_X(t)$.*

Proof. Let $s, t \in J$. We fix s and let t vary in a some small open set $J_0 \subset J$ containing 0 such that $s + t \in J$. Then consider the curves in G, $\alpha(t) = \varphi_X(s+t)$ and $\beta(t) = \varphi_X(s)\varphi_X(t)$. These curves fulfil that $\alpha(0) = \beta(0) = \varphi_X(s)$ since $\varphi_X(0) = e$.
Now we differentiate these curves. On the one hand:

$$\alpha(t)' = \varphi_X(s+t)' = \nu(X)_{\varphi_X(s+t)} = \nu(X)_{\alpha(t)}.$$

To differentiate $\beta(t)$ we express $\beta(t) = (L_{\varphi_X(s)} \circ \varphi_X)(t)$. Then:

$$\begin{aligned}\beta(t)' &= (d_{\varphi_X(t)} L_{\varphi_X(s)})(\varphi_X(t)') \\ &= (d_{\varphi_X(t)} L_{\varphi_X(s)})(\nu(X)_{\varphi_X(t)}) \\ &= \nu(X)_{\varphi_X(s)\cdot\varphi_X(t)} = \nu(X)_{\beta(t)},\end{aligned}$$

where we have used (2.19) in the third equality. Thus, $\alpha(t)$ and $\beta(t)$ are both integral curves for $\nu(X)$ through $\varphi_X(s)$. By uniqueness of integral curves, they correspond to the same curve and consequently $\varphi_X(s+t) = \varphi_X(s)\varphi_X(t)$. $\square$

Any continuous mapping $f : (-\varepsilon, \varepsilon) \to G$ such that $f(0) = e \in G$ and $f(s+t) = f(s)f(t)$ is called a one-parameter subgroup of G. If f is defined for all $t \in \mathbb{R}$, that is, $f : R \to G$, then f is called the one-parameter group of G [4]. Thus, we are going to refer to φ_X as the one-parameter subgroup (or the one-parameter group) associated to X.
The following Lemma provides a tool to extend a one-parameter subgroup to a one-parameter group.

Lemma 2.33. *Let $J \subset \mathbb{R}$. We can extend $\varphi_X : J \to G$ to $\widetilde{\varphi}_X : \mathbb{R} \to G$ such that for all $t \in J, \widetilde{\varphi}_X(t) = \varphi_X(t)$.*

Proof. Let $t \in \mathbb{R}$, and $m, n \in \mathbb{N}$ such that $\frac{t}{m}, \frac{t}{n} \in J$. Then

$$\begin{aligned}\varphi_X\left(\frac{t}{m}\right), \varphi_X\left(\frac{t}{m}\right)^m \in G, \\ \varphi_X\left(\frac{t}{n}\right), \varphi_X\left(\frac{t}{n}\right)^n \in G.\end{aligned}$$

Thus, since $mn \geq m, n$ we have $\frac{t}{mn} \in J$ and therefore

$$\varphi_X\left(\frac{t}{n}\right)^n = \varphi_X\left(\frac{mt}{mn}\right)^n = \varphi_X\left(\frac{t}{mn}\right)^{mn} = \varphi_X\left(\frac{nt}{mn}\right)^m = \varphi_X\left(\frac{t}{m}\right)^m.$$

Thus, the following extension is well defined:

$$\widetilde{\varphi}_X : \mathbb{R} \to G; \qquad \widetilde{\varphi}_X(t) = \varphi_X\left(\frac{t}{n}\right)^n,$$

for large $n \in N$. $\square$

We are going to denote $\widetilde{\varphi}_X$ simply by φ_X.

Corollary 2.34. *$\varphi_X : \mathbb{R} \to G$ is an homomorphism of Lie Groups. It is the one-parameter group associated to X.*

Then, the exponential map can be defined. As we have stated before, this map maps the Lie algebra $\mathfrak{g}$ associated to a Lie group G with the Lie group. The key to its definition is Lemma 2.31 and the existence of the one-parameter group associated to any $X \in \mathfrak{g}$.

Definition 2.35. ***The exponential map.*** *The exponential map is defined by:*

$$\begin{aligned}\exp : \mathfrak{g} &\to G \\ X &\mapsto \exp(X) = \varphi_X(1).\end{aligned} \tag{2.22}$$

Now consider a matrix Lie group G and the corresponding Lie algebra $\mathfrak{g}$. We are going to derive an expression for the left invariant vector field $\nu(X)$ associated to any $X \in \mathfrak{g}$.

Lemma 2.36. *Let G be a matrix Lie group and let $\mathfrak{g}$ be the associated Lie algebra. Let $X \in \mathfrak{g}$ and let $\nu(X)$ the associated left invariant vector field. Then we have that*

$$\nu(X)_g = gX, \qquad \text{for all } g \in G, \tag{2.23}$$

where gX is a matrix product.

Proof. Remember that $\nu(X)_g = d_e L_g(X)$. We can explicitly compute this differential for the matrix Lie groups. Let $\gamma_X : J \to G$ be the one-parameter subgroup associated to X. Then, $\gamma_X(0) = e \in G$ and $\gamma'_X(0) = X$ and from Proposition 2.6, $d_e L_g(X) = (L_g \circ \gamma_X)'(0) = Lg \circ \gamma'_X(0) = gX$. □

Now we are going to focus on $G = GL(n, \mathbb{R})$ and its corresponding Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ and we are going to obtain an expression for the exponential map. In order to do this we can take for each $X \in \mathfrak{gl}(n, \mathbb{R})$ the mapping $\gamma_X : \mathbb{R} \to GL(n, \mathbb{R})$ defined by:

$$\gamma_X(t) = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!}.$$

Noting by e^A the matrix exponential:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

We can express $\gamma_X(t)$ as:

$$\gamma_X(t) = e^{tX}.$$

It can be shown that $\gamma_X(t) \in GL(n, \mathbb{R})$ since it has an inverse: $\gamma_X(t)^{-1} = \gamma_{-X}(t)$, where $-X \in \mathfrak{g}$. To prove this result we should prove before the following lemma.

Lemma 2.37. *Let A be a real square matrix and let $p, q \in \mathbb{R}$. Then*

$$e^{A(p+q)} = e^{Ap} e^{Aq}.$$

Proof. From the definition of the matrix exponential:

$$e^{Ap} e^{Aq} = \left(\sum_{j=0}^{\infty} \frac{A^j p^j}{j!} \right) \left(\sum_{k=0}^{\infty} \frac{A^k q^k}{k!} \right) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^{j+k} p^j q^k}{j! k!}.$$

Now let $n = j + k$ and therefore $k = n - j$. Thus:

$$e^{Ap} e^{Aq} = \sum_{j=0}^{\infty} \sum_{n=j}^{\infty} \frac{A^n p^j q^{n-j}}{j!(n-j)!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} \sum_{j=0}^{n} \frac{n!}{j!(n-j)!} p^j q^{n-j} = \sum_{n=0}^{\infty} \frac{A^n (p+q)^n}{n!} = e^{A(p+q)}.$$

□

By taking $p = 1$ and $q = -1$ on the one hand, and $p = -1$ and $q = 1$ on the other hand we get:

$$e^A e^{-A} = e^{A(1-1)} = e^0 = Id;$$
$$e^{-A} e^A = e^{A(-1+1)} = e^0 = Id.$$

Now we replace A by tA and we obtain:

$$\gamma_X(t)\gamma_{-X}(t) = e^{tX} e^{-tX} = Id;$$
$$\gamma_{-X}(t)\gamma_X(t) = e^{-tX} e^{tX} = Id.$$

Consequently, $\gamma_X(t)^{-1} = \gamma_{-X}(t)$ and therefore $\gamma_X(t) \in GL(n, \mathbb{R})$.

Indeed, from Lemma 2.37 we have also shown that γ_X is a one-parameter group since $\gamma_X(0) = Id_{n\times n} \in GL(n, \mathbb{R})$. To show that it is the one-parameter subgroup associated to X we have to show that it is the integral curve corresponding to $\nu(X)$. First, by computing the derivative of $\gamma_X(t)$ at $t = 0$ we get $\gamma'_X(0) = X$. Now let $t' \in \mathbb{R}$, $\gamma_X(t') \in GL(n, \mathbb{R})$ and define $s \in \mathbb{R}$ by $s = t - t'$:

$$\left.\frac{d\gamma_X(t)}{dt}\right|_{t=t'} = \left.\frac{d\gamma_X(t'+s)}{ds}\right|_{s=0} = \left.\frac{d}{ds}(\gamma_X(t')\gamma_X(s))\right|_{s=0} = \gamma_X(t')\gamma'_X(0) = \gamma_X(t')X = \nu(X)_{\gamma_X(t')}.$$

In the last equality we have used the Lemma 2.36. Thus, it is an integral curve to $\nu(X)$ and since $\gamma'_X(0) = X$, by the uniqueness of integral curves γ_X is the one-parameter group associated to $X \in \mathfrak{gl}(n, \mathbb{R})$. Thus, the exponential map in this case is the usual matrix exponential:

$$\exp(X) = \gamma_X(1) = \sum_{k=0}^{\infty} \frac{X^k}{k!} = e^X. \tag{2.24}$$

We are going to finish this section obtaining a result that will be useful in the following Chapter 3. First we prove that the Lie bracket of two left invariant vector fields is a left invariant vector field. Let G be a Lie group, let $X, Y \in Vect^L(G)$, let $f \in \mathcal{D}(G)$ and $h \in G$, then

$$\begin{aligned}((g\cdot[X,Y])(f))(h) &= (g\cdot([X,Y](g^{-1}\cdot f)))(h) = ([X,Y](g^{-1}\cdot f))(g^{-1}h)\\ &= ((XY - YX)(g^{-1}\cdot f)(g^{-1}h) = X(Y(g^{-1}\cdot f))(g^{-1}h) - Y(X(g^{-1}\cdot f))(g^{-1}h)\\ &= X((g\cdot Y)(f))(h) - Y((g\cdot X)(f))(h).\end{aligned}$$

Since $X, Y \in Vect^L(G)$, $g \cdot X = X$ and $g \cdot Y = Y$. Thus:

$$((g\cdot[X,Y])(f))(h) = X(Y(f))(h) - Y(X(f))(h) = [X,Y](f)(h). \tag{2.25}$$

Therefore, $[X, Y] \in Vect^L(G)$.

From Lemma 2.31, since X, Y and also $[X, Y]$ are left invariant vector fields, they are uniquely defined by $X_e, Y_e, [X, Y]_e \in \mathfrak{g}$. Using the notation of this Lemma, $X = \nu(X_e)$, $Y = \nu(Y_e)$ and $[X, Y] = \nu([X, Y]_e)$. It can be shown that $[X, Y]_e = [X_e, Y_e] = ad(X_e, Y_e)$ and therefore we obtain $[X, Y] = \nu([X_e, Y_e])$. Consequently:

$$\nu([X_e, Y_e]) = [\nu(X_e), \nu(Y_e)]. \tag{2.26}$$

2.4 Riemannian geometry

Consider a curve $\gamma(t)$ on a manifold M. If we assume that $\gamma(t)$ represents a trajectory, then its velocity at an arbitrary point is the tangent vector to the curve at this point ($\gamma'(t) \in T_{\gamma(t)}M$). To obtain the acceleration we have to proceed more carefully: we should differentiate the vector field that defines the velocity along the curve $\gamma(t)$. The problem that arises here is that the different tangent spaces $T_{\gamma(t)}M$ defined by the different points of the curve are not interrelated.
Thus, to properly define the acceleration (and higher order derivatives) we have to introduce the Riemannian connection. This section provides the theoretical notions of Riemannian geometry that are needed to do this. We have used [2] as the main reference.

We begin with the definition of the Riemannian manifolds.

Definition 2.38. ***Riemannian manifold and the Riemannian metric.*** *A Riemannian metric on a differentiable manifold M is a correspondence which associates to each point $p \in M$ an inner product $\langle -, -\rangle_x$ defined on the tangent space T_pM, which varies differentially, that is, let $\widetilde{x} : U \subset \mathbb{R}^n \to M$ be a system of coordinates around p, $\widetilde{x}(x_1, \ldots, x_n) = q \in \widetilde{x}(U)$, then $\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q)\rangle_q = g_{ij}(x_1, \ldots, x_n)$ is a differentiable function on U. Additionally, the manifold is said to be Riemannian.*

This definition clearly does not depend on the choice of $\widetilde{x}$.

We are going to denote by $\mathcal{X}(M) \subset Vect(M)$ the set of all vector fields of class $\mathcal{C}^\infty$ on M and now we are going to use the notation $\mathcal{D}^\infty(M)$ to refer to the set of functions of class $\mathcal{C}^\infty$ defined on M. Then, we define the affine connection.

Definition 2.39. ***Affine connection.*** *An affine connection ∇ on a differentiable manifold M is a mapping*

$$\begin{aligned} \nabla : \mathcal{X}(M) \times \mathcal{X}(M) &\to \mathcal{X}(M) \\ (X, Y) &\mapsto \nabla_X Y, \end{aligned}$$

which satisfies the following properties for $X, Y, Z \in \mathcal{X}(M)$ and $f, g \in \mathcal{D}^\infty(M)$:

i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.

ii) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$.

iii) $\nabla_X (fY) = f\nabla_X Y + X(f)Y$.

The concept of the affine connection may not be clear enough and therefore it should be complemented with the covariant derivative. The following proposition introduce it.

Proposition 2.40. *Let M be a differentiable manifold with an affine connection ∇. There exists a unique correspondence which associated to a vector field V along the differentiable curve $\gamma : I \to M$ another vector field $\frac{DV}{dt}$ along γ such that:*

a) $\frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$.

b) $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$, *where W is a vector field along γ and f is a differentiable function on I.*

c) If V is induced by a vector field $Y \in \mathcal{X}(M)$, i.e. $V(t) = Y(\gamma(t))$, then $\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y$.

$\frac{DV}{dt}$ is called the covariant derivative of V along γ.

Proof. First of all we suppose that there exists a vector field $\frac{DV}{dt}$ along γ satisfying the conditions of the proposition. Let $\widetilde{x} : U \subset \mathbb{R}^n \to M$ be a system of coordinates such that $\gamma(I) \cap \widetilde{x}(U) \neq \emptyset$. Let $(x_1(t), x_2(t), \ldots, x_n(t))$ be the local expression of $\gamma(t)$, $t \in I$, in the basis $X_i = \frac{\partial}{\partial x_i}$ associated to $\widetilde{x}$. Additionally, we can express V locally as:

$$V = \sum_{j=1}^{n} v^j X_j,$$

where $v^j = v^j(t)$ and also $X_j = X_j(\gamma(t))$, i.e., $X_j \in \mathcal{X}(M)$. By conditions (a) and (b) we have:

$$\frac{DV}{dt} = \sum_{j=1}^{n} \left(\frac{dv^j}{dt} X_j + v^j \frac{DX_j}{dt} \right).$$

By condition (c) and also using (i) of Definition 2.39 we get for $j = 1, \ldots, n$:

$$\frac{DX_j}{dt} = \nabla_{\frac{d\gamma}{dt}} X_j = \nabla_{\left(\sum_i \frac{dx_i}{dt} X_i\right)} X_j = \sum_{i=1}^{n} \frac{dx_i}{dt} \nabla_{X_i} X_j.$$

Therefore we obtain:

$$\frac{DV}{dt} = \sum_{j=1}^{n} \left(\frac{dv^j}{dt} X_j + v^j \sum_{i=1}^{n} \frac{dx_i}{dt} \nabla_{X_i} X_j \right). \tag{2.27}$$

This expression shows that if there is a vector field $\frac{DV}{dt}$ satisfying the conditions of the proposition, then it is unique.
To prove the existence we define $\frac{DV}{dt}$ in $\widetilde{x}(U)$ by Eq. (2.27). We suppose that it satisfies the desired

properties. Then, if $\widetilde{y}(W)$ is another coordinate neighbourhood, with $\widetilde{y}(W)\cap\widetilde{x}(U)\neq\emptyset$ and we can define $\frac{DV}{dt}$ in $\widetilde{y}(W)$ by Eq. (2.27), both definitions agree in $\widetilde{y}(W)\cap\widetilde{x}(U)$ by the uniqueness of $\frac{DV}{dt}$ in $\widetilde{x}(U)$. Thus, the definition can be extended over all M.

Now we show the that if we define $\frac{DV}{dt}$ in $\widetilde{x}(U)$ by Eq. (2.27), it satisfies the properties. Let $V=\sum_j v^j X_j$ and $W=\sum_j w^j X_j$ be two vector fields, then $V+W=\sum_j (v^j+w^j)X_j$. From Eq. (2.27):

$$\frac{D}{dt}(V+W)=\sum_{j=1}^{n}\left(\frac{dv^j}{dt}X_j+\frac{dw^j}{dt}X_j+(v^j+w^j)\sum_{i=1}^{n}\frac{dx_i}{dt}\nabla_{X_i}X_j\right)=\frac{DV}{dt}+\frac{DW}{dt}.$$

Now if f is a differentiable function on I, $fV=\sum_j (fv^j)X_j$ and thus:

$$\frac{D}{dt}(fV)=\sum_{j=1}^{n}\left(f\frac{dv^j}{dt}X_j+\frac{df}{dt}v^jX_j+fv^j\sum_{i=1}^{n}\frac{dx_i}{dt}\nabla_{X_i}X_j\right)=\frac{df}{dt}V+f\frac{DV}{dt}.$$

Finally we have to show the condition (c). Let V be a vector field induced by $Y=\sum_j y^j X_j\in\mathcal{X}(M)$, then using the conditions of Definition 2.39:

$$\nabla_{\frac{d\gamma}{dt}}Y=\nabla_{\frac{d\gamma}{dt}}\left(\sum_{j=1}^{n}y^jX_j\right)=\sum_{j=1}^{n}\nabla_{\frac{d\gamma}{dt}}y^jX_j=\sum_{j=1}^{n}\left(y_j\nabla_{\frac{d\gamma}{dt}}X_j+\frac{d\gamma}{dt}(y^j)X_j\right).$$

Then, since

$$\frac{d\gamma}{dt}(y^j)=\sum_{i=1}^{n}\frac{dx_i}{dt}X_i(y^j)=\sum_{i=1}^{n}\frac{dx_i}{dt}\frac{\partial y^j}{\partial x_i}=\frac{dy_j}{dt},$$

we have shown that $\nabla_{\frac{d\gamma}{dt}}Y=\frac{DV}{dt}$. $\square$

Subsequently, we are going to define the Riemannian connection. Prior to this we need to show some previous concepts. We begin with the concept of parallelism.

Definition 2.41. ***Parallel vector field.*** *Let M be a differentiable manifold with an affine connection ∇. A vector field V along a curve $\gamma: I\to M$ is called parallel when $\frac{DV}{dt}=0$, for all $t\in I$.*

The following proposition complement this definition.

Proposition 2.42. *Let M be a differentiable manifold with an affine connection ∇. Let $\gamma: I\to M$ be a differentiable curve in M and let V_0 be a vector tangent to M at $\gamma(t_0)$, $t_0\in I$ ($V_0\in T_{\gamma(t_0)}M$). Then there exists a unique parallel vector field V along γ such that $V(t_0)=V_0$.*

$V(t)$ is called the parallel transport of $V(t_0)$ along γ.

Proof. First of all, suppose that this proposition has been proved for the case in which $\gamma(I)$ is contained in a local coordinate neighbourhood. Then by compactness, for any $t_1\in I$, the segment $\gamma([t_0,t_1])\subset M$ can be covered by a finite number of coordinate neighbourhoods, in each of which V can be defined, by hypothesis. From uniqueness, the definitions coincide when the intersections are not empty and consequently we can obtain the definition of V along all of $[t_0,t_1]$.

Thus, we have to prove the proposition when $\gamma(I)$ is contained in a coordinate neighbourhood $\widetilde{x}(U)$ of a system of coordinates $\widetilde{x}: U\subset\mathbb{R}^n\to M$. Now let $\widetilde{x}^{-1}(\gamma(t))=(x_1(t),\ldots,x_n(t))$ be the local expression for $\gamma(t)$ and let $V_0=\displaystyle\sum_j v_0^jX_j$, where $X_j=X_j(\gamma(t_0))$ is $X_j=\frac{\partial}{\partial x_j}$.

Now suppose that there exists a parallel vector field V in $\widetilde{x}(U)$ along γ such that $V(t_0)=V_0$. Then $V=\displaystyle\sum_j v^jX_j$ satisfies:

$$0=\frac{DV}{dt}=\sum_j\frac{dv^j}{dt}X_j+\sum_{i,j}\frac{dx_i}{dt}v^j\nabla_{X_i}X_j.$$

Defining the functions Γ_{ij}^k (called Christoffel symbols) by $\nabla_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k$ we get:

$$0 = \frac{DV}{dt} = \sum_k \left(\frac{dv^k}{dt} + \sum_{i,j} v^j \frac{dx_i}{dt} \Gamma_{ij}^k \right) X_k.$$

This expression is equivalent to the following linear system for $v^1(t), \ldots, v^n(t)$:

$$0 = \frac{dv^k}{dt} + \sum_{i,j} \Gamma_{ij}^k v^j \frac{dx_i}{dt}, \qquad k = 1, \ldots, n.$$

This system has a unique solution satisfying the initial conditions $v^k(t_0) = v_0^k$. Then, if V exists, it is unique. Furthermore, since the system is linear, any solution is defined for all $t \in I$ and consequently, V exists (and, as we have stated, is unique). □

Then, we can establish when a connection is compatible with a Riemannian metric.

Definition 2.43. *Let M be a differentiable manifold with an affine connection ∇ and a Riemannian metric $\langle , \rangle$. The connection is said to be compatible with the metric when for any smooth curve γ and any pair of parallel vector field X and X' along γ, they satisfy $\langle X, X' \rangle = constant$.*

Proposition 2.44. *Let M be a Riemannian manifold. A connection ∇ on M is compatible with a metric if and only if for any vector fields V and W along the differentiable curve $\gamma : I \to M$ we have:*

$$\frac{d}{dt}\langle V, W \rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle, \qquad t \in I. \tag{2.28}$$

Proof. On the one hand, Eq. (2.28) implies that, for any pair of parallel vector fields V and W along γ:

$$\frac{d}{dt}\langle V, W \rangle = 0,$$

and therefore ∇ is compatible with the metric $\langle , \rangle$.
On the other hand, let $t_0 \in I$ and consider an orthonormal basis $\{P_1(t_0), \ldots, P_n(t_0)\}$ of $T_{\gamma(t_0)}(M)$. By the Proposition 2.42 we can extend the vectors $P_i(t_0)$ $(i = 1, \ldots, n)$ along γ by parallel translation. Since ∇ is compatible with the metric, $\{P_1(t), \ldots, P_n(t)\}$ is an orthonormal basis of $T_{\gamma(t)}M$ for any $t \in I$. Thus, any vector fields V and W along γ can be expressed as:

$$V = \sum_i v^i P_i, \qquad W = \sum_i w^i P_i, \qquad i = 1, \ldots, n,$$

where v^i and w^i are differentiable functions on I. Since $P_i(t)$ is parallel, $\frac{DP_i(t)}{dt} = 0$ for any $i = 1, \ldots, n$ and $t \in I$. Consequently:

$$\frac{DV}{dt} = \sum_i \frac{dv^i}{dt} P_i, \qquad \frac{DW}{dt} = \sum_i \frac{dw^i}{dt} P_i.$$

Therefore we can show Eq. (2.28):

$$\left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle = \sum_i \left(\frac{dv^i}{dt} w_i + \frac{dw^i}{dt} v^i \right) = \frac{d}{dt} \left(\sum_i v^i w^i \right) = \frac{d}{dt} \langle V, W \rangle.$$

□

Using this proposition we can show the following corollary.

Corollary 2.45. *A connection ∇ on a Riemannian manifold M is compatible with the metric if and only if*

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \qquad X, Y, Z \in \mathcal{X}(M) \tag{2.29}$$

Proof. Suppose that ∇ is compatible with the metric. Let $m \in M$ and let $\gamma : I \to M$ be a differentiable curve with $\gamma(t_0) = m$, $t_0 \in I$, and $\gamma'(t_0) = X(m)$. Then

$$X(m)\langle Y, Z\rangle = \gamma'(t_0)\langle Y, Z\rangle = \frac{d}{dt}\langle Y(\gamma(t)), Z(\gamma(t))\rangle\big|_{t=t_0}.$$

Then, considering $Y(t) = Y(\gamma(t))$ and $Z(t) = Z(\gamma(t))$ as the restrictions of Y and Z to γ we can assume Y and Z to be vector fields along γ. Applying the proposition 2.44 we get:

$$X(m)\langle Y, Z\rangle = \frac{d}{dt}\langle Y(t), Z(t)\rangle\big|_{t=t_0} = \langle \nabla_{X(m)} Y, Z\rangle_m + \langle Y, \nabla_{X(m)} Z\rangle_m.$$

Since m is arbitrary, we have shown Eq. (2.29). The converse is immediate. □

In addition we establish when an affine connection is said to be symmetric.

Definition 2.46. ***Symmetric affine connections.*** *An affine connection ∇ on a smooth manifold M is said to be symmetric when*

$$\nabla_X Y - \nabla_Y X = [X, Y] \textit{ for all } X, Y \in \mathcal{X}(M).$$

Then, we can define the Riemannian connection.

Theorem 2.47. ***Levi-Civita.*** *Given a Riemannian manifold M, there exists a unique affine connection ∇ on M satisfying the conditions:*

a) ∇ is symmetric.

b) ∇ is compatible with the Riemannian metric.

This connection is known as the Levi-Civita or Riemannian connection on M.

Proof. First of all we are going to suppose the existence of an affine connection ∇ satisfying the stated properties. Thus, using the corollary 2.45 we have:

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \tag{2.30}$$
$$Y\langle Z, X\rangle = \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle, \tag{2.31}$$
$$Z\langle X, Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle. \tag{2.32}$$

Adding (2.30) and (2.31) and subtracting (2.32) we obtain:

$$X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle = \langle [Y, Z], X\rangle + \langle [X, Z], Y\rangle + \langle [X, Y], Z\rangle + 2\langle Z, \nabla_Y X\rangle.$$

Then:

$$\begin{aligned}\langle Z, \nabla_Y X\rangle =& \frac{1}{2}\Big(X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle\Big)\\ &\frac{1}{2}\Big(-\langle [Y, Z], X\rangle - \langle [X, Z], Y\rangle - \langle [X, Y], Z\rangle\Big).\end{aligned} \tag{2.33}$$

Eq. (2.33) shows that ∇ is uniquely determined from the metric $\langle , \rangle$. Thus, if it exists, it will be unique.
To prove the existence we can define the affine connection ∇ by Eq. (2.33). Now we have to check that it satisfies the properties of the theorem.
From the definition of ∇ (Eq. (2.33)) we can immediately check that Eq. (2.29) from Corollary 2.45 is satisfied and therefore, ∇ is compatible with the metric. On the other hand, again from Eq. (2.33) we get:

$$\langle Z, \nabla_X Y\rangle - \langle Z, \nabla_Y X\rangle = \langle Z, [X, Y]\rangle.$$

Since this equation is true for all $Z \in \mathcal{X}(M)$ we have:

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

And therefore ∇ is symmetric. □

Now we can get an expression for the Christoffel symbols. Let $(U, \widetilde{x})$ be a coordinate system. Then we have defined the Christoffel symbols Γ_{ij}^k on U by $\nabla_{X_i} X_j = \sum_k \Gamma_{ij}^k X_k$. From Eq. (2.33) we get:

$$\sum_l \Gamma_{ij}^l g_{lk} = \frac{1}{2}\left(\frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij}\right),$$

where $g_{ij} = \langle X_i, X_j \rangle$. Since the matrix with elements g_{ij} admits an inverse (with elements g^{ij}) we obtain:

$$\Gamma_{ij}^m = \frac{1}{2}\sum_k \left(\frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij}\right) g^{km}. \tag{2.34}$$

From Eq. (2.34) it follows that for the Euclidean space $\mathbb{R}^n$, all the Christoffel symbols are null $\Gamma_{ij}^k = 0$.
In addition, as we have previously shown, if $V = \sum_j v^j X_j$, then

$$\frac{DV}{dt} = \sum_k \left(\frac{dv^k}{dt} + \sum_{i,j} v^j \frac{dx_i}{dt}\Gamma_{ij}^k\right) X_k.$$

Thus, in the Euclidean space $\mathbb{R}^n$, $\frac{DV}{dt}$ is the usual time derivative of V.

We continue with the definition of the curvature of a Riemannian manifold.

Definition 2.48. ***The curvature of a Riemannian manifold.*** *Let ∇ be the Riemannian connection of a Riemannian manifold M. The curvature R of M is a correspondence that associates to every pair $X, Y \in \mathcal{X}(M)$ a mapping $R(X,Y) : \mathcal{X}(M) \to \mathcal{X}(M)$ given by*

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z, \qquad Z \in \mathcal{X}(M).$$

Now let $M = \mathbb{R}^n$. Let X, Y and Z be three arbitrary vector fields. If the vector field Z expressed in the natural coordinates of $\mathbb{R}^n$ is given by $Z = (z_1, \ldots, z_n)$ we get:

$$\nabla_X Z = (Xz_1, \ldots, Xz_n), \qquad \nabla_Y \nabla_X Z = (YXz_1, \ldots, YXz_n).$$

Then, from the Definition (2.35) we get $R(X,Y)Z = 0$.

We are going to finish this section with the definition of the left invariant metrics, that are going to be important in the next Chapter.

Definition 2.49. ***Left invariant metric.*** *A Riemannian metric on a Lie group G is left invariant if $\langle u, v \rangle_y = \langle d_y L_x(u), d_y L_x(v) \rangle_{L_x(y)}$ for all $x, y \in G$, $u, v \in T_y G$. Remember that L_x is the left action $L_x(y) = xy$.*

Now let G be a Lie group and let $\mathfrak{g}$ be the associated Lie algebra. We can extend an inner product defined on the Lie algebra to a left invariant metric (defined on the whole manifold). Let $y \in G$ and let $u, v \in T_y G$, then we can define the inner product of u and v by:

$$\langle u, v \rangle_y = \langle d_y L_{y^{-1}}(u), d_y L_{y^{-1}}(v) \rangle_e. \tag{2.35}$$

From the Definition 2.49 we observe that the metric defined by Eq. (2.35) is left invariant.

Indeed, if G is a matrix Lie group we can obtain a more clear expression for $d_y L_{y^{-1}}(u)$. We first differentiate the equality $L_e = L_{y-1} L_y$ at e obtaining

$$d_e L_e = d_y L_{y^{-1}} \circ d_e L_y.$$

Therefore, let $u \in \mathfrak{g}$:

$$d_e L_e(u) = d_y L_{y^{-1}}(d_e L_y(u)).$$

Since $d_e L_e = id_{\mathfrak{g}}$ is the identity map in $\mathfrak{g}$ and from Lemma 2.36 $d_e L_y(u) = yu$ we get

$$d_y L_{y^{-1}}(u) = y^{-1} u.$$

Thus, the Riemannian metric corresponding to Eq. (2.35) is:

$$\langle u, v \rangle_y = \langle y^{-1} u, y^{-1} v \rangle_e. \tag{2.36}$$

Chapter 3

Interpolation in $SE(3)$

This Chapter studies the groups $SE(3)$ and $SO(3)$ as Lie groups and derives algorithms to interpolate multiple points in $SE(3)$. In order to do this we are going to provide $SE(3)$ with a chosen left invariant metric and we are going also to introduce the variational calculus on $SE(3)$.

3.1 $SE(3)$ and $SO(3)$ as Lie groups

The motion of a rigid body can be characterized using the special Euclidean group $SE(3)$. Indeed, $SE(3)$ is the group of the rigid motions and includes the translations and the rotations. Thus, we can choose the following matrices to represent an element $A \in SE(3)$:

$$A = \begin{pmatrix} R & \mathbf{d} \\ 0_{1\times 3} & 1 \end{pmatrix}, \tag{3.1}$$

where $R \in SO(3)$ and $\mathbf{d} \in \mathbb{R}^3$. Recall that $SO(3)$ is the group of rotations:

$$SO(3) = \{R \in GL(3,\mathbb{R}) | R^T R = Id, RR^T = Id \text{ and } \det R = 1\}. \tag{3.2}$$

We can show that $SE(3)$ is the semidirect product of $SO(3)$ and $\mathbb{R}^3$. In order to do that we first define the semidirect product [4].

Definition 3.1. ***Semidirect product.*** *Let G be a group and let H and N be two subgroups satisfying:*

i) N is normal in G,

ii) $H \cap N = e$,

iii) $HN = G$.

iv) Let $\theta : H \times N \to N$ be a mapping such that for each $h \in H$, $\theta_h = \theta(h, -) : N \to N$ is a group automorphism (i.e. an isomorphism from N to N). Then, the group operation on G is defined by:

$$(h, n) \cdot_G (h', n') = (h \cdot_H h', n \cdot_N \theta_h(n')), \tag{3.3}$$

If these statements are fulfilled, G is called the semidirect product of H and N and is denoted by $G = H \ltimes_\theta N$.

Choosing the following representations for the elements $R \in SO(3)$ and $\mathbf{d} \in \mathbb{R}^3$ we can verify the conditions $(i), (ii)$ and (iii) of Definition 3.1:

$$\overline{R} = \begin{pmatrix} R & 0_{3\times 1} \\ 0_{1\times 3} & 1 \end{pmatrix}, \qquad \overline{\mathbf{d}} = \begin{pmatrix} Id_{3\times 3} & \mathbf{d} \\ 0_{1\times 3} & 1 \end{pmatrix}.$$

In addition, let $A_1, A_2 \in SE(3)$:

$$A_1 A_2 = \begin{pmatrix} R_1 & \mathbf{d}_1 \\ 0_{1\times 3} & 1 \end{pmatrix} \begin{pmatrix} R_2 & \mathbf{d}_2 \\ 0_{1\times 3} & 1 \end{pmatrix} = \begin{pmatrix} R_1 R_2 & R_1\mathbf{d}_2 + \mathbf{d}_1 \\ 0_{1\times 3} & 1 \end{pmatrix}.$$

Thus, condition (iv) of Definition 3.1 is satisfied for

$$\theta : SO(3) \times \mathbb{R}^3 \to \mathbb{R}^3, \qquad \theta(R, \mathbf{d}) = R\mathbf{d}.$$

Consequently, $SE(3) = SO(3) \ltimes \mathbb{R}^3$.

Moreover, we are going to show that both $SO(3)$ and $SE(3)$ are Lie groups. Since both the orthogonality and the property of having determinant one are preserved under limits, $SO(3)$ is a closed subgroup of $GL(3, \mathbb{R})$ and from Theorem 2.15, $SO(3)$ is a Lie group.
In addition, let A_m be a sequence of matrix in $SE(3)$, we can express these matrices by

$$A_m = \begin{pmatrix} R_m & \mathbf{d}_m \\ 0 & 1 \end{pmatrix}$$

where $R_m \in SO(3)$ and $\mathbf{d}_m \in R^3$. Thus, this sequences converges to a matrix A:

$$A = \begin{pmatrix} R & \mathbf{d} \\ 0 & 1 \end{pmatrix}. \tag{3.4}$$

Indeed, R_m converges to R and $\mathbf{d}_m$ converges to $\mathbf{d}$. As we have previously stated $R \in SO(3)$ and clearly, $\mathbf{d} \in \mathbb{R}^3$. Thus, $A \in SE(3)$ and consequently $SE(3)$ is a closed subgroup of $GL(4, \mathbb{R})$ and therefore it is a Lie group by Theorem 2.15. Indeed, as a manifold, $SE(3)$ is just the Cartesian product $SO(3) \times \mathbb{R}^3$ [4]. It is easy to show that the Cartesian product ot two differentiable manifolds M and N of dimensions m and n, respectively, is a differentiable manifold of dimension $m+n$ [2]. Therefore $\dim(SE(3)) = \dim(SO(3)) + \dim(\mathbb{R}^3) = 3 + 3 = 6$.

The following Lemma shows that the Lie algebra of $SO(3)$ consists of the 3×3 skew-symmetric matrices.

Lemma 3.2. *The Lie algebra of $SO(3)$ is the space of the 3×3 skew-symmetric matrices.*

Proof. First of all, the Lie algebra of $SO(3)$ is $T_{Id}SO(3)$. Then, let $R(t)$ be a curve embedded in $SO(3)$ such that $R(0) = Id$ and consequently $\dot{R}(0) \in so(3)$. From the property $R^T(t)R(t) = Id$ we obtain:

$$R^T(t)R(t) = Id \Rightarrow \dot{R}^T(t)R(t) + R^T(t)\dot{R}(t) = 0 \Rightarrow \dot{R}^T(0) + \dot{R}(0) = Id \Rightarrow \dot{R}(0) = -\dot{R}^T(0).$$

Thus, $T_{Id}SO(3)$ is a vector subspace of the space of 3×3 skew-symmetric matrices. As both vector spaces have dimension 3 we conclude that the Lie algebra of $SO(3)$ is indeed the space of 3×3 skew-symmetric matrices ($so(3) = T_{Id}SO(3)$). □

Notice that we have used the notation $\dot{R}$ to refer to the time derivative of R. Indeed, in this Chapter we are going to use $\dot{f}$ to refer to the time derivative of a function f.

Now we introduce the following notation to represent an skew-symmetric matrix in terms of a vector $\mathbf{r} = (r_1, r_2, r_3)^T \in \mathbb{R}^3$:

$$[\mathbf{r}] := \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}.$$

Then, it follows for any $\mathbf{r}, \mathbf{v} \in \mathbb{R}^3$:

$$[\mathbf{r}]\mathbf{v} = \mathbf{r} \times \mathbf{v}.$$

The following Lemma shows the structure of the Lie algebra corresponding to $SE(3)$, denoted by $se(3)$.

Lemma 3.3. *$se(3)$ can be expressed as:*

$$se(3) = \left\{ \begin{pmatrix} [\boldsymbol{r}] & \boldsymbol{v} \\ 0_{1\times 3} & 0 \end{pmatrix} \mid [\boldsymbol{r}] \in so(3); \boldsymbol{v} \in \mathbb{R}^3 \right\}. \tag{3.5}$$

Proof. Remember that $se(3) = T_{Id}SE(3)$. Let $\gamma(t)$ be a curve embedded in $SE(3)$ such that $\gamma(0) = Id$ and therefore, $\dot{\gamma}(0) \in se(3)$. As we have previously stated, $\gamma(t)$ can be expressed as:

$$\gamma(t) = \begin{pmatrix} R(t) & d(t) \\ 0_{1\times 3} & 1 \end{pmatrix}, \tag{3.6}$$

where $R(t)$ and $d(t)$ are curves embedded in $SO(3)$ and $\mathbb{R}^3$ respectively. Thus, differentiating $\gamma(t)$ at $t = 0$ we get:

$$\dot{\gamma}(0) = \begin{pmatrix} \dot{R}(0) & \dot{d}(0) \\ 0_{1\times 3} & 0 \end{pmatrix} \tag{3.7}$$

By definition, $\dot{R}(0) \in so(3)$ and $\dot{d}(0) \in T_{0_{3\times 1}}\mathbb{R}^3 = \mathbb{R}^3$. □

Attending to Lemma 3.3 we can define the basis $\{L_i\}_{i=1,\dots,6}$ for $se(3)$, where

$$L_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$L_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

As we have previously shown in the Lemma 2.31, for all $S \in se(3)$ there exists a left invariant vector field $\nu(S)$ and furthermore we have an isomorphism between $Vect^L(SE(3))$ and $se(3)$. Thus, since $\{L_i\}_{i=1,\dots,6}$ is a basis of $se(3)$, $\{\nu(L_i)\}_{i=1,\dots,6}$ is a basis of $Vect^L(SE(3))$.

From Lemma 2.36, $\nu(L_i)_A = AL_i$. Moreover, if we consider a left invariant metric of the form of Eq. (2.36), it follows that for any point $A \in SE(3)$, the vectors $\{\nu(L_i)_A\}_{i=1,\dots,6}$ form a basis of the tangent space at that point. Consequently, we can express any vector field X of $SE(3)$ as:

$$X = \sum_{i=1}^{6} X^i \nu(L_i); \qquad X_A = \sum_{i=1}^{6} X^i(A)\nu(L_i)_A = \sum_{i=1}^{6} X^i(A)AL_i, \tag{3.8}$$

where the coefficients $X^i : SE(3) \to \mathbb{R}$ can vary over the manifold $SE(3)$ (if they are constant then the vector field X is left invariant).

3.2 Exponential map corresponding to $SO(3)$

In order to obtain the interpolation algorithms, in further sections we are going to need an expression for the exponential map that maps the Lie algebra $so(3)$ to $SO(3)$.
To show that the exponential map is indeed the exponential of matrices we just have to show that the exponential of an skew-symmetric matrix is a rotation matrix. The other necessary steps of the proof are the same as the steps that we have shown at the end of Section 2.3 to prove that the exponential map for $GL(n, \mathbb{R})$ is also the matrix exponential.

Proposition 3.4. *Given a skew-symmetric matrix $[\boldsymbol{r}] \in so(3)$, $R = e^{[\boldsymbol{r}]} \in SO(3)$.*

Proof. We must verify that $R^T R = Id$, $RR^T = Id$ and $\det R = 1$. First, we have:

$$\left(e^{[\mathbf{r}]}\right)^{-1} = e^{-[\mathbf{r}]} = e^{[\mathbf{r}]^T} = \left(e^{[\mathbf{r}]}\right)^T.$$

Thus $R^{-1} = R^T$ and therefore $R^T R = Id$ and $RR^T = Id$. From this result it also follows that $\det R = \pm 1$. To show that $\det R = 1$ we use the continuity of the determinant as a function of the entries of a matrix combined with the continuity of the map $[\mathbf{r}] \mapsto e^{[\mathbf{r}]}$ and the fact that $\det(e^{[0_{3\times 1}]}) = 1$. $\square$

Thus, the exponential map $\exp : so(3) \to SO(3)$ is the usual matrix exponential. The following lemma gives an expression for the exponential of an skew-symmetric matrix.

Lemma 3.5. *Let $[\boldsymbol{r}] \in so(3)$, then:*

$$\exp[\boldsymbol{r}] = Id + \left(\frac{\sin ||\boldsymbol{r}||}{||\boldsymbol{r}||}\right)[\boldsymbol{r}] + \left(\frac{1-\cos ||\boldsymbol{r}||}{||\boldsymbol{r}||^2}\right)[\boldsymbol{r}]^2, \tag{3.9}$$

where $||-||$ is the Euclidean norm.

Proof. Assume $||\mathbf{r}|| \neq 0$. Remember that for any matrix $M \in GL(n, \mathbb{R})$:

$$\exp M = \sum_{k=0}^{+\infty} \frac{M^k}{k!}.$$

Thus, first of all we compute some powers of $[\mathbf{r}]$:

$$[\mathbf{r}]^2 = \begin{pmatrix} -(r_2^2 + r_3^2) & r_1 r_2 & r_1 r_3 \\ r_1 r_2 & -(r_1^2 + r_3^2) & r_2 r_3 \\ r_1 r_3 & r_2 r_3 & -(r_1^2 + r_2^2) \end{pmatrix};$$

$$[\mathbf{r}]^3 = \begin{pmatrix} 0 & r_3(r_1^2 + r_2^2 + r_3^2) & -r_2(r_1^2 + r_2^2 + r_3^2) \\ -r_3(r_1^2 + r_2^2 + r_3^2) & 0 & r_1(r_1^2 + r_2^2 + r_3^2) \\ r_2(r_1^2 + r_2^2 + r_3^2) & -r_1(r_1^2 + r_2^2 + r_3^2) & 0 \end{pmatrix} = -||\mathbf{r}||^2[\mathbf{r}];$$

$$[\mathbf{r}]^4 = -||\mathbf{r}||^2[\mathbf{r}]^2; \qquad [\mathbf{r}]^5 = -||\mathbf{r}||^2[\mathbf{r}]^3 = ||\mathbf{r}||^4[\mathbf{r}]; \qquad [\mathbf{r}]^6 = ||\mathbf{r}||^4[\mathbf{r}]^2.$$

Consequently, we derive the following formula:

$$\forall n \in \mathbb{N}, \qquad [\mathbf{r}]^n = \begin{cases} (-1)^k ||\mathbf{r}||^{2k}[\mathbf{r}], & \text{for } n = 2k+1 \quad (k \in \mathbb{N}) \\ (-1)^{k-1}||\mathbf{r}||^{2(k-1)}[\mathbf{r}]^2, & \text{for } n = 2k \quad (k \in \mathbb{N}) \end{cases}$$

Then:

$$\begin{aligned} \exp[\mathbf{r}] &= \sum_{k=0}^{+\infty} \frac{[\mathbf{r}]^k}{k!} = \sum_{k=0}^{+\infty} \frac{[\mathbf{r}]^{2k}}{(2k)!} + \sum_{k=0}^{+\infty} \frac{[\mathbf{r}]^{2k+1}}{(2k+1)!} \\ &= Id + \left[\sum_{k=1}^{+\infty} \frac{(-1)^{k-1}||\mathbf{r}||^{2(k-1)}}{(2k)!}\right][\mathbf{r}]^2 + \left[\sum_{k=0}^{+\infty} \frac{(-1)^k ||\mathbf{r}||^{2k}}{(2k+1)!}\right][\mathbf{r}] \\ &= Id + \left(\frac{1-\cos ||\mathbf{r}||}{||\mathbf{r}||^2}\right)[\mathbf{r}]^2 + \left(\frac{\sin ||\mathbf{r}||}{||\mathbf{r}||}\right)[\mathbf{r}]. \end{aligned}$$

Now, let $[\mathbf{r}]$ be the null matrix and therefore $\mathbf{r} = 0_{3\times 1}$. Then we get:

$$\lim_{||\mathbf{r}|| \to 0} \frac{\sin ||\mathbf{r}||}{||\mathbf{r}||} = 1; \qquad \lim_{||\mathbf{r}|| \to 0} \frac{1-\cos ||\mathbf{r}||}{||\mathbf{r}||^2} = \frac{1}{2}.$$

Consequently, $\exp[0_{3\times 1}] = Id$. $\square$

In addition, to properly interpolate the rotation matrices we have to prove that the exponential map is surjective onto $SO(3)$.

Proposition 3.6. *Given $R \in SO(3)$, there exists $\boldsymbol{r} \in \mathbb{R}$ such that $R = \exp([\boldsymbol{r}])$.*

Proof. We are going to equate the terms of $R \in SO(3)$ and $\exp([\mathbf{r}])$ and show that the corresponding equations have solution. We are going to denote the entries of the matrix R by R_{ij}. To compute $\exp([\mathbf{r}])$ we define $\theta = ||\mathbf{r}||$ and $\hat{\mathbf{r}} = (r_1, r_2, r_3)^T$ such that $\mathbf{r} = \theta\hat{\mathbf{r}}$. Thus, from Lemma 3.5:

$$\exp([\hat{\mathbf{r}}]\theta) = Id + \sin\theta[\hat{\mathbf{r}}] + (1 - \cos\theta)[\hat{\mathbf{r}}]^2. \tag{3.10}$$

We now define $\nu_\theta = 1 - \cos\theta$, $c_\theta = \cos\theta$ and $s_\theta = \sin\theta$. From Eq. (3.10) we get:

$$\exp([\hat{\mathbf{r}}]) = \begin{pmatrix} r_1^2\nu_\theta + c_\theta & r_1r_2\nu_\theta - r_3s_\theta & r_1r_3\nu_\theta + r_2s_\theta \\ r_1r_2\nu_\theta + r_3s_\theta & r_2^2\nu_\theta + c_\theta & r_2r_3\nu_\theta - r_1s_\theta \\ r_1r_3\nu_\theta - r_2s_\theta & r_2r_3\nu_\theta + r_1s_\theta & r_3^2\nu_\theta + c_\theta \end{pmatrix} \tag{3.11}$$

Equating (3.11) with the entries of R we obtain:

$$Tr(R) = R_{11} + R_{22} + R_{33} = 1 + 2\cos\theta. \tag{3.12}$$

Now recall that the trace of R is equal to the sum of its eigenvalues. Let $\mathbf{v} \in \mathbb{R}^3$ be an eigenvector of R and let λ be the corresponding eigenvalue. Then we have $R\mathbf{v} = \lambda\mathbf{v}$ and therefore $||R\mathbf{v}|| = ||\lambda\mathbf{v}||$. Since the norm of a rotated vector equals the norm of this vector ($||R\mathbf{v}|| = ||\mathbf{v}||$) we have $|\lambda| = 1$ and therefore $\lambda = e^{i\phi}$. In addition, it is easy to show that $\det(R - Id) = 0$ for any $R \in SO(3)$ and therefore 1 is an eigenvalue of R. Then, let λ_+ and λ_- be the other eigenvalues

$$\lambda_+ + \lambda_- = Tr(R) - 1 \in \mathbb{R},$$

and therefore λ_+ and λ_- are complex conjugates. Then it follows that $-1 \leq Tr(R) \leq 3$ and hence Eq. (3.12) has solution:

$$\theta = \cos^{-1}\left(\frac{Tr(R) - 1}{2}\right). \tag{3.13}$$

We are going to assume that $0 \leq \theta \leq \pi$. Indeed, $\theta \pm 2\pi n$ and $-\theta \pm 2\pi n$ are also valid solutions. Equating the off-diagonal terms of R and Eq. (3.11) we get:

$$R_{32} - R_{23} = 2r_1\sin\theta, \qquad R_{13} - R_{31} = 2r_2\sin\theta, \qquad R_{21} - R_{12} = 2r_3\sin\theta. \tag{3.14}$$

If $\theta \neq 0$ and $\theta \neq \pi$ we can take:

$$\hat{\mathbf{r}} = \frac{1}{2\sin\theta}\begin{pmatrix} R_{32} - R_{23} \\ R_{13} - R_{31} \\ R_{21} - R_{12} \end{pmatrix}. \tag{3.15}$$

If $R = Id$ then $Tr(R) = 3$, $\theta = 0$ and therefore $\hat{\mathbf{r}}$ can be arbitrary chosen. If $\theta = \pi$, from Eq. (3.10) $\hat{\mathbf{r}}$ has to be solution of the equation $[\hat{\mathbf{r}}]^2 = \frac{1}{2}(R - Id)$. □

Thus, we have shown that the exponential map is surjective onto $SO(3)$.
In addition, from proposition 3.6 we also get the expression for the logarithm of a rotation matrix $R \in SO(3)$, defined as the matrix $[\mathbf{r}] \in so(3)$ such that $R = exp([\mathbf{r}])$ and we have shown that this logarithm is not uniquely defined.

Lemma 3.7. *Let $R \in SO(3)$, then exists $[\boldsymbol{r}] \in so(3)$ such that $R = e^{[\boldsymbol{r}]}$. Indeed the norm $||\boldsymbol{r}||$ of $\boldsymbol{r} \in \mathbb{R}^3$ is solution of the following equation:*

$$Tr(R) = 1 + 2\cos||\boldsymbol{r}||. \tag{3.16}$$

Assume $||\boldsymbol{r}|| \in [0, \pi]$. If $||\boldsymbol{r}|| \neq 0$ and $||\boldsymbol{r}|| \neq \pi$ we get

$$[\boldsymbol{r}] = \log R = \frac{||\boldsymbol{r}||}{2\sin||\boldsymbol{r}||}(R - R^T). \tag{3.17}$$

If $||\boldsymbol{r}|| = 0$, $\boldsymbol{r} = 0_{3\times1}$ and if $||\boldsymbol{r}|| = \pi$, $\boldsymbol{r}$ is the solution of $[\boldsymbol{r}]^2 = \frac{\pi^2}{2}(R - Id)$. From Eq. (3.16) we get that the logarithm is not uniquely defined, since if $||\boldsymbol{r}|| \neq 0$, $[\boldsymbol{r}]\left(1 + \left(\frac{2\pi n}{||\boldsymbol{r}||}\right)\right)$ is also a valid solution to $\log R$ and if $||\boldsymbol{r}|| = 0$, then $[2\pi n\hat{\boldsymbol{r}}]$ is a solution to $\log(Id)$, where $\hat{\boldsymbol{r}}$ is an arbitrary unitary vector.

We finish this section providing a physical interpretation of the exponential map. For this purpose we are going to illustrate that every rotation $R \in SO(3)$ can be expressed as a rotation about a given axis by some amount. Let $\hat{\mathbf{r}} \in \mathbb{R}^3$. If we rotate a body with coordinates $\mathbf{q}(t)$ at a constant unit velocity about the axis $\hat{\mathbf{r}}$, the velocity of the point $\dot{\mathbf{q}}$ can be expressed as:

$$\dot{\mathbf{q}}(t) = \hat{\mathbf{r}} \times \mathbf{q}(t) = [\hat{\mathbf{r}}]\mathbf{q}(t).$$

By integrating this equation we obtain:

$$\mathbf{q}(t) = e^{[\hat{\mathbf{r}}]t}\mathbf{q}(0), \tag{3.18}$$

where $\mathbf{q}(0)$ is the initial position of the point. Then, if we rotate about $\hat{\mathbf{r}}$ at unit velocity for θ units of time, it results in a rotation of an angle θ given by

$$R(\hat{\mathbf{r}}, \theta) = e^{[\hat{\mathbf{r}}]\theta} = e^{[\theta\hat{\mathbf{r}}]}. \tag{3.19}$$

Indeed, we have shown that the exponential map is surjective onto $SO(3)$ and therefore any $R \in SO(3)$ can be expressed as a rotation about a fixed axis $\hat{\mathbf{r}} \in \mathbb{R}^3$ through an angle $\theta \in [0, \pi]$. We call the pair $(\hat{\mathbf{r}}, \theta)$ the exponential coordinates of R [6].
Thus, the exponential map maps the exponential coordinates with the corresponding rotation matrix.
In addition, the exponential coordinates allow us to picture $SO(3)$ as a three-dimensional solid ball of radius π, centred at the origin with the antipodal points identified: each point $\mathbf{r} \in \mathbb{R}^3$ of the solid ball represents a rotation by $\theta = ||\mathbf{r}||$ radians (in the right-hand sense) about the line directed from the origin through $\mathbf{r}$. Indeed, this is the description of the projective space $\mathbb{RP}^3$ and consequently, we have illustrated that $SO(3)$ is homeomorphic to $\mathbb{RP}^3$.

3.3 Angular velocity and linear velocity associated to a curve on $SE(3)$

In this section we are going to derive expressions for the angular velocity and linear velocity of a rigid body that will be useful to define an alternative manner to represent the elements of $se(3)$. The main reference has been [7].
Consider a system described in coordinates $\mathbf{q}$, and let $\mathbf{Q}$ be a moving coordinate system. The following equation provides a relation between these coordinates systems:

$$\mathbf{q}(t) = R(t)\mathbf{Q}(t) + \mathbf{d}(t), \tag{3.20}$$

where $R(t) \in SO(3)$ is a rotation and $\mathbf{d}(t) \in \mathbb{R}^3$ is a translation. By differentiating Eq. (3.20) respect to the time we obtain the following expression for the addition of velocities:

$$\dot{\mathbf{q}} = \dot{R}\mathbf{Q} + R\dot{\mathbf{Q}} + \dot{\mathbf{d}} \tag{3.21}$$

Now, we assume that the object under study is at rest in the moving coordinate system ($\dot{\mathbf{Q}} = 0$) and also that this system only rotates ($\dot{\mathbf{d}} = 0$). Indeed, we will suppose that both systems have the same origin ($\mathbf{d} = 0$). Thus, $R(t)$ is the rotation between the systems $\mathbf{Q}$ and $\mathbf{q}$. With these conditions, Eq. (3.21) leads to $\dot{\mathbf{q}} = \dot{R}\mathbf{Q}$. Expressing $\mathbf{Q}$ in terms of $\mathbf{q}$ using (3.20) we obtain:

$$\dot{\mathbf{q}} = \dot{R}R^{-1}\mathbf{q}. \tag{3.22}$$

Similarly, we can assume that the object is at rest in the system $\mathbf{q}$ ($\dot{\mathbf{q}} = 0$) and suppose again that the moving coordinate system only rotates ($\dot{\mathbf{d}} = 0$) and that these systems have the same origin ($\mathbf{d} = 0$). Then, from Eq. (3.21) we get:

$$\dot{\mathbf{Q}} = -R^{-1}\dot{R}\mathbf{Q}. \tag{3.23}$$

The following definition formalizes the latter result.

Definition 3.8. ***Angular velocity.*** *Let $\boldsymbol{Q}$ be the coordinates attached to a rigid body and let $\boldsymbol{q}$ another coordinate system. Let $R(t)$ be the rotation between $\boldsymbol{Q}$ and $\boldsymbol{q}$. Then the angular velocity with respect to $\boldsymbol{Q}$ is defined by:*

$$[\omega_b] = R^{-1}\dot{R}. \tag{3.24}$$

To define the angular velocity we have used the usual skew-symmetric notation since $[\omega_b] \in so(3)$ ($\omega_b \in \mathbb{R}^3$). This result can be easily proved:

$$R^T R = Id \Rightarrow \dot{R}^T R + R^T \dot{R} = 0 \Rightarrow R^T \dot{R} = -\dot{R}^T R = -(R^T \dot{R})^T.$$

On the other hand, if we assume that the object is at rest in the system $\mathbf{q}$ ($\dot{\mathbf{q}} = 0$) and we let the moving coordinate system to have a translation movement with respect to $\mathbf{q}$ ($\mathbf{d} \neq 0$ and $\dot{\mathbf{d}} \neq 0$) we get

$$\dot{\mathbf{Q}} = -R^{-1}\dot{R}\mathbf{Q} - R^{-1}\dot{\mathbf{d}}.$$

Thus, we can define the linear velocity with respect to the moving frame.

Definition 3.9. ***Linear velocity.*** *Let $\boldsymbol{Q}$ be the coordinates attached to a rigid body and let $\boldsymbol{q}$ another coordinate system. Let $R(t)$ be the rotation between $\boldsymbol{Q}$ and $\boldsymbol{q}$ and let $\boldsymbol{d}(t)$ be the corresponding translation. Then the linear velocity with respect to $\boldsymbol{Q}$ is defined by:*

$$\boldsymbol{v}_b = R^{-1}\dot{\boldsymbol{d}}.$$

Now let $\gamma(t) \in SE(3)$ be a curve describing the motion of a rigid body. The velocity of the curve is $V(t) = \dot{\gamma}(t) \in T_{\gamma(t)}SE(3)$, i.e., the vector field tangent to $\gamma(t)$. Then, we can associate to $V(t)$ an element $S(t) \in se(3)$ by

$$S(t) = \gamma^{-1}(t)\dot{\gamma}(t) = \begin{pmatrix} R^{-1}\dot{R} & R^{-1}\dot{\mathbf{d}} \\ 0 & 0 \end{pmatrix}.$$

From the above definitions, if $(\omega(t), \mathbf{v}(t))^T$ are the coordinates of $S(t)$ in the basis $\{L_i\}_{i=1,\ldots,6}$, then $\omega = \omega_b$ is the angular velocity of the rigid body and $\mathbf{v} = \mathbf{v}_b$ is its linear velocity, both expressed in the body frame.
These elements $S(t)$ are called twists and they provide a representation of the velocity of the curve, $\dot{\gamma}$. Indeed, since $\nu(L_i)_{\gamma(t)} = \gamma(t)L_i$, it follows that $(\omega(t), \mathbf{v}(t))^T$ are the coordinates of $\dot{\gamma}$ in the basis $\{\nu(L_i)_{\gamma(t)}\}_{i=1,\ldots,6}$. We are going to refer to $\{\omega, \mathbf{v}\}$ as the vector pair associated to $\dot{\gamma}$ and in addition we are going to use the notation $\dot{\gamma} = \{\omega, \mathbf{v}\}$.

We can extend this notation to any vector field using Eq. (3.8). Let $A \in SE(3)$, then $X_A \in T_A SE(3)$ and

$$A^{-1}X_A = A^{-1}\sum_{i=1}^{6} X^i(A)AL_i = \sum_{i=1}^{6} X^i(A)L_i \in se(3). \tag{3.25}$$

Hence, $(X_1(A), \ldots, X_6(A))^T$ are the coordinates of $A^{-1}X_A$ in the basis $\{L_i\}_{i=1,\ldots,6}$ and therefore we can associate a pair $\{\omega, \mathbf{v}\}$ to X:

$$\omega = (X^1, X^2, X^3)^T, \qquad \mathbf{v} = (X^4, X^5, X^6)^T.$$

As we have stated above, we are going to use the notation $X = \{\omega, \mathbf{v}\}$. ω and $\mathbf{v}$ are usually referred as the rotational and the translational components of X, respectively.

3.4 The Riemannian metric and the Riemannian connection on $SE(3)$

In this section we are going to propose Riemannian metrics associated to $SE(3)$ in order to characterize the corresponding Riemannian connection, that is, to find an expression for $\nabla_X Y$ for any vector fields X and Y.

At the end of Chapter 2 we have described the technique to extend the inner product defined on $se(3)$ to a left invariant Riemannian metric. First, let $T_1, T_2 \in se(3)$ and let t_1, t_2 be the 6×1 vectors of components of T_1 and T_2 with respect to some basis (for example $\{L_i\}_{i=1,\ldots,6}$). Then the inner product of T_1 and T_2 is defined by:

$$\langle T_1, T_2 \rangle_{Id_{4\times4}} = t_1^T W t_2,$$

where W is a positive definite matrix. From Eq. (2.36) we get that for any $V_1, V_2 \in T_A SE(3)$ ($A \in SE(3)$) the inner product of $V1$ and V_2 can be defined as:

$$\langle V_1, V_2 \rangle_A = \langle A^{-1} V_1, A^{-1} \rangle_{Id_{4\times4}}.$$

Thus, let X and Y be two vector fields and let $\{\omega_X, \mathbf{v}_X\}$ and $\{\omega_Y, \mathbf{v}_Y\}$ be the associated vector pairs to X and Y, respectively. For any $A \in SE(3)$ we get:

$$\langle X_A, Y_A \rangle_A = \langle A^{-1} X_A, A^{-1} Y_A \rangle_{Id_{4\times4}} = t_X^T W t_Y, \tag{3.26}$$

where $t_X^T = (X_1, \ldots, X_6)^T = (\omega_X, \mathbf{v}_X)$ and $t_Y^T = (Y_1, \ldots, Y_6)^T = (\omega_Y, \mathbf{v}_Y)$.

From now on we are going to ignore the subscript of the Riemannian metric.

From the above discussion, to choose a Riemannian metric on $SE(3)$ we just have to define the matrix of the inner product of the elements of $se(3)$. In this regard, it is reasonable to propose the following matrix for the inner product:

$$W = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \qquad A, B \in \mathfrak{gl}(3, \mathbb{R}), \tag{3.27}$$

so that the rotational and the translational components do not interact and therefore we reduce the complexity of Eq. (3.26). More to the point, we are going to study the two following matrices

$$W = \begin{pmatrix} H & 0 \\ 0 & \alpha Id \end{pmatrix}, \tag{3.28}$$

where H is a 3×3 diagonal matrix, and

$$W = \begin{pmatrix} \alpha Id & 0 \\ 0 & \beta Id \end{pmatrix}. \tag{3.29}$$

Indeed, these matrices induce metrics with physical interpretation. Consider a rigid body and let W be a matrix of the form of (3.27). Thus we set $A = M$ and $B = mId$, where M is the inertia tensor and m is the mass of the rigid body. Let $V \in se(3)$ be the vector field corresponding to the velocity of the rigid body, and let $\{\omega, \mathbf{v}\}$ be the vector pair associated to the vector V. Then, the kinetic energy of the rigid body is:

$$K = \frac{1}{2}\langle V, V \rangle = \frac{1}{2}\omega^T M \omega + \frac{1}{2} m \mathbf{v}^T \mathbf{v}.$$

If we assume that the body reference frame is attached at the centroid of the rigid body and also that it is aligned with the principal axes, then M becomes a diagonal matrix

$$M = \begin{pmatrix} M_{xx} & 0 & 0 \\ 0 & M_{yy} & 0 \\ 0 & 0 & M_{zz} \end{pmatrix},$$

where M_{xx}, M_{yy} and M_{zz} are the moments of inertia about the principal axis. The obtained matrix corresponds to the matrix (3.28). Furthermore, if we assume that the rigid body is regular enough, we can suppose that these moments are equal and therefore $M = \alpha Id$, obtaining a matrix of the form of (3.29).

Subsequently, we are going to derive an expression for the Riemannian connection in terms of the proposed metrics. Prior to this, we are going to prove a Lemma that is valid for any left invariant Riemannian metric.

In addition, we are going to use the Einstein summation convention to simplify the notation.

Lemma 3.10. *Let $X = X^i\nu(L_i)$, $Y = Y^i\nu(L_i)$, and $Z = Z^i\nu(L_i)$ be three arbitrary vector fields and let the corresponding vector pairs be $\{\omega_X, \boldsymbol{v}_X\}$, $\{\omega_Y, \boldsymbol{v}_Y\}$ and $\{\omega_Z, \boldsymbol{v}_Z\}$, respectively. Let ∇ be the Riemannian connection corresponding to a left invariant Riemannian metric $\langle -, - \rangle$, then*

$$\begin{aligned}\langle Z, \nabla_X Y\rangle = &\ \langle Z, X(Y^i)\nu(L_i)\rangle \\ &+ \frac{1}{2}\langle\{(\omega_Z \times \omega_Y), (\omega_Z \times \boldsymbol{v}_Y + \boldsymbol{v}_Z \times \omega_Y)\}, \{\omega_X, \boldsymbol{v}_X\}\rangle \\ &+ \frac{1}{2}\langle\{(\omega_Z \times \omega_X), (\omega_Z \times \boldsymbol{v}_X + \boldsymbol{v}_Z \times \omega_X)\}, \{\omega_Y, \boldsymbol{v}_Y\}\rangle \\ &+ \frac{1}{2}\langle\{(\omega_X \times \omega_Y), (\omega_X \times \boldsymbol{v}_Y + \boldsymbol{v}_X \times \omega_Y)\}, \{\omega_Z, \boldsymbol{v}_Z\}\rangle. \end{aligned} \tag{3.30}$$

Proof. Eq. (2.33), developed in Chapter 2, can also be written as:

$$\begin{aligned}\langle Z, \nabla_X Y\rangle =& \frac{1}{2}\Big(X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle\Big) \\ &+ \frac{1}{2}\Big(\langle [Z, Y], X\rangle + \langle [Z, X], Y\rangle + \langle [X, Y], Z\rangle\Big). \end{aligned} \tag{3.31}$$

Thus, we have to compute all the terms of this expression. First of all, from Proposition 2.29 the Lie bracket of any two vector fields holds:

$$[X, Y] = X^iY^j[\nu(L_i), \nu(L_j)] + X(Y^i)\nu(L_i) - Y(X^i)\nu(L_i). \tag{3.32}$$

In addition, from Eq. (2.26) we have $\nu([L_i, L_j]) = [\nu(L_i), \nu(L_j)]$. Moreover, we can compute $[L_i, L_j]$ for all $i, j \in \{1, \dots, 6\}$ as the usual matrix commutator:

$$[L_i, L_i] = 0, \text{ for all } i \in \{1, \dots 6\};$$

$$\begin{array}{lllll} [L_1, L_2] = L_3; & [L_1, L_3] = -L_2; & [L_1, L_4] = 0_{3\times3}; & [L_1, L_5] = L_6; & [L_1, L_6] = -L_5; \\ [L_2, L_3] = L_1; & [L_2, L_4] = -L_6; & [L_2, L_5] = 0_{3\times3}; & [L_2, L_6] = L_4; & [L_3, L_4] = L_5; \\ [L_3, L_5] = -L_4; & [L_3, L_6] = 0_{3\times3}; & [L_4, L_5] = 0_{3\times3}; & [L_4, L_6] = 0_{3\times3}; & [L_5, L_6] = 0_{3\times3}. \end{array}$$

$0_{3\times3}$ stands for the null 3×3 matrix. Thus, the first term of Eq. (3.32) is:

$$\begin{aligned} X^iY^j[\nu(L_i), \nu(L_j)] = &(X^2Y^3 - X^3Y^2)\nu(L_1) \\ &+ (X^3Y^1 - X^1Y^3)\nu(L_2) \\ &+ (X^1Y^2 - X^2Y^1)\nu(L_3) \\ &+ (X^2Y^6 - X^3Y^5 + X^5Y^3 - X^6Y^2)\nu(L_4) \\ &+ (X^3Y^4 - X^1Y^6 + X^6Y^1 - X^4Y^3)\nu(L_5) \\ &+ (X^1Y^5 - X^2Y^4 + X^4Y^2 - X^5Y^1)\nu(L_6) \end{aligned}$$

Therefore, in terms of $\{\omega_X, \mathbf{v}_X\}$ and $\{\omega_Y, \mathbf{v}_Y\}$ we have:

$$X^iY^j[\nu(L_i), \nu(L_j)] = \{\omega_X \times \omega_Y, \omega_X \times \mathbf{v}_Y + \mathbf{v}_X \times \omega_Y\}.$$

Thus, using Eq. (3.32):

$$\begin{aligned}\langle Z, [X, Y]\rangle = &\ \langle Z, X(Y^i)\nu(L_i)\rangle - \langle Z, Y(X^i)\nu(L_i)\rangle \\ &+ \langle\{\omega_Z, \mathbf{v}_Z\}, \{\omega_X \times \omega_Y, \omega_X \times \mathbf{v}_Y + \mathbf{v}_X \times \omega_Y\}\rangle. \end{aligned} \tag{3.33}$$

Furthermore, let W be the matrix associated to the metric $\langle -, - \rangle$ and let W_{ij} be the elements of this matrix. We have:

$$\begin{aligned} X\langle Y, Z\rangle &= X(Y^iW_{ij}Z^j) \\ &= X(Y^i)W_{ij}Z^j + Y^iW_{ij}X(Z^j) \\ &= \langle X(Y^i)\nu(L_i), Z\rangle + \langle Y, X(Z^i)\nu(L_i)\rangle. \end{aligned} \tag{3.34}$$

By computing all the terms in Eq. (3.31) as we have done in (3.33) and (3.34) we get Eq. (3.30). □

Now we are going to focus on the metrics previously introduced.

Proposition 3.11. *Let $X = X^i\nu(L_i)$ and $Y = Y^i\nu(L_i)$ be two arbitrary vector fields. Let ∇ be the Riemannian connection corresponding to the Riemannian metric of Eq. (3.28), then*

$$\nabla_X Y = \left\{ \frac{d\omega_Y}{dt} + \frac{1}{2}\Big(\omega_X \times \omega_Y + H^{-1}(\omega_X \times (H\omega_Y)) + H^{-1}(\omega_Y \times (H\omega_X))\Big), \frac{d\boldsymbol{v}_Y}{dt} + \omega_X \times \boldsymbol{v}_Y \right\},$$

where $\frac{d}{dt}$ is the derivative along the integral curve of X.

Proof. By applying lemma 3.10 to the Riemannain metric of Eq. (3.28) we get:

$$\begin{aligned}
\langle Z, \nabla_X Y\rangle &= \langle Z, X(Y^i)\nu(L_i)\rangle \\
&\quad + \frac{1}{2}\Big((\omega_Z \times \omega_Y)\cdot(H\omega_X) + m(\omega_Z \times \mathbf{v}_Y + \mathbf{v}_Z \times \omega_Y)\cdot \mathbf{v}_X\Big) \\
&\quad + \frac{1}{2}\Big((\omega_Z \times \omega_X)\cdot(H\omega_Y) + m(\omega_Z \times \mathbf{v}_X + \mathbf{v}_Z \times \omega_X)\cdot \mathbf{v}_Y\Big) \\
&\quad + \frac{1}{2}\Big((\omega_X \times \omega_Y)\cdot(H\omega_Z) + m(\omega_X \times \mathbf{v}_Y + \mathbf{v}_X \times \omega_Y)\cdot \mathbf{v}_Z\Big) \\
&= \langle Z, X(Y^i)\nu(L_i)\rangle + m(\omega_X \times \mathbf{v}_Y)\cdot \mathbf{v}_Z \\
&\quad + \frac{1}{2}\Big((\omega_Y \times (H\omega_X))\cdot \omega_Z + (\omega_X \times (H\omega_Y))\cdot\omega_Z + (\omega_X \times \omega_Y)\cdot(H\omega_Z)\Big).
\end{aligned}$$

Here, $\cdot$ represents the scalar product.
Since H is a diagonal matrix, $(\omega_Y \times (H\omega_X))\cdot\omega_Z = H^{-1}(\omega_Y \times (H\omega_X))\cdot(H\omega_Z)$. Thus:

$$\begin{aligned}
\langle Z, \nabla_X Y\rangle &= \langle Z, X(Y^i)\nu(L_i)\rangle + m(\omega_X \times \mathbf{v}_Y)\cdot \mathbf{v}_Z + \frac{1}{2}H^{-1}(\omega_Y \times (H\omega_X))\cdot(H\omega_Z) \\
&\quad + \frac{1}{2}\Big(H^{-1}(\omega_X \times (H\omega_Y))\cdot(H\omega_Z) + (\omega_X \times \omega_Y)\cdot(H\omega_Z)\Big) \\
&= \left\langle Z, \left\{\frac{1}{2}\Big(H^{-1}(\omega_Y \times (H\omega_X)) + H^{-1}(\omega_X \times (H\omega_Y)) + \omega_X \times \omega_Y\Big), \omega_X \times \mathbf{v}_Y\right\}\right\rangle \\
&\quad + \langle Z, X(Y^i)\nu(L_i)\rangle.
\end{aligned}$$

This equality is true for all Z and therefore we have proved the Proposition. □

Corollary 3.12. *Let $X = X^i\nu(Li)$ and $Y = Y^i\nu(L_i)$ be two arbitrary vector fields. If ∇ is the Riemannian connection corresponding to the Riemannian metric of Eq. (3.29), then*

$$\nabla_X Y = \left\{ \frac{d\omega_Y}{dt} + \frac{1}{2}(\omega_X \times \omega_Y), \frac{d\boldsymbol{v}_Y}{dt} + \omega_X \times \boldsymbol{v}_Y \right\}, \tag{3.35}$$

where $\frac{d}{dt}$ is the derivative along the integral curve of X.

Proof. This corollary immediately follows from Proposition 3.11 if we set $H = \alpha Id$. □

Now let $\gamma(t)$ be a curve describing the motion of a rigid body, let V be the velocity (tangent to the curve) and $\{\omega, \mathbf{v}\}$ the corresponding velocity pair. Then, as we have reasoned in the Chapter 2, the acceleration of $\gamma(t)$ corresponds to the covariant derivative $\frac{DV}{dt}$ and from Proposition 2.40, $\frac{DV}{dt} = \nabla_V V$. The expression of the acceleration corresponding to the metric defined by Eq. (3.29) can be obtained using Eq. (3.35):

$$\nabla_V V = \{\dot{\omega}, \dot{\mathbf{v}} + \omega \times \mathbf{v}\}. \tag{3.36}$$

We are going to end this section providing the expression of the Riemannian curvature with the metric of Eq. (3.29). Thus, the following Proposition follows from the definition of the Riemannian curvature (Eq. (2.35)) using Eq. (3.35).

Proposition 3.13. *Let X, Y, and Z be three arbitrary vector fields on $SE(3)$ with the associated vector pairs $\{\omega_X, \boldsymbol{v}_X\}$, $\{\omega_Y, \boldsymbol{v}_Y\}$ and $\{\omega_Z, \boldsymbol{v}_Z\}$. The Riemannian curvature for $SE(3)$ with the Riemannian connection defined by Eq. (3.35) is*

$$R(X,Y)Z = \left\{\frac{1}{4}(\omega_X \times \omega_Y) \times \omega_Z, 0\right\}. \tag{3.37}$$

3.5 Variational calculus on $SE(3)$

The goal of this section is to obtain trajectories between two points in $SE(3)$ minimizing an integral cost function. We are going to consider trajectories minimizing the distance and the acceleration. Before that, we describe the corresponding theoretical framework [2].

Definition 3.14. ***Piecewise differentiable curve.*** *Let M be a differentiable manifold. A piecewise differentiable curve is a continuous mapping $\gamma : [a,b] \to M$ of a closed interval $[a,b] \subset \mathbb{R}$ into M satisfying the following condition: there exists a partition $a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$ of $[a,b]$ such that the restrictions $\gamma|_{[t_i,t_{i+1}]}$, $i = 0,\ldots,k-1$, are differentiable. We say that γ joins the points $\gamma(a)$ and $\gamma(b)$.*

Definition 3.15. ***Variation of a curve.*** *Let $\gamma : [a,b] \to M$ be a piecewise differentiable curve in a manifold M. A variation of γ is a continuous mapping $f : (-\varepsilon, \varepsilon) \times [a,b] \to M$ such that:*

a) $f(0,t) = \gamma(t)$, for all $t \in [a,b]$.

b) There exists a subdivision of $[a,b]$ by points

$$a = t_0 < t_1 < \cdots < t_{k+1} = b,$$

such that the restriction of f to each $(-\varepsilon, \varepsilon) \times [t_i, t_{i+1}]$, $i = 0, 1, \ldots, k$ is differentiable.

A variation is said to be proper if

$$f(s,a) = \gamma(a); \qquad f(s,b) = \gamma(b),$$

for all $s \in (-\varepsilon, \varepsilon)$. If f is differentiable, the variation is said to be differentiable.

Thus, we can define for each $s \in (-\varepsilon, \varepsilon)$ the parametrized curve $f_s : [a,b] \to M$ given by $f_s(t) = f(s,t)$.
In addition, for a variation f, we can define the vector fields $V = \frac{\partial f(s,t)}{\partial f}$ and $S = \frac{\partial f(s,t)}{\partial s}$. Indeed, V is the velocity of $f_s(t)$.
Moreover, if the manifold M is Riemannian, a cost functional on a curve $f_s(t)$ can be defined by

$$J(s) = \int_a^b \left\langle h\left(\frac{\partial f(s,t)}{\partial t}\right), h\left(\frac{\partial f(s,t)}{\partial t}\right)\right\rangle dt,$$

where h is a function of V. Then, it follows the next Lemma.

Lemma 3.16. *If the curve $\gamma(t) = f_0(t)$ is a stationary point of J,*

$$J(s) = \int_a^b \left\langle h\left(\frac{\partial f(s,t)}{\partial t}\right), h\left(\frac{\partial f(s,t)}{\partial t}\right)\right\rangle dt,$$

then $\frac{dJ(s)}{ds}$ must vanish for $s = 0$.

Once we have defined this general frame we can focus on $M = SE(3)$. Subsequently, we are going to study the geodesics.

Definition 3.17. ***Geodesics.*** *A parametrized curve $\gamma : I \to M$ is a geodesic at $t_0 \in I$ if $\frac{D}{dt}(\frac{d\gamma}{dt}) = 0$ at the point t_0. If γ is a geodesic at t, for all $t \in I$, then γ is called a geodesic. If $[a,b] \subset I$ and $\gamma : I \to M$ is a geodesic, the restriction of γ to $[a,b]$ is called a geodesic segment joining $\gamma(a)$ to $\gamma(b)$.*

From Proposition 2.40, if $V = \dot{\gamma}$, then $\frac{DV}{dt} = \nabla_V V$ and therefore, $\gamma : I \to M$ is a geodesic at $t_0 \in I$ if $\nabla_V V = 0$.
Moreover, if $\gamma : I \to M$ is a geodesic, then we have from Proposition 2.44:

$$\frac{d}{dt}\left\langle \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right\rangle = 2\left\langle \frac{D}{dt}\left(\frac{d\gamma}{dt}\right), \frac{d\gamma}{dt} \right\rangle = 0.$$

Thus, the length of the tangent vector $\frac{d\gamma}{dt}$ is constant, $|\frac{d\gamma}{dt}| = c$. If we don't consider the geodesics which reduce to points we can also assume that $c \neq 0$.
Furthermore, the arc length $s : \mathbb{R} \to \mathbb{R}$ of a curve $f : I \to M$ starting from $t_0 = a$ can be defined as

$$s(t) = \int_a^t \left|\frac{df}{dt}\right| dt.$$

Thus, the arc length of a geodesic γ is $s_\gamma = c(t - t_0)$. It can be shown [2] that geodesics locally minimize the arc length.

Moreover, let $\gamma : [a, b] \to M$ be a geodesic and let $f : (-\varepsilon, \varepsilon) \times [a, b] \to M$ be a variation of γ. We are going to show that γ is a critic point of the so-called energy functional:

$$E(s) = \int_a^b \left\langle \frac{df(s,t)}{dt}, \frac{df(s,t)}{dt} \right\rangle dt = \int_a^b \langle V, V\rangle dt. \tag{3.38}$$

Theorem 3.18. *Let γ be a geodesic, then γ is a critical point of the energy function, i.e.* $\frac{\partial E}{\partial s}(0) = 0$.

Proof. Let $f : (-\varepsilon, \varepsilon) \times [a, b] \to M$ be a variation of γ and let $V = \frac{\partial f(s,t)}{\partial t}$ and $S = \frac{\partial f(s,t)}{\partial s}$. Indeed, since the vector fields V and S can be thought as differential operators we can express them as $V = \frac{\partial}{\partial t}$ and $S = \frac{\partial}{\partial s}$ and therefore $Vf = \frac{\partial f}{\partial t}$ and $Sf = \frac{\partial f}{\partial s}$. We also have that $[V, S] = 0$ and since ∇ is symmetric we get $\nabla_V S = \nabla_S V$. Using also that ∇ is compatible with the metric we obtain:

$$\begin{aligned} -\frac{1}{2}\frac{d}{ds}\Big(E(s)\Big) &= -\frac{1}{2} S \int_a^b \langle V, V\rangle dt = -\int_a^b \langle \nabla_S V, V\rangle dt \\ &= -\int_a^b \langle \nabla_V S, V\rangle dt \\ &= \int_a^b \Big(\langle \nabla_V V, S\rangle - V\langle S, V\rangle\Big) dt \\ &= \int_a^b \langle \nabla_V V, S\rangle dt - \langle S, V\rangle|_a^b. \end{aligned} \tag{3.39}$$

Since the initial and final positions are fixed, S vanishes at the endpoints. Thus, when we evaluate the variation $\frac{dE(s)}{ds}$ at $s = 0$ we have that the integral in the last equality must vanish and this is only possible if $\nabla_V V = 0$. Recall that for $s = 0$ we have that $V = \frac{d\gamma}{dt}$ and therefore we have proved the Theorem. □

Now we can derive the equations associated to a geodesic γ in terms of the vector pair $\{\omega, \mathbf{v}\}$ associated to $V = \dot{\gamma}$ by imposing $\nabla_V V = 0$.
From Proposition 3.11, if ∇ is the Riemannian connection associated to the Riemannian metric of Eq. (3.28) and $\{\omega, \mathbf{v}\}$ is the vector pair associated to V, then:

$$\nabla_V V = \{\dot{\omega} + H^{-1}(\omega \times (H\omega)), \dot{\mathbf{v}} + \omega \times \mathbf{v}\}. \tag{3.40}$$

Now remember that $[\omega] = R^T \dot{R}$ and $\mathbf{v} = R^T \dot{\mathbf{d}}$. Thus, using also that $\dot{R}^T = -R^T \dot{R} R^T$, the translation component of $\nabla_V V$ can be written as:

$$\dot{\mathbf{v}} + \omega \times \mathbf{v} = \dot{\mathbf{v}} + [\omega]\mathbf{v} = \dot{R}^T \dot{\mathbf{d}} + R^T \ddot{\mathbf{d}} + R^T \dot{R} R^T \dot{\mathbf{d}} = R^T \ddot{\mathbf{d}}. \tag{3.41}$$

Then, $\nabla_V V = 0$ is equivalent to:

$$\begin{aligned} \dot{\omega} + H^{-1}(\omega \times (H\omega)) &= 0, \\ \ddot{\mathbf{d}} &= 0. \end{aligned} \tag{3.42}$$

On the other hand, if ∇ is the Riemannian connection associated to the Riemannian metric of Eq. (3.29), then:

$$\nabla_V V = \{\dot{\omega}, \dot{\mathbf{v}} + \omega \times \mathbf{v}\}.$$

Consequently, $\nabla_V V = 0$ implies:

$$\begin{aligned} \dot{\omega} &= 0, \\ \ddot{\mathbf{d}} &= 0. \end{aligned} \tag{3.43}$$

This system can easily be solved. The following Proposition shows the obtained analytical solution.

Proposition 3.19. *Given two position of a rigid body,*

$$A(0) = A_1 = \begin{pmatrix} R_1 & \boldsymbol{d}_1 \\ 0 & 1 \end{pmatrix}, \qquad A(1) = A_2 = \begin{pmatrix} R_2 & \boldsymbol{d}_2 \\ 0 & 1 \end{pmatrix},$$

the geodesic between them,

$$A(t) = \begin{pmatrix} R(t) & \boldsymbol{d}(t) \\ 0 & 1 \end{pmatrix},$$

with respect to the metric of Eq. (3.29) is given by:

$$\begin{aligned} R(t) &= R_1 \exp(t \cdot \log(R_1^T R_2)), \\ \boldsymbol{d}(t) &= (\boldsymbol{d}_2 - \boldsymbol{d}_1) \cdot t + \boldsymbol{d}_1, \end{aligned} \tag{3.44}$$

where $t \in [0, 1]$.

Proof. By integrating the first equation of system (3.43) we get $\omega(t) = \omega_0$ and therefore $[\omega(t)] = [\omega_0] = \Omega_0$. On the other hand, we have $[\omega(t)] = R^T(t)\dot{R}(t)$ and consequently $R^T(t)\dot{R}(t) = \Omega_0$. If we integrate this equation we get

$$R(t) = R_0 \exp(\Omega_0 t).$$

From the initial condition $R(0) = R_1$, we have $R_0 = R_1$ and from the boundary condition $R(1) = R_2$ we obtain $\Omega_0 = \log(R_1^T R_2)$.
The second equation of system (3.43) can be integrated twice obtaining $\mathbf{d}(t) = \mathbf{c}_1 t + \mathbf{c}_0$ and from the initial and final conditions we get $\mathbf{d}(t) = (\mathbf{d}_2 - \mathbf{d}_1) \cdot t + \mathbf{d}_1$. The exponential and the logarithm can be computed using the results of Lemma 3.5 and Lemma 3.7. □

Now we are going to study the minimum acceleration trajectories. Let $\gamma : [a, b] \to M$ be the minimum acceleration trajectory between two points in $SE(3)$, and let $f : (-\varepsilon, \varepsilon) \times [a, b] \to M$ be a variation of γ. As we have defined above, $V = \frac{\partial f(s,t)}{\partial t}$. Thus, γ minimize the square of the L^2 norm of the acceleration $\nabla_V V$. Considering the following functional:

$$J_{acc}(s) = \int_a^b \langle \nabla_V V, \nabla_V V \rangle dt, \tag{3.45}$$

γ is a critical point of J_{acc}, i.e. $\frac{dJ(s)}{ds}$ vanishes for $s = 0$. The initial and final points and the initial and final velocities shall be provided.

The following theorem is crucial to obtain the desired trajectories.

Theorem 3.20. *Let $\gamma(t)$ be a curve on a Riemannian manifold that starts and ends at the prescribed points with the prescribed velocities and let $V = \frac{d\gamma}{dt}$. If $\gamma(t)$ minimizes the functional J_{acc}, then*

$$\nabla_V \nabla_V \nabla_V V + R(V, \nabla_V V)V = 0. \tag{3.46}$$

Proof. Let $f : (-\varepsilon, \varepsilon) \times [a, b] \to M$ be a variation of γ and let $V = \frac{\partial f(s,t)}{\partial t}$ and $S = \frac{\partial f(s,t)}{\partial s}$. Again, we can also express V and S as as $V = \frac{\partial}{\partial t}$ and $S = \frac{\partial}{\partial s}$ and therefore $Vf = \frac{\partial f}{\partial t}$ and $Sf = \frac{\partial f}{\partial s}$. We also have that $[V, S] = 0$. We are going to enumerate a set of identities that we will use in the proof. Let X, Y, Z, T, U be arbitrary vector fields, then

1. Since ∇ is a Riemannian connection, it is compatible with the metric: $\langle \nabla_V S, U\rangle = V\langle S, U\rangle - \langle S, \nabla_V U\rangle$.
2. ∇ is also symmetric: $\nabla_V S = \nabla_S V + [V, S] = \nabla_S V$.
3. Let $g(t)$ be an arbitrary differentiable function, then $\int_a^b V(g)dt = g(t)|_a^b$.
4. If $[S, T] = 0$, then $\nabla_S \nabla_T U = \nabla_T \nabla_S U + R(T, S)U$.
5. It can be shown from the definition of the Riemannian curvature that $\langle R(X, Y)Z, T\rangle = \langle R(Z, T)X, Y\rangle$.

Then we have:

$$\begin{aligned}
\frac{1}{2}\frac{d}{ds}\Big(J_{acc}(s)\Big) &= \frac{1}{2}S\int_a^b \langle \nabla_V V, \nabla_V V\rangle dt \overset{(1)}{=} \int_a^b \langle \nabla_S \nabla_V V, \nabla_V V\rangle dt \\
&\overset{(4)}{=} \int_a^b \langle \nabla_V \nabla_S V + R(V, S)V, \nabla_V V\rangle dt \\
&\overset{(1,5)}{=} \int_a^b \Big(V\langle \nabla_S V, \nabla_V V\rangle - \langle \nabla_S V, \nabla_V^2 V\rangle + \langle R(V, \nabla_V V)V, S\rangle\Big)dt \\
&\overset{(2,3)}{=} \int_a^b \Big(\langle R(V, \nabla_V V)V, S\rangle - \langle \nabla_V S, \nabla_V^2 V\rangle\Big)dt + \langle \nabla_S V, \nabla_V V\rangle\Big|_a^b \\
&\overset{(1,3)}{=} \int_a^b \langle \nabla_V^3 V + R(V, \nabla_V V)V, S\rangle dt + \langle \nabla_S V, \nabla_V V\rangle\Big|_a^b - \langle S, \nabla_V^2 V\rangle\Big|_a^b.
\end{aligned}$$

Since the initial and final positions, velocities and accelerations are fixed, S and $\nabla_S V = \nabla_V S$ vanish at the endpoints. Thus, when we evaluate the variation $\frac{dJ_{acc}}{ds}$ at $s = 0$ we have that the integral in the last equality must vanish and this is only possible if Eq. (3.46 holds. Notice that for $s = 0$ we have that $V = \frac{d\gamma}{dt}$ and therefore we have proved the Theorem. □

Thus, we apply this theorem to $SE(3)$ with the Riemannian metric of Eq. (3.29).

Proposition 3.21. *Let*

$$\gamma(t) = \begin{pmatrix} R(t) & \boldsymbol{d}(t) \\ 0 & 1 \end{pmatrix}$$

be a curve between two prescribed points on $SE(3)$ that has prescribed initial and final velocities. Let $\{\omega, \boldsymbol{v}\}$ be the vector pair corresponding to $V = \frac{d\gamma}{dt}$. Assume that ω and $\boldsymbol{d}$ are differentiable. Then, $\gamma(t)$ minimizes the cost function J_{acc} derived from the metric of Eq. (3.29) only if:

$$\begin{aligned}
\omega^{(3)} + \omega \times \ddot{\omega} &= 0, \\
\boldsymbol{d}^{(4)} &= 0,
\end{aligned} \tag{3.47}$$

where $(-)^{(n)}$ denotes the nth derivative of $(-)$.

Proof. We have to compute Eq. (3.46) using the metric of Eq. (3.28). Thus, we have to use the results of Corollary 3.12 and Proposition 3.13.
The second term of Eq. (3.46) can be easily computed:

$$0 = \nabla_V \nabla_V \nabla_V V + R(V, \nabla_V V)V = \nabla_V \nabla_V \nabla_V V + \left\{ \frac{1}{4}(\omega \times \dot{\omega}) \times \omega, 0_{1\times 3} \right\}. \tag{3.48}$$

The rotational component of $\nabla_V \nabla_V \nabla_V V$ can be obtained by applying consecutively Eq. (3.36) and is $\omega^{(3)} + \omega \times \ddot{\omega} + \frac{1}{4}\omega \times (\omega \times \dot{\omega})$. Combining this result with Eq. (3.48) we obtain the desired expression for the rotational component:

$$\omega^{(3)} + \omega \times \ddot{\omega} = 0.$$

To derive the translational part remember that $[\omega] = R^T \dot{R}$ and $\dot{\mathbf{v}} + \omega \times \mathbf{v} = R^T \ddot{\mathbf{d}}$ (where $\mathbf{v} = R^T \dot{\mathbf{d}}$). Using also that $\dot{R}^T = -R^T \dot{R} R^T$ we get that the translational component of $\nabla_V \nabla_V V$ is:

$$\begin{aligned} \frac{d}{dt}\left(R^T \ddot{\mathbf{d}}\right) + \omega \times (R^T \ddot{\mathbf{d}}) &= \frac{d}{dt}\left(R^T \ddot{\mathbf{d}}\right) + [\omega](R^T \ddot{\mathbf{d}}) \\ &= \dot{R}^T \ddot{\mathbf{d}} + R^T \mathbf{d}^{(3)} + R^T \dot{R} R^T \ddot{\mathbf{d}} = R^T \mathbf{d}^{(3)}. \end{aligned}$$

Finally, the translational part of $\nabla_V \nabla_V \nabla_V V$ is:

$$\begin{aligned} \frac{d}{dt}\left(R^T \mathbf{d}^{(3)}\right) + \omega \times (R^T \mathbf{d}^{(3)}) &= \frac{d}{dt}\left(R^T \mathbf{d}^{(3)}\right) + [\omega](R^T \mathbf{d}^{(3)}) \\ &= \dot{R}^T \mathbf{d}^{(3)} + R^T \mathbf{d}^{(4)} + R^T \dot{R} R^T \mathbf{d}^{(3)} = R^T \mathbf{d}^{(4)}. \end{aligned}$$

Thus, the translational component of Eq. (3.48) is:

$$\mathbf{d}^{(4)} = 0.$$

□

The second equation of the system (3.47) can be easily integrated three times obtaining:

$$\mathbf{d} = \mathbf{c}_3 t^3 + \mathbf{c}_2 t^2 + \mathbf{c}_1 t + \mathbf{c}_0, .$$

where $\mathbf{c}_i \in \mathbb{R}^3, i = 0, \ldots, 3$. Thus, to interpolate the position we can use a cubic spline algorithm obtaining a C^2 trajectory [8].
On the other hand, in general, the rotational component Eq. (3.47) can not be solved analytically. Nevertheless, we can prove the following Lemma that provides a particular case to solve this attitude interpolation.

Lemma 3.22. *Given an initial point A_1 and A_2 on $SE(3)$, let $\gamma : [0,1] \to SE(3)$ be a geodesic connecting these two points. Let $V_0 = \dot{\gamma}|_{t_0}$ and $V_1 = \dot{\gamma}|_{t_1}$. If the boundary condition for the minimum acceleration trajectory are of the form $V(0) = \sigma V_0$ and $V(1) = \mu V_1$ and we assume that the minimum acceleration curve is given by $\gamma(p(t))$, then $p(t)$ is a third degree polynomial that satisfies:*

$$\begin{aligned} p(0) &= 0, & p(1) &= 1, \\ \dot{p}(0) &= \sigma, & \dot{p}(1) &= \mu. \end{aligned} \tag{3.49}$$

where $\dot{p} = \frac{dp}{dt}$.

Proof. As the Lemma states, we assume that the minimum acceleration curve, $\beta(t) = \gamma(p(t))$, where $p : \mathbb{R} \to \mathbb{R}$ is an arbitrary scalar function. Thus, $V = \frac{d\beta}{dt} = \dot{p}\frac{d\gamma}{dt}$. We now denote $T = \frac{d\gamma}{dt}$. Since γ is a geodesic (Definition 3.17), $\nabla_T T = 0$. Thus, using conditions (i) and (iii) of Definition 2.39 of the affine connection we get:

$$\nabla_V V = V(\dot{p})T + \dot{p}\dot{p}\nabla_T T = V(\dot{p})T. \tag{3.50}$$

Since $V(\dot{p})$ is the derivative of $\dot{p}$ along β, we have $V(\dot{p}) = \ddot{p}$ and therefore $\nabla_V^n V = p^{(n+1)}T$. From its definition (Definition 2.48), we observe that the curvature is linear and consequently we get:

$$R(V, \nabla_V V) = R(\dot{p}T, \ddot{p}T) = \dot{p}\ddot{p}R(T,T) = 0.$$

Thus, Eq. (3.46) reduces to $p^{(4)}T = 0$. Since T is a tangent vector for a geodesic, it never vanishes and therefore we have $p^{(4)} = 0$. Hence, we obtain a third degree polynomial and applying the boundary conditions we obtain (3.49). □

In the section A of the Appendix we have compared the geodesics with the minimum acceleration trajectories given by Lemma 3.22.

3.6 Solutions for optimal trajectories on $SE(3)$

In this section we are going to derive the algorithms needed to obtain an optimal smooth interpolated trajectory through multiple points.
In section 3.6.1 we describe an attempt to obtain an iterative algorithm to interpolate multiple points by imposing the continuity of the angular velocity and angular acceleration, following the work of F. C. Park and B. Ravani [9].
On the other hand, in section 3.6.2 we are going to propose another procedure: we are going to use a two-point interpolation algorithm to interpolate the trajectory between the ith point and the $(i+1)$th point, together with a discrete convolution in order to smooth the trajectory. Thus, the smoothed trajectory does not go through the initial points, but in fact, interpolation is not a requirement for our purposes, but a way to increase the frequency of the trajectory and therefore this is a valid proceeding.

3.6.1 First procedure to interpolate through multiple points

To obtain an algorithm to interpolate multiple points we are going to first take the minimum acceleration trajectories as a reference.

Let $\gamma(t) \in SE(3)$,

$$\gamma(t) = \begin{pmatrix} R(t) & \mathbf{d}(t) \\ 0 & 1 \end{pmatrix},$$

be the minimum acceleration trajectory between two points in $SE(3)$ and let $V(t) = \dot{\gamma}$ be the velocity, then we have proved that the vector pair $\{\omega, \mathbf{v}\}$ associated to V satisfies:

$$\begin{aligned} \omega^{(3)} + \omega \times \ddot{\omega} &= 0, \\ \mathbf{d}^{(4)} &= 0. \end{aligned} \tag{3.51}$$

As we have previously stated, to interpolate the position we can use a cubic spline algorithm.
On the other hand, we have to study the rotational part. Instead of following the assumption of Lemma 3.22 we are going to make other simplifications.
First, we are going to introduce a set of matrices that provide a decomposition of the skew-symmetric matrices as a sum of mutually annihilating idempotents and will be useful later. These matrices are the following [10]:

$$\begin{aligned} P_0 &= \frac{1}{||\mathbf{r}||^2}\left([\mathbf{r}] + i||\mathbf{r}||Id\right)\left([\mathbf{r}] - i||\mathbf{r}||Id\right) = \frac{1}{||\mathbf{r}||^2}[\mathbf{r}]^2 + Id; \\ P_+ &= -\frac{1}{2||\mathbf{r}||^2}[\mathbf{r}]\left([\mathbf{r}] - i||\mathbf{r}||Id\right) = -\frac{1}{2||\mathbf{r}||^2}[\mathbf{r}]^2 + \frac{i}{2||\mathbf{r}||}[\mathbf{r}]; \\ P_- &= -\frac{1}{2||\mathbf{r}||^2}[\mathbf{r}]\left([\mathbf{r}] + i||\mathbf{r}||Id\right) = -\frac{1}{2||\mathbf{r}||^2}[\mathbf{r}]^2 - \frac{i}{2||\mathbf{r}||}[\mathbf{r}]. \end{aligned}$$

It is easy to show that any skew-symmetric matrix satisfies $[\mathbf{r}]^3 + ||\mathbf{r}||^2[\mathbf{r}] = 0$. Then we can prove that $P_0P_+ = P_0P_- = P_+P_- = 0$:

$$\begin{aligned}
P_0P_+ &= -\frac{1}{2||\mathbf{r}||^4}[\mathbf{r}]\left([\mathbf{r}] + i||\mathbf{r}||Id\right)\left([\mathbf{r}] - i||\mathbf{r}||Id\right)^2 = -\frac{\left([\mathbf{r}]^3 + ||\mathbf{r}||^2[\mathbf{r}]\right)\left([\mathbf{r}] - i||\mathbf{r}||Id]\right)}{2||\mathbf{r}||^4} = 0;\\
P_0P_- &= -\frac{1}{2||\mathbf{r}||^4}[\mathbf{r}]\left([\mathbf{r}] + i||\mathbf{r}||Id\right)\left([\mathbf{r}] - i||\mathbf{r}||Id\right)\left([\mathbf{r}] + i||\mathbf{r}||Id\right)\\
&= -\frac{\left([\mathbf{r}]^3 + ||\mathbf{r}||^2[\mathbf{r}]\right)\left([\mathbf{r}] + i||\mathbf{r}||Id]\right)}{2||\mathbf{r}||^4} = 0;\\
P_+P_- &= \frac{1}{4||\mathbf{r}||^4}[\mathbf{r}]^2\left([\mathbf{r}] - i||\mathbf{r}||Id\right)\left([\mathbf{r}] + i||\mathbf{r}||Id\right) = \frac{[\mathbf{r}]\left([\mathbf{r}]^3 + ||\mathbf{r}||^2[\mathbf{r}]\right)}{4||\mathbf{r}||^4} = 0.
\end{aligned}$$

From their definition we can also observe that:

$$P_0 + P_+ + P_- = Id.$$

We use this property to prove that these matrices are idempotents:

$$\begin{aligned}
P_0 &= (P_0 + P_+ + P_-)P_0 = P_0^2;\\
P_+ &= (P_0 + P_+ + P_-)P_+ = P_+^2;\\
P_0 &= (P_0 + P_+ + P_-)P_- = P_-^2.
\end{aligned} \tag{3.52}$$

As we have previously stated, these matrices provide a decomposition of a general skew-symmetric matrix:

$$[\mathbf{r}] = i||\mathbf{r}||P_- - i||\mathbf{r}||P_+. \tag{3.53}$$

Since these matrices mutually annihilate, the powers of $[\mathbf{r}]$ do not include cross terms. Using also that they are idempotent we obtain the following Proposition.

Proposition 3.23. *Let $[\boldsymbol{r}]$ be a skew-symmetric matrix, its powers satisfy the following equation:*

$$[\boldsymbol{r}]^n = (i||\boldsymbol{r}||)^n P_- + (-i||\boldsymbol{r}||)^n P_+. \tag{3.54}$$

The following Lemma allows us to find a closed analytical expression for the angular velocity.

Lemma 3.24. *Let $\boldsymbol{x}(t)$ be a curve on a matrix Lie algebra and $\boldsymbol{X}(t) = e^{-\boldsymbol{x}(t)}$ a curve on the corresponding matrix Lie group. Then:*

$$\boldsymbol{X}^{-1}(t)\dot{\boldsymbol{X}}(t) = \int_0^1 e^{-s\boldsymbol{x}(t)}\dot{\boldsymbol{x}}(t)e^{s\boldsymbol{x}(t)}ds; \tag{3.55}$$

Proof. Consider the expression $e^{s\mathbf{x}(t)}$, where $s \in \mathbb{R}$ and consider also $\omega(s) = e^{-s\mathbf{x}(t)}\frac{\partial}{\partial t}\left(e^{s\mathbf{x}(t)}\right)$. Then, differentiate $e^{s\mathbf{x}(t)}$ respect to s and t in the two possible orders:

$$\begin{aligned}
\frac{\partial^2}{\partial t \partial s}\left(e^{s\mathbf{x}(t)}\right) &= \frac{\partial}{\partial t}\left(\mathbf{x}(t)e^{s\mathbf{x}(t)}\right) = \mathbf{x}(t)\frac{\partial}{\partial t}\left(e^{s\mathbf{x}(t)}\right) + \dot{\mathbf{x}}(t)e^{s\mathbf{x}(t)}\\
&= \mathbf{x}(t)e^{s\mathbf{x}(t)}e^{-s\mathbf{x}(t)}\frac{\partial}{\partial t}\left(e^{s\mathbf{x}(t)}\right) + \dot{\mathbf{x}}(t)e^{s\mathbf{x}(t)}\\
&= \mathbf{x}(t)e^{s\mathbf{x}(t)}\omega(s) + \dot{\mathbf{x}}(t)e^{s\mathbf{x}(t)},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial s \partial t}\left(e^{s\mathbf{x}(t)}\right) &= \frac{\partial}{\partial s}\left(\frac{\partial}{\partial t}\left(e^{s\mathbf{x}(t)}\right)\right) = \frac{\partial}{\partial s}\left(e^{s\mathbf{x}(t)}e^{-s\mathbf{x}(t)}\frac{\partial}{\partial t}\left(e^{s\mathbf{x}(t)}\right)\right)\\
&= \frac{\partial}{\partial s}\left(e^{s\mathbf{x}(t)}\omega(s)\right)\\
&= e^{s\mathbf{x}(t)}\frac{\partial\omega(s)}{\partial s} + \mathbf{x}(t)e^{s\mathbf{x}(t)}\omega(s),
\end{aligned}$$

By the equality of mixed partials we obtain:

$$\frac{\partial \omega(s)}{\partial s} = e^{-s\mathbf{x}(t)}\dot{\mathbf{x}}(t)e^{s\mathbf{x}(t)}. \tag{3.56}$$

We finally integrate Eq. (3.56):

$$\int_0^1 e^{-s\mathbf{x}(t)}\dot{\mathbf{x}}(t)e^{s\mathbf{x}(t)}ds = \omega(1) - \omega(0) = e^{-\mathbf{x}(t)}\frac{\partial}{\partial t}\left(e^{\mathbf{x}(t)}\right) = \mathbf{X}^{-1}(t)\dot{\mathbf{X}}(t).$$

□

Consequently, if $[\mathbf{r}(t)] = \mathbf{x}(t)$ is a curve on $so(3)$ and $R(t) = e^{[\mathbf{r}(t)]}$ is a curve on $SO(3)$, Eq. (3.55) can be read as:

$$[\omega_b] = R^{-1}\dot{R} = \int_0^1 e^{-s[\mathbf{r}(t)]}[\dot{\mathbf{r}}(t)]e^{s[\mathbf{r}(t)]}ds. \tag{3.57}$$

Now we are going to solve Eq. (3.57) using the adjoint representation of $SO(3)$ and $so(3)$. From the definitions of the adjoint representation of a Lie group (Definition 2.18) and the adjoint representation of the associated Lie algebra (Definition 2.21) we get:

$$[\omega_b] = \int_0^1 Ad\left(e^{-s[\mathbf{r}]}\right)[\dot{\mathbf{r}}]ds = \int_0^1 e^{-ad(s[\mathbf{r}])}[\dot{\mathbf{r}}]ds.$$

Then we integrate term a term the exponential:

$$[\omega_b] = \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k ad^k(s[\mathbf{r}])}{k!}[\dot{\mathbf{r}}]ds = \sum_{k=0}^{\infty} \frac{(-1)^k ad^k([\mathbf{r}])}{(k+1)!}[\dot{\mathbf{r}}],$$

where $ad^0([\mathbf{r}])[\dot{\mathbf{r}}] = [\dot{\mathbf{r}}]$.

The following Lemma is useful to obtain a more simple expression for $[\omega_b]$.

Lemma 3.25. *Let $[\boldsymbol{r}(t)]$ be a curve in $so(3)$, then:*

$$ad^k([\boldsymbol{r}])[\dot{\boldsymbol{r}}] = \left[[\boldsymbol{r}]^k\dot{\boldsymbol{r}}\right] \text{ for all } k \geq 1. \tag{3.58}$$

Notice that $[\boldsymbol{r}]^k\dot{\boldsymbol{r}} \in \mathbb{R}^3$ and then $\left[[\boldsymbol{r}]^k\dot{\boldsymbol{r}}\right] \in so(3)$ is the associated skew-symmetric matrix.

Proof. We are going to prove this lemma by induction on k.

i) $k = 1$.

$$ad([\mathbf{r}])[\dot{\mathbf{r}}] = [\mathbf{r}][\dot{\mathbf{r}}] - [\dot{\mathbf{r}}][\mathbf{r}] =$$

$$\begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -\dot{r}_3 & \dot{r}_2 \\ \dot{r}_3 & 0 & -\dot{r}_1 \\ -\dot{r}_2 & \dot{r}_1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -\dot{r}_3 & \dot{r}_2 \\ \dot{r}_3 & 0 & -\dot{r}_1 \\ -\dot{r}_2 & \dot{r}_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & -(r_1\dot{r}_2 - r_2\dot{r}_1) & r_3\dot{r}_1 - r_1\dot{r}_3 \\ r_1\dot{r}_2 - r_2\dot{r}_1 & 0 & -(r_2\dot{r}_3 - r_3\dot{r}_2) \\ -(r_3\dot{r}_1 - r_1\dot{r}_3) & r_2\dot{r}_3 - r_3\dot{r}_2 & 0 \end{pmatrix} = [\mathbf{r} \times \dot{\mathbf{r}}] = [[\mathbf{r}]\dot{\mathbf{r}}].$$

ii) Now we assume that the lemma is true for $k - 1$.

$$\begin{aligned} ad^k([\mathbf{r}])[\dot{\mathbf{r}}] &= ad([\mathbf{r}])ad^{k-1}([\mathbf{r}])[\dot{\mathbf{r}}] = ad([\mathbf{r}])\left[[\mathbf{r}]^{k-1}\dot{\mathbf{r}}\right] \\ &= [\mathbf{r}]\left[[\mathbf{r}]^{k-1}\dot{\mathbf{r}}\right] - \left[[\mathbf{r}]^{k-1}\dot{\mathbf{r}}\right][\mathbf{r}] \\ &= \left[\mathbf{r} \times [\mathbf{r}]^{k-1}\dot{\mathbf{r}}\right] = \left[[\mathbf{r}][\mathbf{r}]^{k-1}\dot{\mathbf{r}}\right] = \left[[\mathbf{r}]^k\dot{\mathbf{r}}\right]. \end{aligned}$$

□

Thus, we have:

$$[\omega_b] = \sum_{k=0}^{\infty} \frac{(-1)^k ad^k([\mathbf{r}])}{(k+1)!}[\dot{\mathbf{r}}] = \sum_{k=0}^{\infty} \frac{(-1)^k \left[[\mathbf{r}]^k \dot{\mathbf{r}}\right]}{(k+1)!}.$$

Now we express this equation in terms of the vectors that define the skew-symmetric matrices introducing also the matrix W:

$$\omega_b = \sum_{k=0}^{\infty} \frac{(-1)^k [\mathbf{r}]^k \dot{\mathbf{r}}}{(k+1)!} = W\dot{\mathbf{r}}.$$

Using the idempotent matrices introduced before, P_+, P_- and P_0 (see Eq. 3.52) and using the Proposition 3.23 we can compute W:

$$\begin{aligned} W = \sum_{k=0}^{\infty} \frac{[\mathbf{r}]^k}{(k+1)!} &= Id + \sum_{k=1}^{\infty} \frac{(-i||\mathbf{r}||)^k}{(k+1)!} P_- + \sum_{k=1}^{\infty} \frac{(i||\mathbf{r}||)^k}{(k+1)!} P_+ \\ &= P_0 + \sum_{k=0}^{\infty} \frac{(-i||\mathbf{r}||)^k}{(k+1)!} P_- + \sum_{k=0}^{\infty} \frac{(i||\mathbf{r}||)^k}{(k+1)!} P_+ \\ &= P_0 - \frac{1}{i||\mathbf{r}||}(e^{-i||\mathbf{r}||} - 1)P_- + \frac{1}{i||\mathbf{r}||}(e^{i||\mathbf{r}||} - 1)P_+ \\ &= Id - \frac{1-\cos||\mathbf{r}||}{||\mathbf{r}||^2}[\mathbf{r}] + \frac{||\mathbf{r}|| - \sin||\mathbf{r}||}{||\mathbf{r}||^3}[\mathbf{r}]^2. \end{aligned} \tag{3.59}$$

In the second equality we have used that $Id = P_0 + P_- + P_+$. Consequently:

$$\omega_b = \left(Id - \frac{1-\cos||\mathbf{r}||}{||\mathbf{r}||^2}[\mathbf{r}] + \frac{||\mathbf{r}|| - \sin||\mathbf{r}||}{||\mathbf{r}||^3}[\mathbf{r}]^2 \right) \dot{\mathbf{r}}. \tag{3.60}$$

This result can also be read in terms of the skew-symmetric matrix $[\omega_b]$:

$$[\omega_b] = \left(Id - \frac{1-\cos||\mathbf{r}||}{||\mathbf{r}||^2}[\mathbf{r}] + \frac{||\mathbf{r}|| - \sin||\mathbf{r}||}{||\mathbf{r}||^3}[\mathbf{r}]^2 \right) [\dot{\mathbf{r}}].$$

Since $\omega_b \in \mathbb{R}^3$, the covariant derivative of ω_b is the usual time derivative and consequently we can differentiate the expression of Eq. (3.60) to obtain the angular acceleration vector $\alpha_b(t)$ relative to the body frame:

$$\begin{aligned} \alpha_b(t) = \ddot{\mathbf{r}} &- \frac{\mathbf{r}^T\dot{\mathbf{r}}}{||\mathbf{r}||^4}(2\cos||\mathbf{r}|| + ||\mathbf{r}||\sin||\mathbf{r}|| - 2)(\mathbf{r}\times\dot{\mathbf{r}}) + \frac{\cos||\mathbf{r}|| - 1}{||\mathbf{r}||^2}(\mathbf{r}\times\ddot{\mathbf{r}}) \\ &+ \frac{\mathbf{r}^T\dot{\mathbf{r}}}{||\mathbf{r}||^5}(3\sin||\mathbf{r}|| - ||\mathbf{r}||\cos||\mathbf{r}|| - 2||\mathbf{r}||)(\mathbf{r}\times(\mathbf{r}\times\dot{\mathbf{r}})) \\ &+ \frac{||\mathbf{r}|| - \sin||\mathbf{r}||}{||\mathbf{r}||^3}(\dot{\mathbf{r}}\times(\mathbf{r}\times\dot{\mathbf{r}}) + \mathbf{r}\times(\mathbf{r}\times\ddot{\mathbf{r}})). \end{aligned} \tag{3.61}$$

To obtain the trajectory $R(t) \in SO(3)$ between two points ($R_0, R_1 \in SO(3)$) minimizing the angular acceleration, we can first express $R(t)$ as

$$R(t) = R_0 \exp([\mathbf{r}(t)]). \tag{3.62}$$

Then, we have to find the curve $\mathbf{r}(t) \in \mathbb{R}^3$ that minimizes $\int_0^1 ||\alpha_b(t)||^2 dt$ (where $||\cdot||$ is the Euclidean norm) with the conditions

$$\begin{aligned} \mathbf{r}(0) &= 0, & \dot{\mathbf{r}}(0) &= W^{-1}(\mathbf{r}(0))\omega_0, \\ \mathbf{r}(1) &= \mathbf{r}_1, & \dot{\mathbf{r}}(1) &= W^{-1}(\mathbf{r}_1)\omega_1, \end{aligned} \tag{3.63}$$

where $[\mathbf{r}_1] = \log(R_0^T R_1)$, W is given by Eq. (3.59) and ω_0 and ω_1 are the initial and final angular velocities and they have to be provided.

In this regard, we should simplify α_b in order to compute the above integral. If R_0 is close to R_1 in the sense that $R_0^T R_1$ is close to Id, then the boundary condition on $\mathbf{r}(1)$ is $\mathbf{r}(1) \simeq 0_{3\times 1}$. In addition, if we assume that the initial and final angular velocities are not too large, then $\mathbf{r}(t)$ can be expected to remain small. Under this assumption, $\alpha_b(t)$ is approximately $\ddot{\mathbf{r}}$.

Now let $V = \dot{\mathbf{r}}$. Since $\mathbf{r} \in \mathbb{R}^3$ and $\mathbb{R}^3$ is an Euclidean space, $\frac{DV}{dt} = \frac{dV}{dt} = \ddot{\mathbf{r}}$. Indeed, we have shown in Chapter 2 that, as $V(t) = V(\gamma(t))$, $\nabla_V V = \frac{DV}{dt}$. Thus, the integral that has to be minimized is the functional J_{acc} in the Euclidean metric:

$$J_{acc} = \int_0^1 \langle \nabla_V V, \nabla_V V \rangle dt = \int_0^1 \langle \ddot{\mathbf{r}}, \ddot{\mathbf{r}} \rangle dt = \int_0^1 ||\alpha_b||^2 dt.$$

From Theorem 3.20, the curve $\mathbf{r}$ satisfies

$$\nabla_V \nabla_V \nabla_V V + R(V, \nabla_V V)V = 0.$$

Since we are working in an Euclidean space, the Riemannian curvature is null and following the above reasoning, $\nabla_V \nabla_V \nabla_V V = \mathbf{r}^{(4)}$. Thus, the curve $\mathbf{r}$ corresponding to a minimum acceleration trajectory satisfies $\mathbf{r}^{(4)} = 0$. Imposing the above conditions (see 3.63) we get:

$$\mathbf{r}(t) = \mathbf{a}t^3 + \mathbf{b}t^2 + \mathbf{c}t, \tag{3.64}$$

where $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ are constant vectors satisfying

$$\begin{aligned} \mathbf{a} + \mathbf{b} + \mathbf{c} &= \mathbf{r}_1, \\ \mathbf{c} &= \omega_0, \\ 3\mathbf{a} + 2\mathbf{b} + \mathbf{c} &= W^{-1}(\mathbf{r}_1)\omega_1. \end{aligned}$$

Now we are going to present an algorithm for interpolation through multiple points in $SO(3)$. We can derive it by imposing the continuity of angular velocities and angular accelerations at the given points.

We require the following inputs: an ordered set of $n+1$ rotation matrices $\{R_0, R_1, \ldots, R_n\}$ (the knot points); a set of $n+1$ scalars $\{t_0, t_1, \ldots, t_n\}$ (the knot times) such that $t_i < t_{i+1}$ for $i = 0, \ldots, n-1$; the initial angular velocity $\omega_0 \in \mathbb{R}^3$ and the initial angular acceleration $\alpha_0 \in \mathbb{R}^3$. We assume that the knot times are equispaced and therefore $h = t_{i+1} - t_i$ for $i = 0, \ldots, n-1$. Finally we also require the frequency of interpolation ν and it must verify that $\delta = \frac{1}{\nu} < h$. In addition, we are going to assume that $Tr(R_i^T R_{i+1}) \neq -1$ for $i = 0, \ldots, n-1$.
Before iterating the algorithm we should compute the following previous variables. For $i = 1$ to n:

$$\begin{aligned} &[\mathbf{r}_i] = \log(R_{i-1}^T R_i), \\ &W_i = Id - \frac{1 - \cos||\mathbf{r}_i||}{||\mathbf{r}_i||^2}[\mathbf{r}_i] + \frac{||\mathbf{r}_i|| - \sin||\mathbf{r}_i||}{||\mathbf{r}_i||^3}[\mathbf{r}_i]^2. \end{aligned}$$

Then initialize the vectors $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i$ corresponding to $i = 1$:

$$\begin{aligned} \mathbf{c}_1 &= \omega_0, \\ \mathbf{b}_1 &= \frac{\alpha_0}{2}, \\ \mathbf{a}_1 &= \mathbf{r}_1 - \mathbf{b}_1 - \mathbf{c}_1. \end{aligned}$$

Thus we iterate for $i = 2$ to n:

- We set some temporary variables.

$$\begin{aligned} \mathbf{s} &= \mathbf{r}_{i-1}, \\ \mathbf{t} &= 3\mathbf{a}_{i-1} + 2\mathbf{b}_{i-1} + \mathbf{c}_{i-1}, \\ \mathbf{u} &= 6\mathbf{a}_{i-1} + 2\mathbf{b}_{i-1}. \end{aligned}$$

- We obtain the vectors $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i$.

$$
\begin{aligned}
\mathbf{c}_i &= W_{i-1}\mathbf{t}, \\
\mathbf{b}_i &= \frac{1}{2}\left(\mathbf{u} - \frac{\mathbf{s}^T\mathbf{t}}{||\mathbf{s}||^4}(2\cos||\mathbf{s}|| + ||\mathbf{s}||\sin||\mathbf{s}|| - 2)(\mathbf{s}\times\mathbf{t}) - \frac{1-\cos||\mathbf{s}||}{||\mathbf{s}||^2}(\mathbf{s}\times\mathbf{u})\right) \\
&+ \frac{1}{2}\left(\frac{\mathbf{s}^T\mathbf{t}}{||\mathbf{s}||^5}(3\sin||\mathbf{s}|| - ||\mathbf{s}||\cos||\mathbf{s}|| - 2||\mathbf{s}||)(\mathbf{s}\times(\mathbf{s}\times\mathbf{t}))\right) \\
&+ \frac{1}{2}\left(\frac{||\mathbf{s}|| - \sin||\mathbf{s}||}{||\mathbf{s}||^3}(\mathbf{t}\times(\mathbf{s}\times\mathbf{t}) + \mathbf{s}\times(\mathbf{s}\times\mathbf{u}))\right) \\
\mathbf{a}_i &= \mathbf{r}_i - \mathbf{b}_i - \mathbf{c}_i.
\end{aligned}
$$

- Finally, for $t \in \{t_{i-1}, t_{i-1}+\delta, t_{i-1}+2\delta, \ldots, t_i\}$ the result can be computed by:

$$R(t) = R_{i-1}\exp(\mathbf{a}_i\tau^3 + \mathbf{b}_i\tau^2 + \mathbf{c}_i\tau), \qquad \tau = \frac{t - t_{i-1}}{t_i - t_{i-1}}.$$

The equations corresponding to $\mathbf{c}_i$ and $\mathbf{b}_i$ can be obtained by imposing the continuity of the angular velocity and the angular acceleration respectively. The equation corresponding to $\mathbf{a}_i$ ensures the continuity of the trajectory.

In the section B of the Appendix we have computed some trajectories using this algorithm. If the algorithm is used to interpolate a few points ($n \lesssim 5$) the results obtained are satisfactory since the trajectory is smooth. Nevertheless, when we increase the number of points, the angular velocity and acceleration increase and the trajectory has no physical sense.

Indeed, [9] showed that the more points we want to interpolate, the more the trajectory obtained with this algorithm will differ from the real minimum acceleration trajectory. In order to solve this limitation, [9] proposes to set some control points reasonably spaced to minimize the deviations. However, this proceeding is not valid for our purposes since we could have thousands of knot points and therefore we can not set so many control points.
Since we have discarded this procedure, we have not implemented the cubic splines to interpolate the position.

The following section describes the alternative way that we have chosen to interpolate the trajectory.

3.6.2 Second procedure to interpolate through multiple points

In this section we are going to use the two-point interpolation and a kernel convolution to generate an algorithm able to interpolate multiple points. First we have to choose which two-point interpolation procedure we want to implement. Although the attitude interpolation corresponding to the minimum acceleration trajectory can not be computed analytically, we have shown in the previous sections 3.5 and 3.6.1 that we can obtain solutions to the minimum acceleration curve under some conditions and simplifications (see [11] and [9] for other particular cases).
On the other hand, we have shown (Proposition 3.19) that the geodesic between two points of a rigid body can be easily computed. Thus, for its simplicity and also because it implies the least computational cost among all the studied algorithms, we have used the geodesics for our two-point interpolation.
Then, given a trajectory, we are going to interpolate the trajectory between the ith point and the $(i+1)$th point using the equations corresponding to the geodesics. Clearly, if we interpolate multiple points using a two-point interpolation we are going to obtain a trajectory that is not smooth. In order to solve it we introduce the concept of convolution.

Definition 3.26. ***Convolution***. *Let $f(t)$ and $g(t)$ be locally integrable functions on $\mathbb{R}^n$. We define the convolution $(f*g)(t)$ by:*

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{\infty} f(t-\tau)g(\tau)d\tau. \tag{3.65}$$

In addition, the convolution satisfies the following theorems [12].

Theorem 3.27. *Let $g \in \mathcal{C}^j(\mathbb{R}^n)$ be a compact supported function and let f be a locally integrable function on $\mathbb{R}^n$, then $f * g \in \mathcal{C}^j(\mathbb{R}^n)$.*

Theorem 3.28. *Let $g \in \mathcal{C}^\infty(\mathbb{R}^n)$ be a compact supported function such that $g \geq 0$ and $\int g(x)dx = 1$. If f is a compact supported function on $\mathbb{R}^n$, then $f * g \in \mathcal{C}^\infty(\mathbb{R}^n)$.*

Thus, from Theorem 3.28 it follows that the convolution can be used to approximate functions by more differentiable ones. This process is known as regularization and g is called the convolution kernel. The bump functions are suitable convolution kernels since they are compact supported and also infinitely differentiable. Since we can assume that the frequency of interpolation is much lower than $1Hz$, we can use the following bump function:

$$G(t) = \begin{cases} \dfrac{\Phi(t)}{N_G}, & \text{for } |t| \leq 1, \\ 0 & \text{for } |t| > 1. \end{cases}, \qquad \Phi(t) = \exp\left(-\frac{1}{1-t^2}\right), \qquad N_G = \int_{-1}^{1} \Phi(t)dt. \tag{3.66}$$

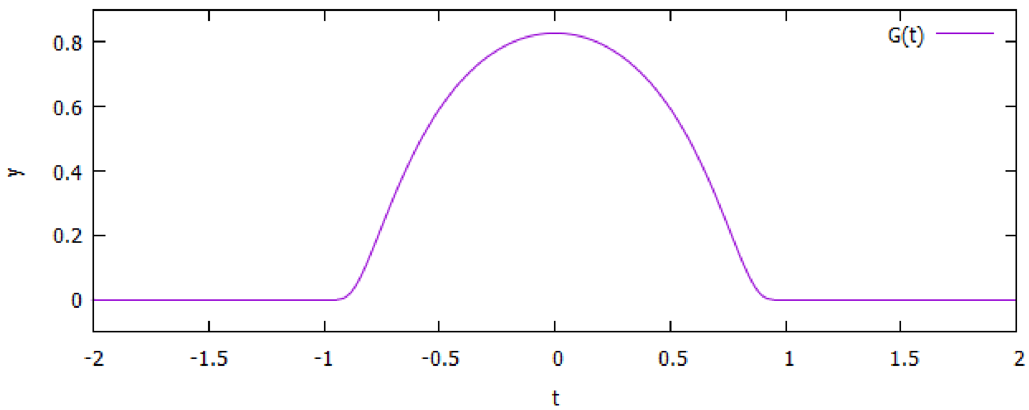

Figure 1: Bump function $G(t)$ corresponding to Eq. (3.66).

Furthermore, since the functions that we want to regularize are discrete, we have to implement a discrete convolution. Let $\{t_0, \ldots, t_n\}$ be the knot times and let $\{t_0, \ldots, t_m\}$ be the interpolation times ($m \geq n$), a discrete convolution can be computed by

$$(f * G)(t_i) = \frac{\sum_{j=0}^{m} f(t_j)\Phi(t_j - t_i)}{\sum_{j=0}^{m} \Phi(t_j)}, \qquad 0 \leq i \leq m. \tag{3.67}$$

Notice that the denominator works as a normalization factor.
In order to reduce the computational cost we have actually implemented the following expressions

to obtain the smoothed function:

$$\tilde{f}_\Delta(t_i) = \frac{\sum_{j=0}^{i+\Delta} f(t_j)\Phi(t_j - t_i)}{\sum_{j=0}^{i+\Delta} \Phi(t_j - t_i)}, \qquad 0 \le i < \Delta,$$
$$\tilde{f}_\Delta(t_i) = \frac{\sum_{j=i-\Delta}^{i+\Delta} f(t_j)\Phi(t_j - t_i)}{\sum_{j=i-\Delta}^{i+\Delta} \Phi(t_j - t_i)}, \qquad \Delta \le i \le m-\Delta, \tag{3.68}$$
$$\tilde{f}_\Delta(t_i) = \frac{\sum_{j=i-\Delta}^{m} f(t_j)\Phi(t_j - t_i)}{\sum_{j=i-\Delta}^{m} \Phi(t_j - t_i)}, \qquad m-\Delta < i \le m,$$

where $\Delta \in \mathbb{N}$ is a parameter that has to be chosen. The more large Δ is, the more smooth the trajectory will be.

Now we present the algorithm to interpolate multiple points in $SE(3)$. We require the following inputs: an ordered set of rotation matrices $\{R_0, R_1, \dots, R_n\}$ and vectors $\{\mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_n\}$ (the knot points); a set of $n+1$ scalars $\{t_0, t_1, \dots, t_n\}$ (the knot times) such that $t_i < t_{i+1}$ for $i = 0, \dots, n-1$. We assume that the knot times are equispaced and therefore $h = t_{i+1} - t_i$ for $i = 0, \dots, n-1$. Finally we require the frequency ν of interpolation such that $\delta = \frac{1}{\nu} < h$. Furthermore, we assume that $Tr(R_{i-1}^T R_i) \neq -1$ for $i = 1, \dots, n$.
Then we iterate for $i = 1$ to n:

- For $t \in \{t_{i-1}, t_{i-1} + \delta, t_{i-1} + 2\delta, \dots, t_i\}$, the interpolated trajectory can be computed by:

$$R(t) = R_{i-1} \exp(\tau \cdot \log(R_{i-1}^T R_i)), \qquad \mathbf{d}(t) = (\mathbf{d}_i - \mathbf{d}_{i-1}) \cdot \tau + \mathbf{d}_{i-1}, \qquad \tau = \frac{t - t_{i-1}}{t_i - t_{i-1}}.$$

- We obtain the exponential coordinates $\mathbf{r}(t)$ of $R(t)$ by:

$$[\mathbf{r}(t)] = \log(R(t)).$$

Then we have to smooth all the components of $\mathbf{r}$ and $\mathbf{d}$ separately, obtaining $\tilde{\mathbf{r}}_\Delta$ and $\tilde{\mathbf{d}}_\Delta$ using Eq. (3.68). We finally obtain the smooth trajectory on $SO(3)$ by computing:

$$\widetilde{R}_\Delta(t) = \exp(\tilde{\mathbf{r}}_\Delta(t)),$$

where $t \in \{t_0, t_0 + \delta, \dots, t_1 - \delta, \dots, t_{n-1}, t_{n-1} + \delta, \dots, t_n\}$.

We have compared this algorithm with the algorithm obtained in the previous section (see section B of the Appendix). We have shown that if we interpolate a few number of points, the trajectory obtained using the algorithm proposed in section 3.6.1 is smoother than the corresponding trajectory for the algorithm derived in this section. However, as we have previously stated, the algorithm of section 3.6.1 is not valid if we want to increase the number of points and therefore we should use the algorithm described in this section.

3.7 Trajectory interpolation applied to the inertial measurement simulation

After having developed an optimal algorithm to increase the frequency of the trajectory, we are going to focus on its application for the simulation of inertial measurements.

As we have stated in the introduction of the Chapter 1, an Inertial Measurement Unit (IMU) is an electronic device consisting of a set of accelerometers and a set of angular rate sensors and therefore it measures the linear acceleration and the angular velocity of a body.
Now we are going to obtain expressions for the IMU's measurements in terms of the rigid body dynamics. First we recall that the acceleration of a rigid body equals the acceleration measured by the accelerometers of the IMU excepting for the Coriolis acceleration and the gravity [1]. On the other hand, to find the angular rate sensor measurements we use the relation $R^T\dot{R} = [\omega_b]$ previously derived.

Thus, the usual mathematical model of an inertial navigation, parametrized in the ECEF coordinate system (e) is [1]:

$$\begin{aligned} \dot{\mathbf{x}}^e &= \mathbf{v}^e, \\ \dot{\mathbf{v}}^e &= R_b^e \mathbf{f}^b - 2\Omega_{ie}^e \mathbf{v}^e + \mathbf{g}^e(\mathbf{x}^e), \\ \dot{R}_b^e &= R_b^e(\Omega_{ei}^b + \Omega_{ib}^b). \end{aligned} \tag{3.69}$$

We have to carefully describe all the terms in (3.69). First, $\mathbf{x}^e$, $\mathbf{v}^e$ and R_b^e are the position, velocity and attitude of the rigid body in motion, respectively.
By $\mathbf{f}^b$ we denote the linear accelerations measured by the IMU's accelerometers and $\Omega_{ib}^b = [\omega_{ib}^b]$ is the skew-symmetric matrix corresponding to the angular velocities of the body b with respect to the inertial frame i, measured in the b frame. Thus, ω_{ib}^b are the measures of the IMU's angular rate sensors. Furthermore, $\Omega_{ie}^e = [\omega_{ie}^e]$ is also a skew-symmetric matrix that represents the Earth rotation rate with respect the i frame observed in the ECEF frame and $\Omega_{ei}^b = [\omega_{ei}^b]$ is minus the Earth rotation rate observed in the b frame. Notice that the term $-2\Omega_{ie}^e\mathbf{v}^e = -2\omega_{ie}^e \times \mathbf{v}^e$ is the Coriolis acceleration.
Finally, $\mathbf{g}^e(\mathbf{x}^e)$ is the gravity vector at $\mathbf{x}^e$ provided by a model of the Earth gravity field. The obtainment and characteristics of this gravity field are beyond of the scope of this work and we are going to assume that it is provided by GeoNumerics software.

Now we can express the matrix rotation R_b^e in the exponential coordinates since $R_b^e = \exp([\mathbf{r}_b^e])$. Notice that we have added the corresponding frames in the notation. Then we get:

$$(R_b^e)^T \dot{R}_b^e = W(\mathbf{r}_b^e)\dot{\mathbf{r}}_b^e,$$

where the matrix W has been defined in Eq. (3.59). Thus, the second and third equations of (3.69) can be written as:

$$\begin{aligned} \dot{\mathbf{v}}^e &= \exp([\mathbf{r}_b^e])\mathbf{f}^b - 2\Omega_{ie}^e \mathbf{v}^e + \mathbf{g}^e(\mathbf{x}^e), \\ W(\mathbf{r}_b^e)\dot{\mathbf{r}}_b^e &= \omega_{ei}^b + \omega_{ib}^b. \end{aligned}$$

In addition, $\omega_{ei}^b = R_e^b \omega_{ei}^e$ and therefore we obtain:

$$\begin{aligned} \mathbf{f}^b &= \exp(-[\mathbf{r}_b^e])\left(\dot{\mathbf{v}}^e + 2\Omega_{ie}^e\mathbf{v}^e - \mathbf{g}^e(\mathbf{x}^e)\right), \\ \omega_{ib}^b &= W(\mathbf{r}_b^e)\dot{\mathbf{r}}_b^e - \exp([\mathbf{r}_b^e])\omega_{ei}^e = W(\mathbf{r}_b^e)\dot{\mathbf{r}}_b^e + \exp([\mathbf{r}_b^e])\omega_{ie}^e. \end{aligned} \tag{3.70}$$

To adapt the algorithm developed in the previous section 3.6.2 in order to obtain all the measures needed to solve the above system (3.70) we have to take into account some considerations.

First of all, the usual input trajectory consists of a set of position points expressed in the e reference frame ($\mathbf{x}^e$) and a set of attitudes parametrized using the heading, pitch and roll angles (ψ, θ, γ). These attitudes correspond to the rotation $R_b^l = \exp([\mathbf{r}_b^l])$. On the other hand, we have to obtain $\mathbf{r}_b^e$ to solve the system (3.70).
Thus, we have to add the following previous step to the algorithm of section 3.6.2: for all the input $(\psi, \theta, \gamma)_i$ we compute the corresponding rotation $(R_i)_b^l$ using Eq. (1.2) and obtain $(R_i)_b^e = (R_i)_l^e (R_i)_b^l$ using Eq. (1.4) and (1.5). To obtain the geodetic coordinates (λ, ϕ, H) from the ECEF coordinates $\mathbf{x}^e$ we use GeoNumerics software. After this computations we obtain a set of rotation matrices $\{(R_0)_b^e, \dots, (R_n)_b^e\}$ and therefore can use the algorithm described in section 3.6.2. Then, from the described two-point interpolation followed by the regularization we get the smoothed $\tilde{\mathbf{x}}^e$ and $\tilde{\mathbf{r}}_b^e$.

To solve Eq. (3.70) we also have to differentiate $\mathbf{r}_b^e$ and differentiate twice $\mathbf{x}^e$. In order to do this we recall some results of the subject *Mètodes Numèrics I* about numerical differentiation.

Proposition 3.29. *The numerical derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at a point $x \in \mathbb{R}$ can be first approximated by a forward differencing*

$$f'(x) \simeq \frac{f(x+h) - (x)}{h}, \tag{3.71}$$

with an error corresponding to $-\frac{h}{2} f''(\xi)$, where $\xi \in (x, x+h)$; or similarly by a backward differencing

$$f'(x) \simeq \frac{f(x) - f(x-h)}{h}, \tag{3.72}$$

with an error corresponding to $-\frac{h}{2} f''(\xi)$, where $\xi \in (x-h, x)$. A more precisely approximation is the centred differencing

$$f'(x) \simeq \frac{f(x+h) - f(x-h)}{2h}. \tag{3.73}$$

The associated error is $-\frac{h^2}{6} f^{(3)}(\xi)$, with $\xi \in (x-h, x+h)$. Finally we also can use the following method:

$$f'(x) \simeq \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h}, \tag{3.74}$$

and the error is $-\frac{h^4}{30} f^{(5)}(\xi)$, where $\xi \in (x-2h, x+2h)$.
Clearly, we require $f \in C^j(\mathbb{R})$ with $j \geq 5$.

Thus, the needed derivatives are computed using Proposition 3.29 and the following reasoning. Let $\{t_0, \ldots, t_m\}$ be the interpolated times and ν be the interpolation frequency such that $\delta = \frac{1}{\nu}$.

- For $t = t_0$ we are going to use the forward differencing (Eq. (3.71)).
- For $t = t_m$ we use the backward differencing (Eq. (3.72)).
- For $t = t_0 + \delta$ and $t = t_m - \delta$ we are going to use the centred differencing (Eq. 3.73).
- Finally, for $t \in \{t_0 + 2\delta, \ldots, t_m - 2\delta\}$ we use the method of Eq. 3.74.

In order to validate both the algorithm of section 3.6.2 and the IMU measurements, we could integrate the IMU measurements using the system (3.69). Thus, providing initial conditions we could obtain the tPVA trajectory of a rigid body (position, velocity and attitude). To integrate the IMU measurements we have used a GeoNumerics software called NEXA (New Extensible Generic State-Space Approach).
Indeed, this integration leads to numerical errors and the computed trajectory eventually diverges. To be more specific, let the position of the rigid body be expressed in geodetic coordinates (λ, ϕ, H). Then, small numerical errors in the height component of the position cause errors in the gravity vector $\mathbf{g}^e(\mathbf{x}^e)$. In turn, a wrongly estimation of the gravity increases the height errors and this loop eventually leads to an exponential divergence of the height. The latitude and the longitude are also affected by numerical errors, but their divergence is slower and is caused by the height divergence.
On the other hand, from the system 3.69, we compute the attitude by obtaining $R_b^e = \exp([\mathbf{r}_b^e])$ and the transformation $R_b^l = R_b^e R_e^l$. Since R_e^l depends on λ and ϕ (see Eq. (1.5)), the divergence of these components also cause an error on the attitude.
In a real navigation mission, these divergences are avoided by computing a least square adjustment to improve the trajectory with the precise GNSS measurements.
Nevertheless, since we want to validate our algorithms we are not going to use the GNSS. In the section C of the Appendix we show the obtained results for a real trajectory of 4 hours. We expect the trajectory computed using NEXA to be similar to the smoothed trajectory obtained with the algorithm of section 3.6.2 for a certain period of time, until the NEXA trajectory diverges.
The results are consistent with this reasoning and therefore we can conclude that the proposed proceeding to interpolate a trajectory in order to obtain inertial measurements is valid in addition to being quite innovative (see Appendix for a detailed analysis of the results).

Conclusions

In this work we have studied the problem of the trajectory interpolation i.e., the interpolation of the position and the attitude (orientation) of a rigid body.
Since the interpolation of the position has been extensively studied, the key consideration of the work was to decide how to interpolate the attitude. Thus, we discarded the traditional attitude interpolation techniques using parametrizations such Euler angles or quaternions, in order to derive a more physically accurate interpolation procedure.
In this regard, we have followed the work of F.C. Park, B. Ravani, M. Zefran et. al., who had studied the trajectory interpolation in $SE(3)$ using the Lie theory [9, 11].

Thus, first of all we have introduced some fundamental concepts of Lie theory such as Lie groups, the associated Lie algebra and the exponential map and also some basic concepts of Riemannian geometry. Then, providing $SE(3)$ with simple metrics, we have found the conditions that have to satisfy the minimum acceleration trajectories and the geodesics between two points in $SE(3)$.
We have used two different techniques to adapt this algorithms to a multiple points interpolation. On the one hand, we have simplified the minimum acceleration equations following the Park and Ravani reasoning [9] obtaining an iterative algorithm imposing the continuity of angular velocity and angular acceleration. They developed this algorithm in the context of robotics and therefore they used it to interpolate a few number of points. Indeed, we have checked that the algorithm works for an interpolation of four points. However, we have shown that if we use this algorithm to interpolate more points it generates a trajectory with no physical sense. Consequently, we have discarded this procedure.
Then, we proposed an alternative proceeding. To interpolate a trajectory we have used a two-point interpolation between the ith and the $(i+1)$th points to obtain a first interpolated trajectory and we have smoothed it using a convolution kernel. Notice that the smoothed trajectory does not interpolate the initial trajectory, but, in fact, the interpolation is not a requirement for our purposes since we just want to obtain a similar trajectory with a greater frequency. We have chosen the geodesics for our two-point interpolation, due to its simplicity and lower computational cost.

We have implemented this algorithm to increase the frequency of a real trajectory. Furthermore, we have also described the algorithms needed to obtain the IMU measurements from the smoothed trajectory and we have implemented them. Finally, using a GeoNumerics software we have computed the trajectory corresponding to the IMU simulated measurements and we have shown that this is close to the smoothed trajectory for a long period of time, as we expected.

To sum up, we have adapted the studies of the robotic community in order to obtain a trajectory interpolation algorithm (actually, an algorithm to increase the frequency) and we have tested it obtaining successfully results.
Further studies could explore whether or not the implementation of other two-point interpolation algorithms instead of the geodesics could improve the obtained results without dramatically increasing the computational costs.

Appendix: interpolation results

In this Appendix we are going to show a set of graphics in order to compare and validate the interpolation algorithms described in Chapter 3. As we have previously stated, the input of these algorithms is a tPA (or tA) trajectory with the attitude parametrized using the heading-pitch-roll parametrization. The first step is to obtain all the corresponding rotation matrices in order to implement the interpolation algorithms. Once the attitude has been interpolated, it is again parametrized using the heading-pitch-roll angles. Notice that these parametrizations have been used to provide graphics with an easy physical interpretation and they do not interfere in the interpolation.

A Two-point interpolation comparison

In this section we are going to compare the algorithms described in section 3.5 to interpolate two points in $SO(3)$: the geodesics and the minimal acceleration trajectories under the assumptions of Lemma 3.22. Table 1 shows the points that have been interpolated. In addition, the frequency of interpolation is 100 Hz.

Figure 2 shows the interpolation using the geodesic equations. Although we can observe a behaviour similar to the linear interpolation (as we expected from Proposition 3.19), we also show Figure 3 to illustrate that the geodesics indeed differ from the linear interpolation.

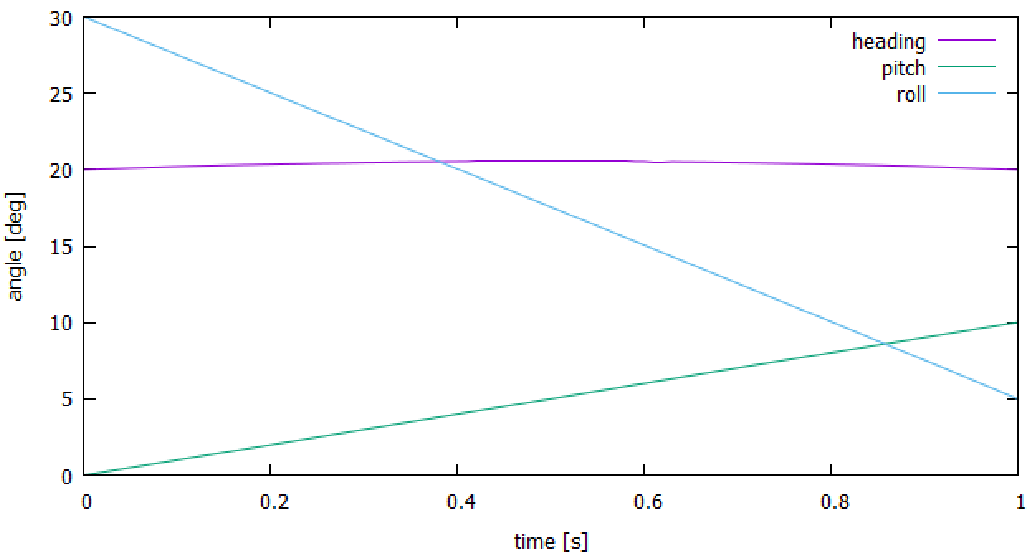

Figure 2: Interpolation of the trajectory corresponding to Table 1 using a geodesic. The graphic shows the heading, pitch and roll angles.

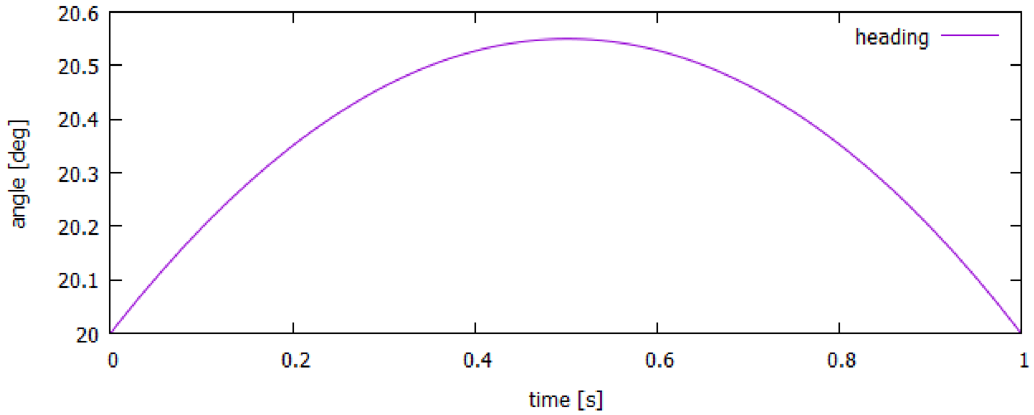

Figure 3: Heading interpolation of the trajectory corresponding to Table 1.

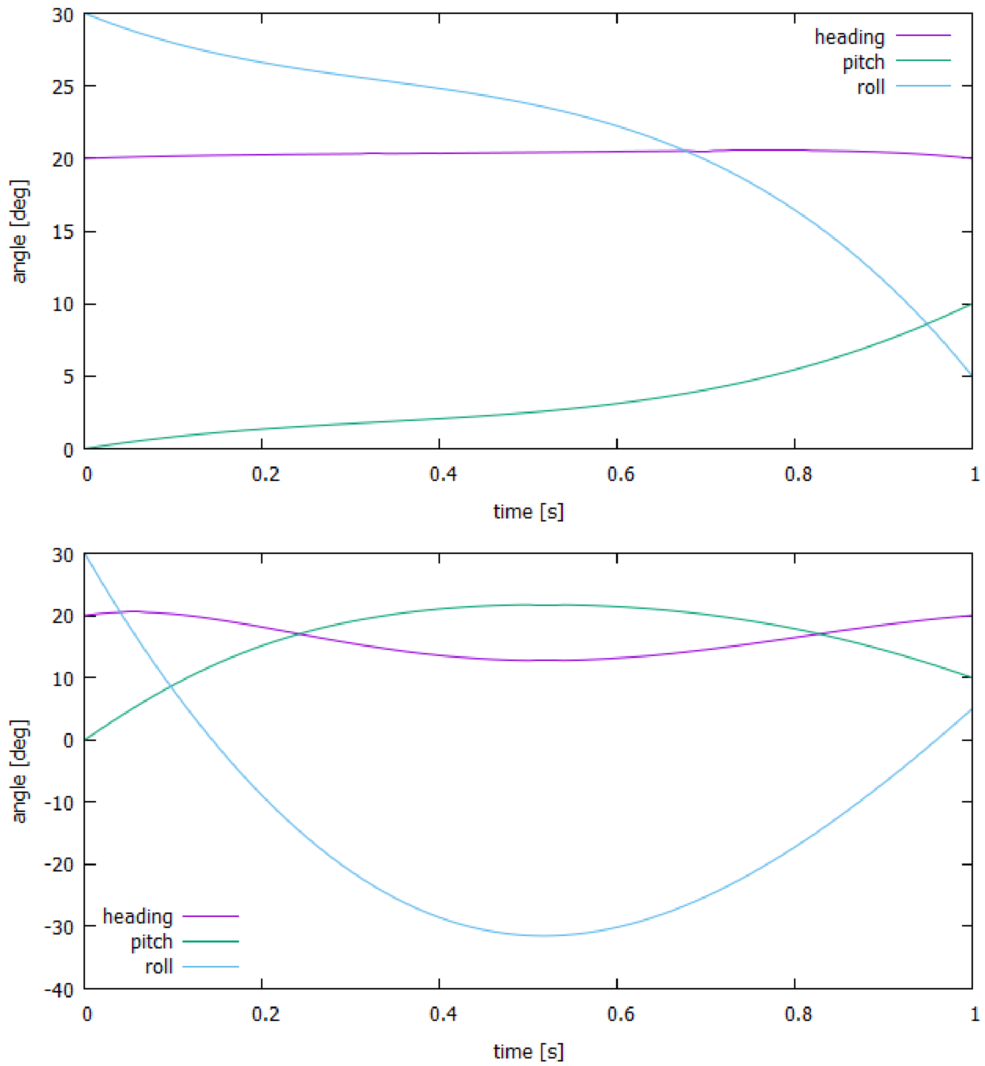

Figure 4: Interpolation of the trajectory corresponding to Table 1 using the minimum acceleration trajectory of Lemma 3.22 and using $(\sigma, \mu) = (1, 3)$ (top image) and $(\sigma, \mu) = (10, -5)$ (bottom image). The graphics show the heading, pitch and roll.

On the other hand, Figure 4 shows the minimum acceleration trajectories corresponding to Lemma 3.22 with two different values for the pair (σ, μ). We observe that if $|\sigma|, |\mu|$ are relatively small, the minimum acceleration trajectory do not significantly differ from the geodesics. On the other hand, if we increase $|\sigma|$ and $|\mu|$, then the trajectory varies widely, even becoming a trajectory with no physical sense (accordingly with the roll variation). Thus, if we want to use the minimum acceleration trajectories to interpolate two points in $SO(3)$, $|\sigma|, |\mu|$ should be not too large.

time (s)	ψ (deg)	θ (deg)	γ (deg)
1	20	0	30
2	20	10	5

Table 1: Two points in $SO(3)$ parametrized using the heading (ψ), pitch (θ) and roll (γ) angles.

B Comparison of the procedures of sections 3.6.1 and 3.6.2

In this section we are going to compare the two procedures described in section 3.6 to interpolate a trajectory through multiple points. First of all we provide an example of the implementation of the multiple point interpolation algorithm based on the minimum acceleration trajectory (see section 3.6.1). The attitude trajectory that has been interpolated is showed in the Table 2 (left). Moreover, the frequency of interpolation is 100Hz.

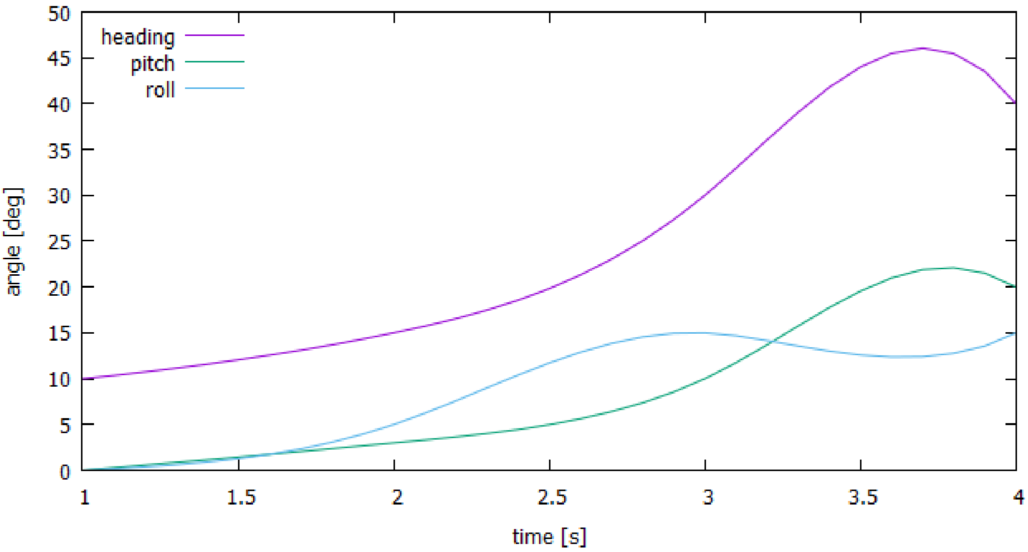

Figure 5: Interpolation of the trajectory corresponding to Table 2 (left) using the algorithm described in section 3.6.1. The graphic shows the heading, pitch and roll angles.

We observe that the interpolated trajectory is smooth and it does not present any abrupt change.
Nevertheless, if we extend this trajectory to a trajectory consisting in ten points (see the right table of Table 2) we obtain a trajectory with no physical sense and, above all, the heading presents an erratic behaviour. The frequency of interpolation is again 100 Hz.

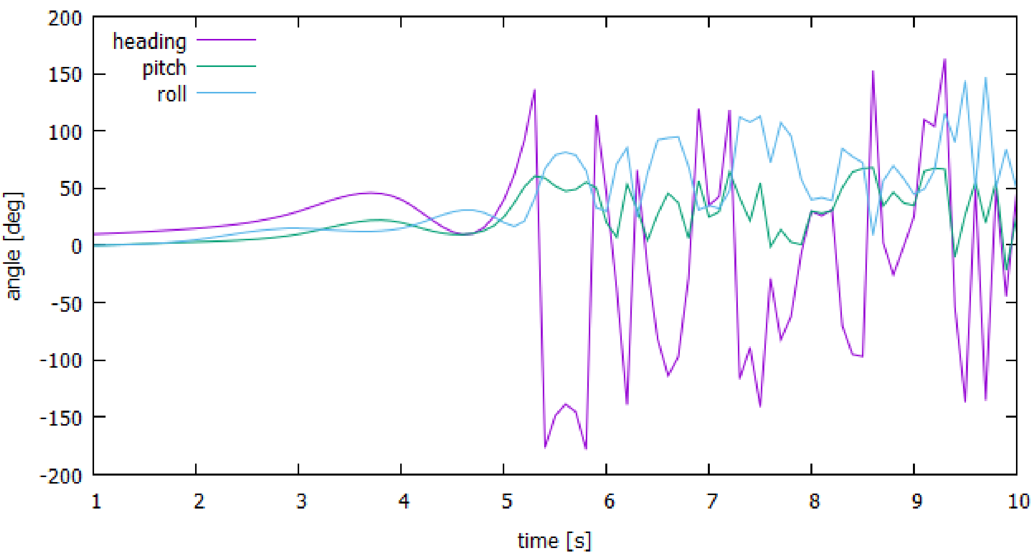

Figure 6: Interpolation of the trajectory corresponding to Table 2 (right) using the algorithm described in section 3.6.1. The graphic shows the heading, pitch and roll angles.

time (s)	ψ (deg)	θ (deg)	γ (deg)
1	10	0	0
2	15	3	5
3	30	10	15
4	40	20	15

time (s)	ψ (deg)	θ (deg)	γ (deg)
1	10	0	0
2	15	3	5
3	30	10	15
4	40	20	15
5	50	25	20
6	40	20	30
7	35	25	35
8	30	30	40
9	25	35	45
10	45	25	50

Table 2: The left table shows a four point attitude trajectory and the right table shows an attitude trajectory of ten points. Both trajectories are parametrized using the heading (ψ) - pitch (θ) - roll (γ) parametrization.

The above results show that the proposed multiple point interpolation (section 3.6.1) is only valid for a few number of points and it does not generates a satisfactory interpolated trajectory when we increase the number of points.

Subsequently, we have interpolated the same trajectories using the interpolation procedure described in section 3.6.2, that is, to interpolate the trajectory using a two-point interpolation and later smooth the trajectory using a kernel convolution. Figures 7 and 8 show the results. The interpolation frequency is 100 Hz and $\Delta = 10$ for both trajectories.

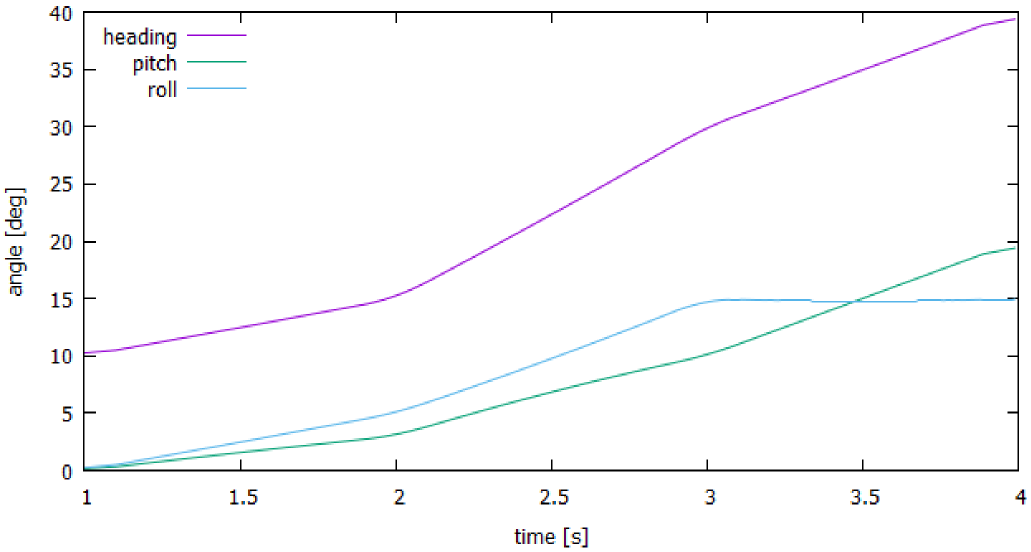

Figure 7: Interpolation of the trajectory corresponding to Table 2 (left) using the algorithm described in section 3.6.2. The graphic shows the heading, pitch and roll angles.

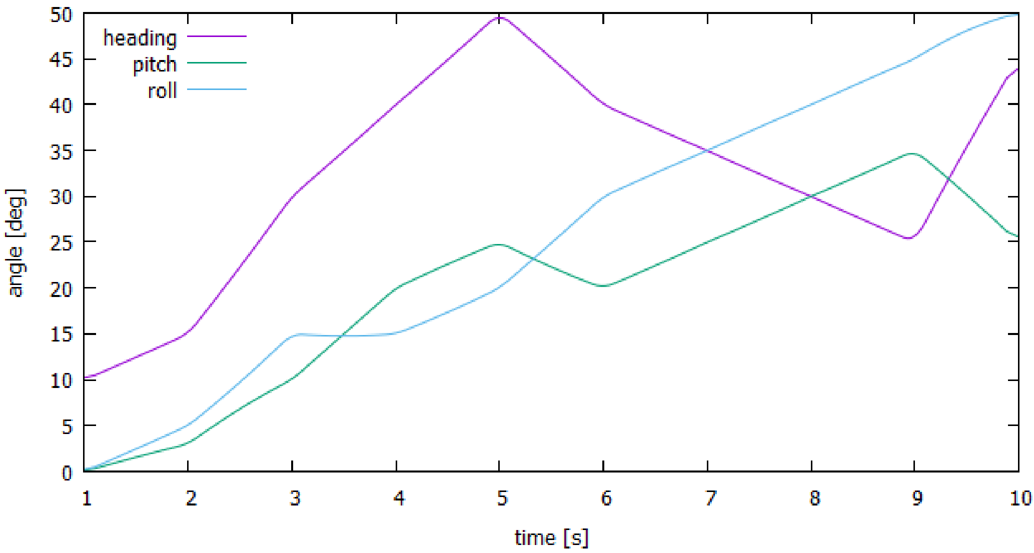

Figure 8: Interpolation of the trajectory corresponding to Table 2 (right) using the algorithm described in section 3.6.2. The graphic shows the heading, pitch and roll angles.

We can observe that the behaviour of the interpolated trajectory is more linear than the corresponding to the previous algorithm, but it is also smooth and in addition it can be used to interpolate an arbitrary large number of points.

C Validation of the algorithm of section 3.6.2

Finally, in this section we validate the algorithm described in section 3.6.2 interpolating a real trajectory of approximately 4 hours provided at a frequency of 1 Hz. The output frequency is 50 Hz and we have used $\Delta = 50$ to regularize the interpolated function. Once we have obtained the smooth trajectory, we have computed the IMU measurements using Eq. (3.70) and also the corresponding trajectory using NEXA.
As we have previously stated, NEXA integrates the IMU measurements using the system (3.69). This integration generates numerical errors and the trajectory eventually diverges.
Figure 9 shows the comparison of the heading angle for the NEXA trajectory and for the regularized trajectory. We have chosen the heading since this angle varies widely along the trajectory while the pitch and roll remain quasi constant. Thus, the heading angle is more suitable to be analysed. We can qualitatively observe that the trajectories are close until $t \simeq 46600$ (corresponding to almost 3 hours after the beginning of the trajectory).
For the sake of precision, we provide the differences in the heading, pitch and roll angles for the two trajectories (for $t \leq 46700$) and also the differences in the position coordinates, see Figures 10 and 11. We have parametrized the position using the geodetic coordinates in order to show the exponential divergence of the height in front of the slower divergence of the latitude and the longitude. Thus, the top image of Figure 11 shows the differences in the latitude and the longitude for $t \leq 46700$ and the bottom image shows the differences in the height for $t \leq 43000$.
We observe that the attitude, the longitude and latitude differences are quite small and that the differences in the height remain small over a long period of time. These results concord with the expected results and in fact, we would expect even an earlier divergence of the position coordinates. Thus, both the algorithm of section 3.6.2 and the algorithm designed to obtain the IMU measurements are very successful.

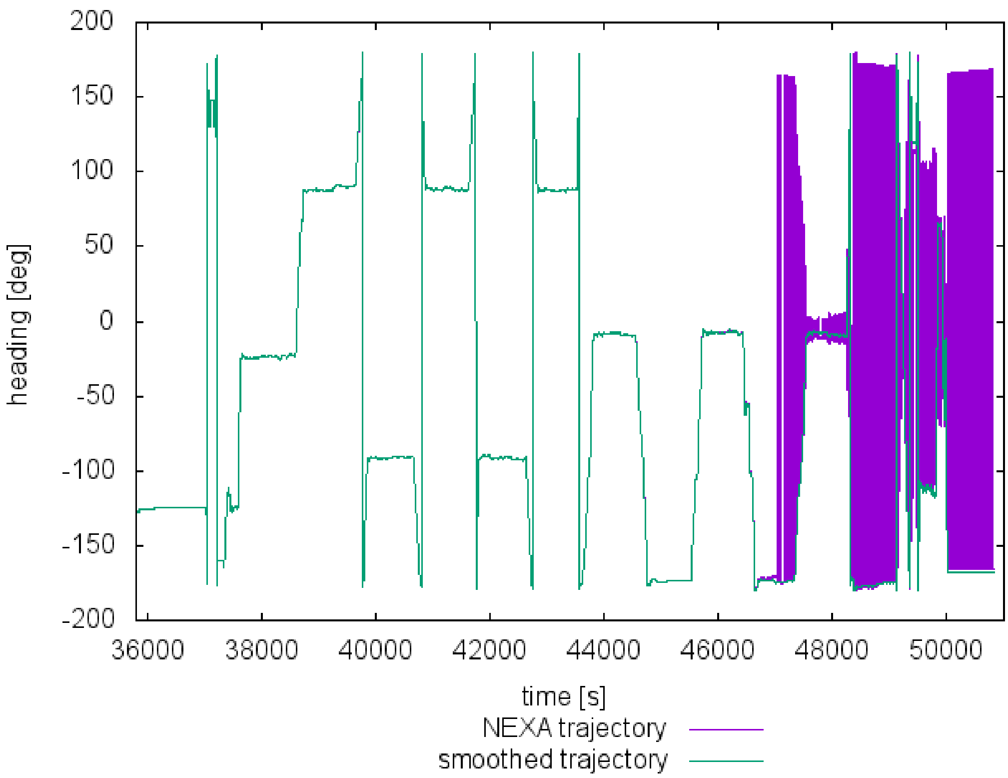

Figure 9: Comparison of the heading angle of the trajectory computed with NEXA and the smoothed trajectory obtained using the algorithm described in section 3.6.2.

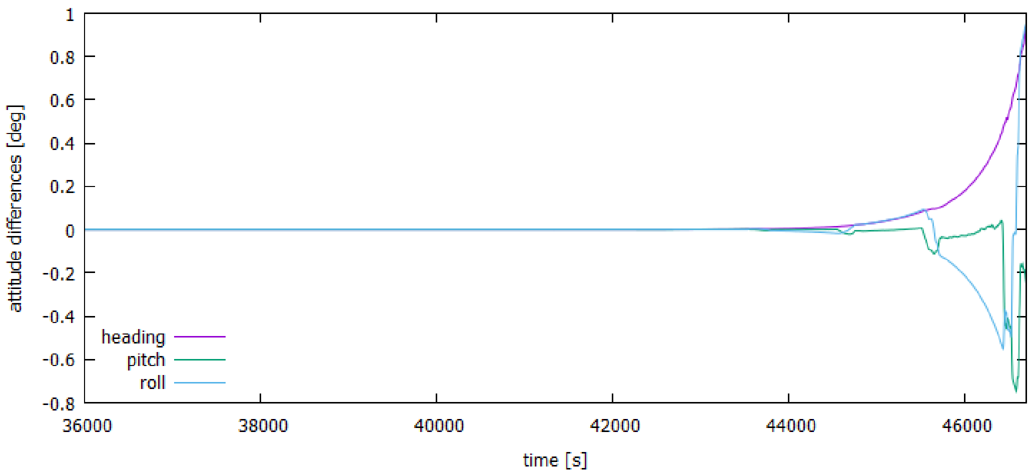

Figure 10: Differences in the heading, pitch and roll angles of the trajectory computed with NEXA and the smoothed trajectory obtained using the algorithm described in section 3.6.2.

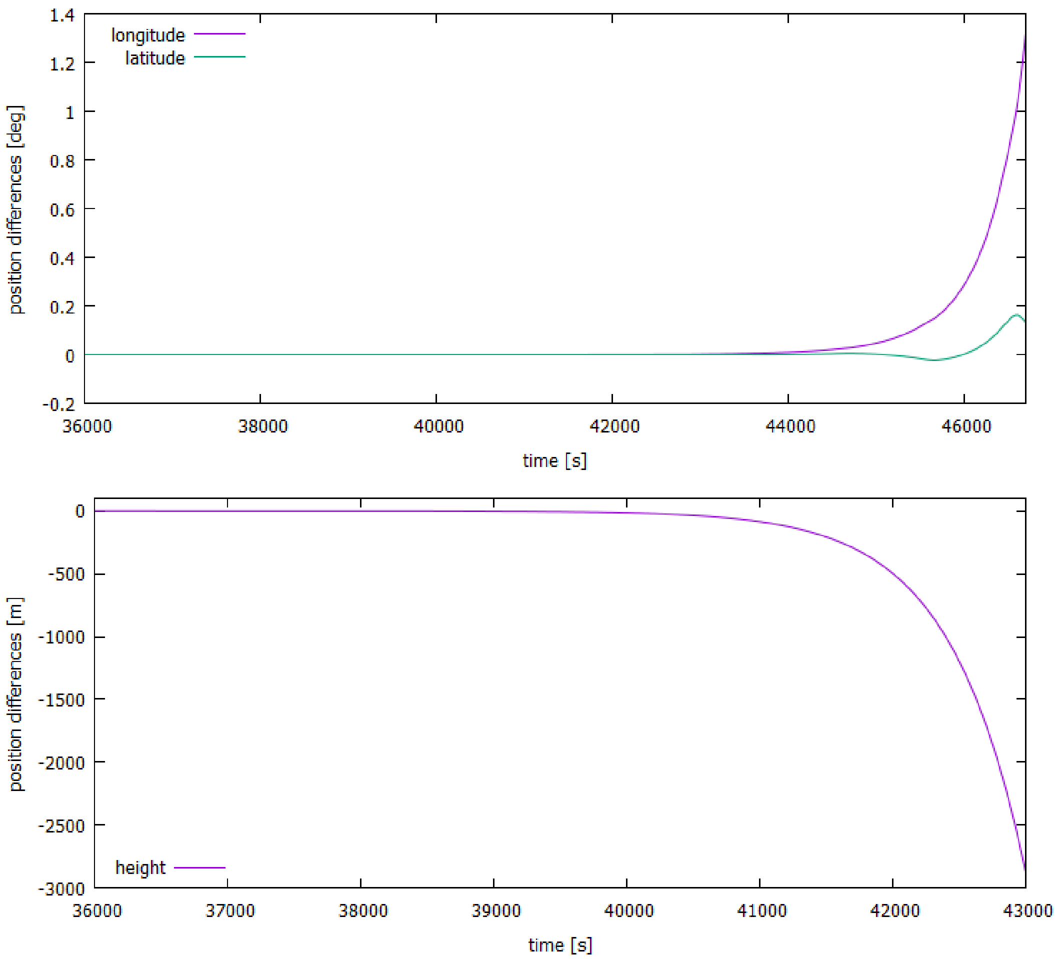

Figure 11: Differences in the position of the trajectory computed with NEXA and the smoothed trajectory obtained using the algorithm described in section 3.6.2. The top image shows the differences in the longitude and latitude and the bottom image shows the differences in the height.

Bibliography

[1] W. Torge and J. Müller, *Geodesy*. Walter de Gruyter (2012).

[2] M.P. do Carmo, *Riemannian geometry (Translated by Francis Flaherty)*, Mathematics: Theory & Applications, Birkhauser Boston Inc. (1992).

[3] G. Bellamy, *Lie groups, Lie algebras, and their representations*, University of Glasgow (2010).

[4] J. Lee, *Introduction to smooth manifolds*, Graduate Text in Mathematics **218**, Springer New York (2012).

[5] D. Bump, *Lie groups*, Graduate Texts in Mathematics **225**, Springer-Verlag New York (2004).

[6] C. Chevalley, *Theory of Lie groups*, Princeton University Press (1946).

[7] V. I. Arnold, *Mathematical methods of classical mechanics*, Graduate Text in Mathematics **60**, Springer-Verlag New York (1989).

[8] J. Stoer and R. Bulirsch, *Introduction on Numerical Analysis*, Springer-Verlag New York (1993).

[9] F. C. Park and B. Ravani *Smooth invariant interpolation of rotations*, ACM Transactions on Graphics (TOG) **16.3** (1997) 277-295.

[10] J. M. Selig, *Geometric fundamentals of robotics*, Springer-Verlag New York (2005).

[11] M. Zefran, V. Kumar, and C. B. Croke., *On the generation of smooth three-dimensional rigid body motions*, IEEE Transactions on Robotics and Automation **14.4** (1998) 576-589.

[12] L. Hormander, *The Analysis of Partial Differential Operators*, Springer-Verlag Berlin Heidelberg (1983).