The SO(3) and SE(3) Lie Algebras of Rigid Body Rotations and Motions and their Application to Discrete Integration, Gradient Descent Optimization, and State Estimation

Eduardo Gallo*†‡

June 2022

Abstract

Common mathematical techniques such as discrete integration, gradient descent optimization, and state estimation (exemplified by the Runge-Kutta method, Gauss-Newton minimization, and extended Kalman filter or EKF, respectively), rely on linear algebra and hence are only applicable to state vectors belonging to Euclidean spaces when implemented as described in the literature. This article describes how to modify these methods so they can be applied to non Euclidean state vectors, such as those containing rotations and full motions of rigid bodies.

To do so, this article provides an in-depth review of the $\mathbb{SO}(3)$ and $\mathbb{SE}(3)$ Lie groups, known as the special orthogonal and special Euclidean groups of $\mathbb{R}^3$, which represent the rigid body rotations and motions, placing special emphasis on the different possible representations, their tangent spaces, the analysis of perturbations, and in particular the definitions of the jacobians required to employ the previously mentioned calculus methods.

Keywords: Lie algebra, SO(3), SE(3), manifold, tangent space, state estimation, EKF, discrete integration, Runge-Kutta, gradient descent optimization, minimization, Gauss-Newton

Acronyms

CDF	Cumulative Distribution Functions	PDF	Probability Density Function
ECEF	Earth Centered Earth Fixed	PMF	Probability Mass Function
EKF	Extended Kalman Filter	PSD	Power Spectral Density
NED	North East Down	ScLERP	Screw linear interpolation
ODE	Ordinary Differential Equation	SLERP	Spherical linear interpolation

1 Introduction and Outline

The widespread implementations of common calculus techniques such as discrete integration (exemplified by the Runge-Kutta method), gradient descent optimization (Gauss-Newton or Levenberg-Marquardt), and state

*The author holds a MSc in Aerospace Engineering by the Polytechnic University of Madrid and has twenty-three years of experience working in aircraft trajectory prediction, modeling, and flight simulation. He is currently a Senior Trajectory Prediction and Aircraft Performance Engineer at Boeing Research & Technology Europe (BR&TE), although he is publishing this article in his individual capacity and time as part of his PhD thesis titled "Autonomous Unmanned Air Vehicle GNSS-Denied Navigation", advised by Dr. Antonio Barrientos within the Centre for Automation and Robotics of the Polytechnic University of Madrid.

†Contact: edugallo@yahoo.com, `https://orcid.org/0000-0002-7397-0425`

‡Affiliation: Centro de Automática y Robótica, Universidad Politécnica de Madrid - Consejo Superior de Investigaciones Científicas. Centre for Automation and Robotics, Polytechnic University of Madrid.

estimation (extended Kalman filter or `EKF`), are designed to work on state vectors belonging to linear Euclidean spaces [1], and hence rely of linear algebra. There exist two possible approaches for those cases in which the state vector contains non Euclidean components, such as rigid body motions:

- The solution most commonly observed in the literature is to incorporate each component of the pose (position plus attitude) as an unconstrained real number into the state vector. The above techniques can then be employed without difficulties, but the resulting state vector does not comply with the constraints imposed by having some of its members being components of a rigid body pose, as these constraints have not been taken into account when integrating, optimizing, or estimating. The solutions hence need to be projected back into the space of valid rigid body poses, but this does not hide the fact that the whole process has been performed without respecting the motion constraints, which often has negative consequences for the accuracy, stability, and consistency of the solution [2].
- The approach taken in some recent robotics literature, in particular in the field of motion estimation for navigation, consists on reformulating the above calculus techniques (integration, optimization, filtering) taking into account that some members of the state vector represent rigid body poses and hence can not be treated as Euclidean. By modeling these states properly, the quality of the solution can be improved. The use of Lie theory, with its manifolds and tangent spaces, enables the construction of a rigorous calculus corpus to handle uncertainties, derivatives, and integrals of non euclidean elements with precision and ease [2].

This article begins with a review of the concepts of random variables, stochastic processes, and white noise in section 2, which are necessary for the section 3 descriptions of the Runge-Kutta discrete integration, Gauss-Newton minimization, and `EKF` when applied to Euclidean spaces. Section 4 introduces the concepts of Lie groups and their tangent spaces or Lie algebras, and adapts the three calculus techniques so they can be applied when the state vector contains non Euclidean Lie group components. Sections 5 and 6 particularize the generic concepts of section 4 to the specific cases of both rigid body rotations as well as full rigid body motions.

2 Random Variables, Stochastic Processes, and White Noise

This section provides an introduction to the random variables and processes required to model those physical systems that can not be represented by deterministic models due to their inherent randomness, which results in the same set of parameter values and initial conditions leading to different outputs.

2.1 Random Variables

Consider a random experiment (one in which the outcome is uncertain) with a sample space Ω (collection of possible elementary outcomes of the experiment), and let ω be a sample point belonging to Ω. A *random variable* $\mathrm{X}(\omega)$ (generally just X) is a single valued real function that assigns a real number, called the value of $\mathrm{X}(\omega)$, to each sample point $\omega \in \Omega$ [3, 4]. A random variable hence represents a map between the sample and real spaces $\{\mathrm{X} : \Omega \rightarrow \mathbb{R} \mid \omega \in \Omega \rightarrow \mathrm{X}(\omega) \in \mathbb{R}\}$. The *realization* of a random variable is the real variable obtained after a given experiment.

A random variable X is completely described by its *cumulative distribution function* (`CDF`) $\mathrm{F_X}$, which represents the probability that the value of X is less or equal than the function input [5]:

$$\mathrm{F_X}(\mathrm{x}) = \mathrm{P}[\mathrm{X} \leq \mathrm{x}] \qquad -\infty < \mathrm{x} < \infty \tag{1}$$

Random variables can also be described by the `CDF` derivative, known as the *probability mass function* (`PMF`) $\mathrm{p_X}$ in case of discrete random variables (those that can take at most a countable number of possible values) or as *probability density function* (`PDF`) $\mathrm{f_X}$ for continuous ones (those that can take an uncountable number of possible values):

$$F_X(x) = \sum_{x_k \le x} p_X(x_k) \tag{2}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy \tag{3}$$

The *expected value* $E[X]$ or *mean* μ_X of a random variable X is a function defined as its average value over a large number of experiments, and represents its central or typical value:

$$E[X] = \mu_X = \begin{cases} \displaystyle\sum_k x_k\, p_X(x_k) & \text{when X is discrete} \\ \displaystyle\int_{-\infty}^{\infty} y\, f_X(y)\, dy & \text{when X is continuous} \end{cases} \tag{4}$$

If a function acts on a random variable, then its output is also a random variable [6], and it is hence possible to compute the expected value of the output random variable[1]. The *variance* σ_X^2, $\mathrm{Var}(X)$, or second central moment of a random variable X, is the expected value of the squared deviation of X from its mean, and measures the spread of its PMF or PDF about its expected value [3], this is, how much the random variable is expected to deviate from its mean. The square root of the variance is called the *standard deviation* σ_X.

$$\mathrm{Var}(X) = \sigma_X^2 = E\left[(X - \mu_X)^2\right] = \begin{cases} \displaystyle\sum_k (x_k - \mu_X)^2\, p_X(x_k) & \text{when X is discrete} \\ \displaystyle\int_{-\infty}^{\infty} (y - \mu_X)^2\, f_X(y)\, dy & \text{when X is continuous} \end{cases} \tag{5}$$

The variance and mean of a random variable are related by the following expression, where $E\left[X^2\right]$ is the second moment of X.

$$\sigma_X^2 = E\left[(X - \mu_X)^2\right] = E\left[X^2\right] - 2\,\mu_X\, E[X] + \mu_X^2 = E\left[X^2\right] - \mu_X^2 = \mu_{X^2} - \mu_X^2 \tag{6}$$

The notation $X \sim (\mu_X,\ \sigma_X^2)$ means that the random variable X has μ_X mean and σ_X^2 variance. A *normal* or *Gaussian* random variable X of parameters μ_X and σ_X^2 [5], represented as $X \sim N(\mu_X,\ \sigma_X^2)$, is one whose PDF responds to:

$$f_X(x) = \frac{1}{\sqrt{2\,\pi\,\sigma_X^2}} \exp\left(-\frac{(x - \mu_X)^2}{2\,\sigma_X^2}\right) \qquad -\infty < x < \infty \tag{7}$$

The expected value and variance of a normal random variable $N(\mu_X,\ \sigma_X^2)$ are μ_X and σ_X^2, respectively. A normal random variable $N(0,\ 1)$ of zero mean and unit variance is called a *standard normal random variable*. It is worth noting that any affine[2] function of a Gaussian random variable results in a Gaussian random variable [6].

The *discrete uniform distribution* assigns the same probability to each of its N possible values [3]. Represented by $X \sim U(a,\ a + N - 1)$, its expected value is $a + (N-1)/2$ and its variance is $(N^2 - 1)/12$. Its PMF responds to:

$$p_X(x_k) = \begin{cases} \dfrac{1}{N} & x_k = a,\ a+1, \ldots,\ a+N-1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Consider now two random variables X and Y defined in the same sample space Ω with expected values μ_X and μ_Y, respectively, and variances σ_X^2 and σ_Y^2. They are called *independent* if their results do not depend on each other:

$$F_{XY}(x,\ y) = P[X \le x,\ Y \le y] = P[X \le x]\ P[Y \le y] = F_X(x)\ F_Y(y) \tag{9}$$

[1]Note that the mean can be considered as the expected value of the $f(X) = X$ function.

[2]In this contest affine means a function of the form $y = a\,x + b$, while linear means $y = a\,x$.

The *central limit theorem* states that the sum of independent random variables tends towards a Gaussian random variable, regardless of the `CDF` of the individual random variables that contribute to the sum [3, 6]. If a given random variable X is realized many times, the *law of large numbers* states that the average of the realizations is close to the random variable expected value μ_X, and tends to it as the numbers of realizations grows [3].

Stochastic simulations, also known as *Monte Carlo simulations*, use randomness to solve complex problems that may have a deterministic nature [7, 8] and that are based on multiple unknown parameters, many of which are difficult to obtain experimentally [9]. They rely on defining the domain of possible inputs, randomly generating inputs from a probability distribution over the domain, performing a deterministic computation based on those inputs, and finally aggregating the results by means of a set of metrics.

The *correlation* R_{XY} and the *covariance* C_{XY} of the random variables X and Y are two measures of the linear correlation between both random variables [3]. They are defined as:

$$R(X, Y) = R_{XY} = E[X \cdot Y] = \mu_{X \cdot Y} \tag{10}$$

$$C(X, Y) = C_{XY} = E[(X - \mu_X) \cdot (Y - \mu_Y)] = E[X \cdot Y] - \mu_X \mu_Y = R_{XY} - \mu_X \mu_Y \tag{11}$$

Two random variables are *uncorrelated* if their covariance is zero ($C_{XY} = 0 \rightarrow R_{XY} = \mu_X \mu_Y$). Independent random variables are always uncorrelated, but not the other way around, as two uncorrelated random variables may not necessarily be independent if there exists a non linear dependence between them. *Orthogonal* random variables are those whose correlation is zero ($R_{XY} = 0$), so they may or may not also be uncorrelated. If they are, at least one of them is zero mean.

The expected value and variance of the sum and product of two random variables are of particular interest. Given two random variables X and Y with expected values $\{\mu_X, \mu_Y\}$ and variances $\{\sigma_X^2, \sigma_Y^2\}$, its sum $X + Y$ and product $X \cdot Y$ are also random variables, as indicated above. The following expressions can be easily obtained by applying the equations above[3] [10]:

$$\mu_{X+Y} = E[X + Y] = \mu_X + \mu_Y \tag{12}$$

$$\sigma_{X+Y}^2 = E\left[(X + Y)^2\right] - \mu_{X+Y}^2 = \sigma_X^2 + 2\,C_{XY} + \sigma_Y^2 \tag{13}$$

$$\mu_{X \cdot Y} = E[X \cdot Y] = R_{XY} = C_{XY} + \mu_X \mu_Y \tag{14}$$

$$\sigma_{X \cdot Y}^2 = E\left[(X \cdot Y)^2\right] - \mu_{X \cdot Y}^2 = C_{X^2 Y^2} - C_{XY}^2 - 2\,C_{XY}\,\mu_X \mu_Y + \sigma_X^2 \sigma_Y^2 + \sigma_X^2 \mu_Y^2 + \mu_X^2 \sigma_Y^2 \tag{15}$$

2.2 Random Vectors

A *random vector* $\mathbf{X} = [X_1, \ldots, X_n]^T$ is a collection of random variables obtained from the same sample space Ω [3, 4]. The random vector joint `CDF`, `PMF`, and `PDF` are defined as follows:

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P[\{X_1 \leq x_1\} \cap \ldots \cap \{X_n \leq x_n\}] = P[X_1 \leq x_1, \ldots, X_n \leq x_n] \tag{16}$$

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \sum_{k_1 \leq x_1} \cdots \sum_{k_n \leq x_n} p_{X_1,\ldots,X_n}(k_1, \ldots, k_n) \tag{17}$$

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,\ldots,X_n}(y_1, \ldots, y_n)\, dy_n \ldots dy_1 \tag{18}$$

When the components $X_1, \ldots, X_n$ of the random vector $\mathbf{X}$ are independent from each other, its joint `CDF`, `PMF`, and `PDF` are just the product of the respective functions of each of the random vector components [5]. The expected value $E[\mathbf{X}]$ or mean $\boldsymbol{\mu}_X$ and the variance $\boldsymbol{\sigma}_X^2$ of a random vector $\mathbf{X}$ are defined as the vectors of those of its components:

$$E[\mathbf{X}] = \boldsymbol{\mu}_X = [\mu_{X1}, \ldots, \mu_{Xn}]^T \tag{19}$$

$$\mathrm{Var}(\mathbf{X}) = \boldsymbol{\sigma}_X^2 = \left[\sigma_{X1}^2, \ldots, \sigma_{Xn}^2\right]^T \tag{20}$$

[3]These expressions can be further simplified when X and Y are uncorrelated.

Given two random vectors $\mathbf{X} \in \mathbb{R}^{\mathrm{m}}$ and $\mathbf{Y} \in \mathbb{R}^{\mathrm{n}}$, their *correlation matrix* $\mathbf{R}_{\mathrm{XY}}$ is defined so $\mathrm{R_{ij}} = \mathrm{R}(\mathrm{X_i}, \mathrm{Y_j})$, while their *covariance matrix* $\mathbf{C}_{\mathrm{XY}}$ verifies that $\mathrm{C_{ij}} = \mathrm{C}(\mathrm{X_i}, \mathrm{Y_j})$ [5]:

$$\mathbf{R}(\mathbf{X}, \mathbf{Y}) = \mathbf{R}_{\mathrm{XY}} = \mathrm{E}\left[\mathbf{X}\,\mathbf{Y}^{\mathrm{T}}\right] \tag{21}$$

$$\mathbf{C}(\mathbf{X}, \mathbf{Y}) = \mathbf{C}_{\mathrm{XY}} = \mathrm{E}\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathrm{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathrm{Y}})^{\mathrm{T}}\right] = \mathrm{E}\left[\mathbf{X}\,\mathbf{Y}^{\mathrm{T}}\right] - \boldsymbol{\mu}_{\mathrm{X}}\,\boldsymbol{\mu}_{\mathrm{Y}}^{\mathrm{T}} = \mathbf{R}_{\mathrm{XY}} - \boldsymbol{\mu}_{\mathrm{X}}\,\boldsymbol{\mu}_{\mathrm{Y}}^{\mathrm{T}} \tag{22}$$

The *autocorrelation* and *autocovariance* matrices $\mathbf{R}_{\mathrm{XX}}$ and $\mathbf{C}_{\mathrm{XX}}$ of a random vector $\mathbf{X} \in \mathbb{R}^{\mathrm{m}}$ are defined as the correlation and covariance matrices of that vector with itself. Both are square, symmetric $\left(\mathbf{R}_{\mathrm{XX}} = \mathbf{R}_{\mathrm{XX}}^{\mathrm{T}}, \mathbf{C}_{\mathrm{XX}} = \mathbf{C}_{\mathrm{XX}}^{\mathrm{T}}\right)$, positive semidefinite $\left(\mathbf{z}^{\mathrm{T}}\,\mathbf{R}_{\mathrm{XX}}\,\mathbf{z} \geq 0, \mathbf{z}^{\mathrm{T}}\,\mathbf{C}_{\mathrm{XX}}\,\mathbf{z} \geq 0, \forall\,\mathbf{z} \in \mathbb{R}^{\mathrm{m}}\right)$, and their diagonals contain the second moments and variances of each of the random variables $\mathrm{X_i}$ within $\mathbf{X}$:

$$\mathbf{R}(\mathbf{X}) = \mathbf{R}_{\mathrm{XX}} = \mathrm{E}\left[\mathbf{X}\,\mathbf{X}^{\mathrm{T}}\right] \tag{23}$$

$$\mathbf{C}(\mathbf{X}) = \mathbf{C}_{\mathrm{XX}} = \mathrm{E}\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathrm{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathrm{X}})^{\mathrm{T}}\right] = \mathrm{E}\left[\mathbf{X}\,\mathbf{X}^{\mathrm{T}}\right] - \boldsymbol{\mu}_{\mathrm{X}}\,\boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{T}} = \mathbf{R}_{\mathrm{XX}} - \boldsymbol{\mu}_{\mathrm{X}}\,\boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{T}} \tag{24}$$

A normal or Gaussian random vector is that whose components are all normal random variables. As in the case of scalar random variables, an affine transformation of a Gaussian random vector results in a new Gaussian random vector.

2.3 Stochastic Processes

A *random process* or *stochastic process* enlarges the concept of random vector (or random variable when the vector size is one) to include time. Given a sample vector $\boldsymbol{\omega}$ of the sample space Ω and a parameter t belonging to a parameter set $\mathbb{T}$ (generally time), a stochastic process assigns a real vector $\{\mathbf{X} : \mathbb{T}, \Omega \to \mathbb{R}^{\mathrm{m}} \mid \mathrm{t} \in \mathbb{T}, \boldsymbol{\omega} \in \Omega \to \mathbf{X}(\mathrm{t}, \boldsymbol{\omega}) \in \mathbb{R}^{\mathrm{m}}\}$ [3, 4, 11]. If the sample vector $\boldsymbol{\omega}$ is fixed, the random process $\mathbf{X}(\mathrm{t})$ behaves as a function of time; on the other hand, if the time is fixed, the stochastic process $\mathbf{X}(\boldsymbol{\omega})$ defaults to a random vector. A stochastic process is thus a family of random vectors (either discrete or continuous) indexed by a continuous parameter $\mathrm{t} \in \mathbb{T}$. If the parameter is discrete $\mathrm{t} \in \mathbb{Z}^{+}$ [4], then the appropriate name is *stochastic sequence* [5, 6].

Size one stochastic processes $\mathrm{X}(\mathrm{t}, \omega)$, generally represented just by $\mathrm{X}(\mathrm{t})$, are completely described by their `CDF`, while if $\mathbf{X}(\mathrm{t})$ is instead a random vector, it is represented by its joint `CDF`:

$$\mathrm{F_X}(\mathrm{x}; \mathrm{t}) = \mathrm{P}[\mathrm{X}(\mathrm{t}) \leq \mathrm{x}] \tag{25}$$

$$\mathrm{F_X}(\mathrm{x}_1, \ldots, \mathrm{x_n}; \mathrm{t}) = \mathrm{P}[\mathrm{X}_1(\mathrm{t}) \leq \mathrm{x}_1, \ldots, \mathrm{X_n}(\mathrm{t}) \leq \mathrm{x_n}] \tag{26}$$

The joint `PMF` and `PDF` are also defined in similar fashion. The *ensemble average* or mean of a random process becomes a function of time:

$$\boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t}) = \mathrm{E}[\mathbf{X}(\mathrm{t})] \tag{27}$$

Note that the random process $\mathbf{X}(\mathrm{t})$ evaluated at different times comprises different random vectors of the same size. It is then possible to apply the concepts of autocorrelation and autocovariance introduced in section 2.2 to any two of these random vectors, providing quantitative measures of the similarity of the random process at two different times, this is, measuring by how much a signal is similar to its time shifted version [3]. This results in the *autocorrelation* $\mathbf{R}_{\mathrm{XX}}(\mathrm{t}, \mathrm{t} + \tau)$ and the *autocovariance* $\mathbf{C}_{\mathrm{XX}}(\mathrm{t}, \mathrm{t} + \tau)$:

$$\mathbf{R}_{\mathrm{XX}}(\mathrm{t}, \mathrm{t} + \tau) = \mathrm{E}\left[\mathbf{X}(\mathrm{t})\;\mathbf{X}^{\mathrm{T}}(\mathrm{t} + \tau)\right] \tag{28}$$

$$\begin{aligned} \mathbf{C}_{\mathrm{XX}}(\mathrm{t}, \mathrm{t} + \tau) &= \mathrm{E}\left[\big(\mathbf{X}(\mathrm{t}) - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\big)\big(\mathbf{X}(\mathrm{t} + \tau) - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t} + \tau)\big)^{\mathrm{T}}\right] \\ &= \mathrm{E}\left[\mathbf{X}(\mathrm{t})\;\mathbf{X}^{\mathrm{T}}(\mathrm{t} + \tau)\right] - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\;\boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{T}}(\mathrm{t} + \tau) = \mathbf{R}_{\mathrm{XX}}(\mathrm{t}, \mathrm{t} + \tau) - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\;\boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{T}}(\mathrm{t} + \tau) \end{aligned} \tag{29}$$

[4] $\mathbb{Z}^{+}$ represents the set of positive integers.

The autocovariance is zero when the two observations of $\mathbf{X}$ are independent, meaning that there is no coupling between $\mathbf{X}(\mathrm{t})$ and $\mathbf{X}(\mathrm{t}+\tau)$, and they are called uncorrelated. As with the random vectors, the reverse is not true, as two uncorrelated observations does not necessarily mean that they are independent.

A *wide sense stationary process* is that in which the mean does not vary with time and the autocorrelation depends exclusively on the time difference[5]:

$$\begin{aligned}\boldsymbol{\mu}_{\mathrm{X_{WSS}}}(\mathrm{t}) &= \mathrm{E}\left[\mathbf{X}_{\mathrm{WSS}}(\mathrm{t})\right] = \boldsymbol{\mu}_{\mathrm{X_{WSS}}} &(30)\\ \mathbf{R}_{\mathrm{XX_{WSS}}}(\mathrm{t},\,\mathrm{t}+\tau) &= \mathrm{E}\left[\mathbf{X}_{\mathrm{WSS}}(\mathrm{t})\;\mathbf{X}^{\mathrm{T}}_{\mathrm{WSS}}(\mathrm{t}+\tau)\right] = \mathbf{R}_{\mathrm{XX_{WSS}}}(\tau) &(31)\end{aligned}$$

Consider now two stochastic processes $\mathbf{X}(\mathrm{t})$ and $\mathbf{Y}(\mathrm{t})$ defined in the same sample space Ω. The *crosscorrelation* $\mathbf{R}_{\mathrm{XY}}(\mathrm{t},\,\mathrm{t}+\tau)$ and *crosscovariance* $\mathbf{C}_{\mathrm{XY}}(\mathrm{t},\,\mathrm{t}+\tau)$ measure how similar two different processes (or signals) are when one is time shifted with respect to the other [3]:

$$\begin{aligned}\mathbf{R}_{\mathrm{XY}}(\mathrm{t},\,\mathrm{t}+\tau) &= \mathrm{E}\left[\mathbf{X}(\mathrm{t})\;\mathbf{Y}^{\mathrm{T}}(\mathrm{t}+\tau)\right] &(32)\\ \mathbf{C}_{\mathrm{XY}}(\mathrm{t},\,\mathrm{t}+\tau) &= \mathrm{E}\left[\big(\mathbf{X}(\mathrm{t})-\boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\big)\big(\mathbf{Y}(\mathrm{t}+\tau)-\boldsymbol{\mu}_{\mathrm{Y}}(\mathrm{t}+\tau)\big)^{\mathrm{T}}\right] &\\ &= \mathrm{E}\left[\mathbf{X}(\mathrm{t})\;\mathbf{Y}^{\mathrm{T}}(\mathrm{t}+\tau)\right] - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\;\boldsymbol{\mu}_{\mathrm{Y}}^{\mathrm{T}}(\mathrm{t}+\tau) = \mathbf{R}_{\mathrm{XY}}(\mathrm{t},\,\mathrm{t}+\tau) - \boldsymbol{\mu}_{\mathrm{X}}(\mathrm{t})\;\boldsymbol{\mu}_{\mathrm{Y}}^{\mathrm{T}}(\mathrm{t}+\tau) &(33)\end{aligned}$$

Two processes $\mathbf{X}(\mathrm{t})$ and $\mathbf{Y}(\mathrm{t})$ are orthogonal if their crosscorrelation is zero for all t and $\mathrm{t}+\tau$, while they are uncorrelated if their crosscovariance is zero. They are jointly wide sense stationary if their crosscorrelation is independent of the absolute time:

$$\mathbf{R}_{\mathrm{XY_{WSS}}}(\mathrm{t},\,\mathrm{t}+\tau) = \mathbf{R}_{\mathrm{XY_{WSS}}}(\tau) \tag{34}$$

Consider also a stochastic process $\mathbf{X}(\mathrm{t})$ that has one realization $\mathbf{x}(\mathrm{t})$. It is then possible to define the *time average* $\mathrm{A}\big[\mathbf{X}(\mathrm{t})\big]$ and the *time autocorrelation* $\mathrm{R}\big[\mathbf{X}(\mathrm{t})\,,\,\tau\big]$ for continuous processes as:

$$\begin{aligned}\mathrm{A}\big[\mathbf{X}(\mathrm{t})\big] &= \lim_{\mathrm{T}\to\infty}\frac{1}{2\,\mathrm{T}}\int_{-\mathrm{T}}^{\mathrm{T}}\mathbf{x}(\mathrm{t})\;\mathrm{dt} &(35)\\ \mathrm{R}\big[\mathbf{X}(\mathrm{t})\,,\,\tau\big] &= \mathrm{A}\left[\mathbf{X}(\mathrm{t})\;\mathbf{X}^{\mathrm{T}}(\mathrm{t}+\tau)\right] &(36)\end{aligned}$$

The discrete time definitions can be derived accordingly. Finally, an *ergodic* process [6] is a stationary random process for which

$$\begin{aligned}\mathrm{A}_{\mathrm{ERG}}\big[\mathbf{X}(\mathrm{t})\big] &= \mathrm{E}_{\mathrm{ERG}}\left[\mathbf{X}\right] &(37)\\ \mathrm{R}_{\mathrm{ERG}}\big[\mathbf{X}(\mathrm{t})\,,\,\tau\big] &= \mathbf{R}_{\mathrm{XX_{ERG}}}(\tau) &(38)\end{aligned}$$

2.4 White Noise

If any two random vectors $\mathbf{X}(\mathrm{t}_1)$ and $\mathbf{X}(\mathrm{t}_2)$ taken from a stochastic process $\mathbf{X}(\mathrm{t})$ are independent for all $\mathrm{t}_1 \neq \mathrm{t}_2$, then the random process $\mathbf{X}(\mathrm{t})$ is called *white noise*. Otherwise, it is known as *colored noise* [6].

The whiteness or color content of a stochastic process can be characterized by its *power spectrum* or *power spectral density* (`PSD`) $\mathrm{S}_{\mathrm{XX}}(\omega)$. For wide sense stationary processes, it is defined as the Fourier transform of its autocorrelation function $\mathrm{R}_{\mathrm{XX}}(\tau)$ [6]:

$$\mathbf{S}_{\mathrm{XX}}(\omega) = \begin{cases}\displaystyle\sum_{\mathrm{k}=-\infty}^{\infty}\mathbf{R}_{\mathrm{XX}}(\mathrm{k})\;\exp(-\mathrm{i}\,\omega\,\mathrm{k}) \quad \omega\in[-\pi,\,\pi] & \text{when X is discrete}\\ \displaystyle\int_{-\infty}^{\infty}\mathbf{R}_{\mathrm{XX}}(\tau)\;\exp(-\mathrm{i}\,\omega\,\tau)\;\mathrm{d}\tau & \text{when X is continuous}\end{cases} \tag{39}$$

[5]Strict sense stationary processes are those in which the complete `CDF` is time invariant, not only its mean and autocorrelation. The definition is usually too restrictive for any practical use.

The autocorrelation can be recovered by means of the inverse Fourier transform:

$$\begin{cases} \mathbf{R}_{\mathrm{XX}}\left(\mathrm{k}\right) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} \mathbf{S}_{\mathrm{XX}}\left(\omega\right)\, \exp\left(\mathrm{i}\,\omega\,\mathrm{k}\right)\, \mathrm{d}\omega & \text{when } \mathbf{X}\left(\mathrm{t}\right) \text{ is discrete} \\ \mathbf{R}_{\mathrm{XX}}\left(\tau\right) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} \mathbf{S}_{\mathrm{XX}}\left(\omega\right)\, \exp\left(\mathrm{i}\,\omega\,\tau\right)\, \mathrm{d}\omega & \text{when } \mathbf{X}\left(\mathrm{t}\right) \text{ is continuous} \end{cases} \tag{40}$$

In case of continuous time wide sense stationary stochastic processes[6], the *power* is defined as:

$$\mathbf{P}_{\mathrm{XX}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{S}_{\mathrm{XX}}\left(\omega\right)\, \mathrm{d}\omega \tag{41}$$

In the case of two continuous time jointly wide sense stationary stochastic processes $\mathbf{X}\left(\mathrm{t}\right)$ and $\mathbf{Y}\left(\mathrm{t}\right)$, the *cross power spectrum* $\mathrm{S}_{\mathrm{XY}}\left(\omega\right)$ is defined as the Fourier transform of their crosscorrelation $\mathrm{R}_{\mathrm{XY}}\left(\tau\right)$ [6]:

$$\mathbf{S}_{\mathrm{XY}}\left(\omega\right) = \int_{-\infty}^{\infty} \mathbf{R}_{\mathrm{XY}}\left(\tau\right)\, \exp\left(-\mathrm{i}\,\omega\,\tau\right)\, \mathrm{d}\tau \tag{42}$$

$$\mathbf{R}_{\mathrm{XY}}\left(\tau\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{S}_{\mathrm{XY}}\left(\omega\right)\, \exp\left(\mathrm{i}\,\omega\,\tau\right)\, \mathrm{d}\omega \tag{43}$$

A white noise process $\mathbf{N}\left(\mathrm{t}\right)$ (continuous time) or $\mathbf{N}\left(\mathrm{k}\right)$ (discrete time) is one whose `PSD` is constant for all frequencies, this is, a random process having equal power at all frequencies [5]. These processes do not have any correlation with themselves except at the present time [6]. The definition for discrete time processes relies on the *Kronecker delta function* δ_{k}[7]:

$$\mathbf{R}_{\mathrm{NN}}\left(\mathrm{k}\right) = \boldsymbol{\sigma}^2\, \delta_{\mathrm{k}} \tag{44}$$

$$\mathbf{S}_{\mathrm{NN}}\left(\omega\right) = \boldsymbol{\sigma}^2 = \mathbf{R}_{\mathrm{NN}}\left(0\right) \quad \forall\, \omega \in \left[-\pi,\, \pi\right] \tag{45}$$

while that for continuous time processes makes use of the *impulse Dirac delta function* $\delta\left(\tau\right)$[8]:

$$\mathbf{R}_{\mathrm{NN}}\left(\tau\right) = \boldsymbol{\sigma}^2\, \delta\left(\tau\right) \tag{46}$$

$$\mathbf{S}_{\mathrm{NN}}\left(\omega\right) = \boldsymbol{\sigma}^2 = \mathbf{R}_{\mathrm{NN}}\left(0\right) \;\; \forall \omega \in \mathbb{R} \tag{47}$$

3 Calculus Methods in Euclidean Space

This section describes the most common approaches to three frequent calculus problems, such as discrete integration, optimization, and state estimation. The solutions, known as the Runge-Kutta integration method, the Gauss-Newton optimization, and the `EKF`, are intended for state vectors in which all components can be considered Euclidean. Note that the main objective of this article is how to modify these three techniques so they can cope with non Euclidean state vectors.

3.1 Discrete Integration in Euclidean Spaces

Let's consider an Euclidean space time varying state vector $\mathbf{x}\left(\mathrm{t}\right) \in \mathbb{R}^{\mathrm{m}}$ in which its value at a given discrete time $\mathbf{x}_{\mathrm{k}} = \mathbf{x}\left(\mathrm{t}_{\mathrm{k}}\right)$ is known. The objective is to determine the state vector value at a later time $\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}\left(\mathrm{t}_{\mathrm{k+1}}\right) = \mathbf{x}\left(\mathrm{t}_{\mathrm{k}} + \Delta\mathrm{t}\right)$ by relying on evaluations of the state vector derivative with time:

$$\dot{\mathbf{x}}\left(\mathrm{t}\right) = \mathrm{f}\big(\mathbf{x}\left(\mathrm{t}\right),\, \mathrm{t}\big) \tag{48}$$

The initial value first order *ordinary differential equation* (`ODE`) problem can be solved with varying degrees of complexity and accuracy [12], three of which are described below:

[6] Similar expressions can be easily obtained for discrete time processes.

[7] The Kronecker delta function $\delta\left(\mathrm{k}\right)$ is valued 0 for all k except at $\mathrm{k} = 0$, where it is 1.

[8] The Dirac delta function $\delta\left(\tau\right)$ is valued 0 everywhere except at $\tau = 0$, where it is ∞. Its integral over any space containing $\tau = 0$ is 1.

- *Euler's method* is a first order approach that relies on evaluating the time derivative at $\mathrm{t_k}$ and considering that its value does not change for the duration of the integration interval $\Delta \mathrm{t}$. Its error is proportional to the square of the integration interval:

$$\mathbf{x}_{k+1} \approx \mathbf{x}_k + \Delta t \; \dot{\mathbf{x}}(\mathbf{x}_k, \, t_k) \tag{49}$$

- *Heun's method* is a second order approach that requires two evaluations of the time derivative. The constant gradient is estimated as the average between the time derivative evaluation at the initial state and that at the result of Euler's method, and results in an error proportional to the cube of the integration interval:

$$\mathbf{v}_1 = \dot{\mathbf{x}}(\mathbf{x}_k, \, t_k) \tag{50}$$
$$\mathbf{v}_2 = \dot{\mathbf{x}}(\mathbf{x}_k + \Delta t \; \mathbf{v}_1, t_k + \Delta t) \tag{51}$$
$$\mathbf{x}_{k+1} \approx \mathbf{x}_k + \frac{\Delta t}{2} \; [\mathbf{v}_1 + \mathbf{v}_2] \tag{52}$$

- The $\mathbf{4}^{\mathrm{th}}$ *order Runge-Kutta method* is the defacto standard and relies on four evaluations of the state vector time derivative to obtain an error proportional to the fifth power of the integration interval:

$$\mathbf{v}_1 = \dot{\mathbf{x}}(\mathbf{x}_k, \, t_k) \tag{53}$$
$$\mathbf{v}_2 = \dot{\mathbf{x}}(\mathbf{x}_k + \frac{\Delta t \; \mathbf{v}_1}{2}, t_k + \frac{\Delta t}{2}) \tag{54}$$
$$\mathbf{v}_3 = \dot{\mathbf{x}}(\mathbf{x}_k + \frac{\Delta t \; \mathbf{v}_2}{2}, t_k + \frac{\Delta t}{2}) \tag{55}$$
$$\mathbf{v}_4 = \dot{\mathbf{x}}(\mathbf{x}_k + \Delta t \; \mathbf{v}_3, t_k + \Delta t) \tag{56}$$
$$\mathbf{x}_{k+1} \approx \mathbf{x}_k + \Delta t \; \left[\frac{\mathbf{v}_1}{6} + \frac{\mathbf{v}_2}{3} + \frac{\mathbf{v}_3}{3} + \frac{\mathbf{v}_4}{6}\right] \tag{57}$$

3.2 Gradient Descent Optimization in Euclidean Spaces

Let's consider an Euclidean vector $\mathbf{x} \in \mathbb{R}^m$, a nonlinear map $\{\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n \mid \mathbf{f}(\mathbf{x}) \in \mathbb{R}^n, \forall \; \mathbf{x} \in \mathbb{R}^m\}$ for which it is also possible to evaluate its jacobian $\{\mathbf{J} : \mathbb{R}^m \rightarrow \mathbb{R}^{nxm} \mid \mathbf{J}(\mathbf{x}) = \partial \mathbf{f}(\mathbf{x}) / \partial \mathbf{x} \in \mathbb{R}^{nxm}, \forall \; \mathbf{x} \in \mathbb{R}^m\}$, and an error or cost function $\boldsymbol{\mathcal{E}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - \mathbf{f}_T \; \in \mathbb{R}^n$ containing the difference between the map $\mathbf{f}$ and a target result $\mathbf{f}_T$.

Let's also consider that the objective is to determine an input vector $\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x}$ in the vicinity of a known initial value $\mathbf{x}_0$, for which the cost function norm $\|\boldsymbol{\mathcal{E}}(\mathbf{x})\| \in \mathbb{R}$ holds a local minimum, this is, $\|\boldsymbol{\mathcal{E}}(\mathbf{x}_0 + \Delta\mathbf{x})\| < \|\boldsymbol{\mathcal{E}}(\mathbf{x}_0)\|, \forall \; \mathbf{x}_0 \in \mathbb{R}^m$. The *Gauss-Newton* optimization method provides a solution to this problem that relies on iteratively advancing the solution per (58) starting with $\mathbf{x}_0$:

$$\mathbf{x}_{k+1} \longleftarrow \mathbf{x}_k + \Delta\mathbf{x}_k \tag{58}$$

Adopting a lighter notation in which $\mathbf{f}_k = \mathbf{f}(\mathbf{x}_k)$, $\mathbf{J}_k = \mathbf{J}(\mathbf{x}_k)$, and $\boldsymbol{\mathcal{E}}_k = \boldsymbol{\mathcal{E}}(\mathbf{x}_k)$, the process concludes when the step diminution of the cost function norm falls below a given threshold ($\|\boldsymbol{\mathcal{E}}_k\| - \|\boldsymbol{\mathcal{E}}_{k+1}\| < \delta$). The Gauss-Newton method consists on linearizing each step by performing a first order Taylor expansion of the cost function before minimizing its norm by equaling its derivative with respect to $\Delta\mathbf{x}_k$ to zero [13]:

$$\boldsymbol{\mathcal{E}}_{k+1} = \mathbf{f}_{k+1} - \mathbf{f}_T \approx \mathbf{f}_k + \mathbf{J}_k \, \Delta\mathbf{x}_k - \mathbf{f}_T = \boldsymbol{\mathcal{E}}_k + \mathbf{J}_k \, \Delta\mathbf{x}_k \tag{59}$$
$$\|\boldsymbol{\mathcal{E}}_{k+1}\| = \boldsymbol{\mathcal{E}}_{k+1}^T \, \boldsymbol{\mathcal{E}}_{k+1} = \boldsymbol{\mathcal{E}}_k^T \, \boldsymbol{\mathcal{E}}_k + \Delta\mathbf{x}_k^T \, \mathbf{J}_k^T \, \mathbf{J}_k \, \Delta\mathbf{x}_k + 2 \, \Delta\mathbf{x}_k^T \, \mathbf{J}_k^T \, \boldsymbol{\mathcal{E}}_k \tag{60}$$
$$\frac{\partial\|\boldsymbol{\mathcal{E}}_{k+1}\|}{\partial\Delta\mathbf{x}_k} = 0 \longrightarrow 2 \, \mathbf{J}_k^T \, \mathbf{J}_k \, \Delta\mathbf{x}_k + 2 \, \mathbf{J}_k^T \, \boldsymbol{\mathcal{E}}_k = 0 \longrightarrow \Delta\mathbf{x}_k = -\left(\mathbf{J}_k^T \, \mathbf{J}_k\right)^{-1} \mathbf{J}_k^T \, \boldsymbol{\mathcal{E}}_k \tag{61}$$

The Gauss-Newton algorithm is just one type of a more generic class of iterative minimization methods grouped under the name of *gradient descent methods*. The *Newton* method relies on minimizing (equaling its $\Delta\mathbf{x}_k$ derivative to zero) a second order Taylor expansion of the cost function norm $\mathrm{N}(\mathbf{x}) = \|\boldsymbol{\mathcal{E}}(\mathbf{x})\| \; \in \mathbb{R}$, which requires the computation of both its gradient $\left\{\boldsymbol{\nabla} : \mathbb{R}^m \rightarrow \mathbb{R}^{1xm} \mid \boldsymbol{\nabla}(\mathbf{x}) = \partial \mathrm{N}(\mathbf{x}) / \partial \mathbf{x} \in \mathbb{R}^{1xm}, \forall \; \mathbf{x} \in \mathbb{R}^m\right\}$ and its Hessian $\left\{\mathbf{H} : \mathbb{R}^m \rightarrow \mathbb{R}^{mxm} \mid \mathbf{H}(\mathbf{x}) = \partial^2 \mathrm{N}(\mathbf{x}) / \partial \mathbf{x}^2 \in \mathbb{R}^{mxm}, \forall \; \mathbf{x} \in \mathbb{R}^m\right\}$ at each step, resulting in:

$$\Delta\mathbf{x}_k = -\mathbf{H}_k^{-1} \; \boldsymbol{\nabla}_k^T \tag{62}$$

As the Hessian $\mathbf{H}$ can be difficult or expensive to compute, there exist several approximations that reduce the computational cost of each step, such as the *steepest descent* method, which replaces the Hessian with the product of a constant and the identity matrix $\mathbf{I}_m \in \mathbb{R}^{mxm}$, and the *diagonal approximation*, which sets to zero all $\mathbf{H}$ components outside its main diagonal. In this sense, the Gauss-Newton method is just another approximation that employs a first order simplification of the Hessian, as proven in [14].

The convergence of none of these methods is guaranteed. In general, both Gauss-Newton and Newton work better near the local minimum, where the quadratic approximation is good, but may diverge when the initial value is further away, where the steepest descent and diagonal approximation methods may be more robust. To ensure that the error gets smaller in each iteration, it may be convenient to advance with a smaller $\Delta\mathbf{x}_k$ step. The *Levenberg-Marquardt* algorithm employs a varying ratio between the Gauss-Newton (or Newton) and diagonal approximations to the Hessian, moving towards the former when the error $\|\boldsymbol{\mathcal{E}}_k\|$ decreases, and towards the later while repeating the step if it increases.

3.3 State Estimation in Euclidean Spaces

State estimation is the problem of determining the value of the state of a dynamic system based on a series of noisy equations that describe the evolution of the state with time, together with a series of noisy measurements or observations of variables that also depend on the state. *State* or *state vector* refers to those variables that provide a representation of the condition or status of the system at a given instant in time. Section 3.3.1 discusses the equations that describe the system dynamics, this is, the state evolution with time, and what is the best possible state estimation that can be obtained from them. Section 3.3.2 describes the measurement or observation equations, and also reaches the best possible state estimate from the information they contain. Both approaches are combined in section 3.3.3, which describes the extended Kalman filter or `EKF`, the most widely used non linear state estimation algorithm.

3.3.1 Sampled Data Systems

A *state space system* is a mathematical representation of a physical process in which the variables (both state and input) are related by first order differential equations (for continuous systems) or difference equations (for discrete ones). If the state of the system (the value of the state variables) is known at a given time, and so are all the present and future inputs (the evolution with time of the input variables), it is then possible to obtain the evolution with time of all the state variables.

A *sampled data system* is one whose dynamics are described by continuous time differential equations, but whose inputs only change at discrete time instants. Additionally, it is only necessary to estimate the state variables, or to be precise its mean and covariance[9], at those same discrete time instants [6]. A continuous time nonlinear state system can be written as

$$\dot{\mathbf{x}}(t) = f\big(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t), t\big) \tag{63}$$

where $\mathbf{x} \in \mathbb{R}^m$ is the state vector, $\mathbf{u} \in \mathbb{R}^n$ is the known *control* or *input vector*, and $\mathbf{w} \in \mathbb{R}^p$ is the *process noise*. These three vectors may have different sizes. Let's also assume that the process noise $\mathbf{w}(t)$ can be modeled by a zero mean continuous time white noise random process[10] of covariance $\mathbf{Q}_c$ (sections 2.1 and 2.4):

$$\mathbf{w}(t) \sim (\mathbf{0}, \mathbf{Q}_c) \tag{64}$$

$$\mathbf{R}_{ww}(t, \tau) = E\big[\mathbf{w}(t)\, \mathbf{w}^T(\tau)\big] = \mathbf{Q}_c\, \delta(t - \tau) \tag{65}$$

Linearization of Continuous Time Systems

[9] Although the term covariance is traditionally employed in state estimation, it is in fact referring to the state random vector autocovariance provided by (24), or to the state random process autocovariance given by (29), depending on context.

[10] Note that the process noise does not need to be Gaussian.

The dynamics represented by (63) can be linearized by performing a Taylor expansion around an unknown nominal state $\mathbf{x}_{\mathrm{N}}(\mathrm{t})$ and process noise $\mathbf{w}_{\mathrm{N}}(\mathrm{t})$[11], assuming without loss of generality that $\mathbf{w}_{\mathrm{N}}(\mathrm{t}) = \mathbf{0}$. If it is not, it can be written as the sum of a zero mean part and a known deterministic part, which can then be added to the control vector. The expansion is truncated so only the first order terms remain, introducing linearization errors; these are higher the more nonlinear that $\mathrm{f}(\mathbf{x}, \mathbf{u}, \mathbf{w}, \mathrm{t})$ is with respect to $\mathbf{x}$ and $\mathbf{w}$, and the farther away that $\mathbf{x}(\mathrm{t})$ lies from $\mathbf{x}_{\mathrm{N}}(\mathrm{t})$ and $\mathbf{w}(\mathrm{t})$ from $\mathbf{w}_{\mathrm{N}} = \mathbf{0}$ [6].

$$\dot{\mathbf{x}}(\mathrm{t}) \approx \mathrm{f}|_{\mathrm{N}} + \left.\frac{\partial \mathrm{f}}{\partial \mathbf{x}}\right|_{\mathrm{N}} (\mathbf{x} - \mathbf{x}_{\mathrm{N}}) + \left.\frac{\partial \mathrm{f}}{\partial \mathbf{w}}\right|_{\mathrm{N}} \mathbf{w} = \left.\frac{\partial \mathrm{f}}{\partial \mathbf{x}}\right|_{\mathrm{N}} \mathbf{x} + \left(\mathrm{f}|_{\mathrm{N}} - \left.\frac{\partial \mathrm{f}}{\partial \mathbf{x}}\right|_{\mathrm{N}} \mathbf{x}_{\mathrm{N}}\right) + \left.\frac{\partial \mathrm{f}}{\partial \mathbf{w}}\right|_{\mathrm{N}} \mathbf{w} \tag{66}$$

where $|_{\mathrm{N}}$ stands for evaluation at $\big(\mathbf{x}_{\mathrm{N}}(\mathrm{t}), \mathbf{u}(\mathrm{t}), \mathbf{0}, \mathrm{t}\big)$. The state system is now continuous time but linear:

$$\begin{aligned}
\dot{\mathbf{x}}(\mathrm{t}) &\approx \mathbf{A}(\mathrm{t})\,\mathbf{x}(\mathrm{t}) + \mathbf{B}(\mathrm{t})\,\widetilde{\mathbf{u}}(\mathrm{t}) + \widetilde{\mathbf{w}}(\mathrm{t}) && (67)\\
\mathbf{A}(\mathrm{t}) &= \frac{\partial \mathrm{f}}{\partial \mathbf{x}}\big(\mathbf{x}_{\mathrm{N}}(\mathrm{t}), \mathbf{u}(\mathrm{t}), \mathbf{0}, \mathrm{t}\big) && (68)\\
\mathbf{B}(\mathrm{t}) &= \mathbf{I} && (69)\\
\mathbf{L}(\mathrm{t}) &= \frac{\partial \mathrm{f}}{\partial \mathbf{w}}\big(\mathbf{x}_{\mathrm{N}}(\mathrm{t}), \mathbf{u}(\mathrm{t}), \mathbf{0}, \mathrm{t}\big) && (70)\\
\widetilde{\mathbf{u}}(\mathrm{t}) &= \mathrm{f}\big(\mathbf{x}_{\mathrm{N}}(\mathrm{t}), \mathbf{u}(\mathrm{t}), \mathbf{0}, \mathrm{t}\big) - \mathbf{A}(\mathrm{t})\,\mathbf{x}_{\mathrm{N}}(\mathrm{t}) && (71)\\
\widetilde{\mathbf{w}}(\mathrm{t}) &= \mathbf{L}(\mathrm{t})\,\mathbf{w}(\mathrm{t}) \sim \Big(\mathbf{0}, \mathbf{L}\,\mathbf{Q}_{\mathrm{c}}\,\mathbf{L}^{\mathrm{T}}\Big) = \Big(\mathbf{0}, \widetilde{\mathbf{Q}}_{\mathrm{c}}(\mathrm{t})\Big) && (72)\\
\mathbf{R}_{\widetilde{\mathbf{w}}\widetilde{\mathbf{w}}}(\mathrm{t}, \tau) &= \mathrm{E}\Big[\widetilde{\mathbf{w}}(\mathrm{t})\,\widetilde{\mathbf{w}}^{\mathrm{T}}(\tau)\Big] = \widetilde{\mathbf{Q}}_{\mathrm{c}}(\mathrm{t})\,\delta(\mathrm{t} - \tau) && (73)
\end{aligned}$$

The above linear state system is based on a unitary *input matrix* $\mathbf{B} \in \mathbb{R}^{\mathrm{mxm}}$ and a *system matrix* $\mathbf{A}(\mathrm{t}) \in \mathbb{R}^{\mathrm{mxm}}$ that is the jacobian of the non linear system with respect to the state vector evaluated at the unknown nominal state. It also employs modified input $\widetilde{\mathbf{u}}(\mathrm{t}) \in \mathbb{R}^{\mathrm{m}}$ and process noise $\widetilde{\mathbf{w}}(\mathrm{t}) \in \mathbb{R}^{\mathrm{m}}$ vectors.

Comparison of Integrated Continuous and Discrete White Noise Processes

Before continuing, let's make a parenthesis to compare the behavior of an integrated continuous white noise process with that of a discrete one, as the result is essential to the discretization of the continuous time state system (67). According to section 2.4, a continuous zero mean white noise is defined by $\mathbf{w}(\mathrm{t}) \sim (\mathbf{0}, \mathbf{Q}_{\mathrm{c}})$ and $\mathrm{E}\big[\mathbf{w}(\mathrm{t})\,\mathbf{w}^{\mathrm{T}}(\tau)\big] = \mathbf{Q}_{\mathrm{c}}\,\delta(\mathrm{t} - \tau)$, while a zero mean discrete time white noise process responds to $\mathbf{w}_{\mathrm{k}} \sim (\mathbf{0}, \mathbf{Q}_{\mathrm{d}})$ and $\mathrm{E}\big[\mathbf{w}_{\mathrm{k}}\,\mathbf{w}_{\mathrm{l}}^{\mathrm{T}}\big] = \mathbf{Q}_{\mathrm{d}}\,\delta_{\mathrm{k-l}}$. Let's integrate the continuous white noise with $\dot{\mathbf{z}}(\mathrm{t}) = \mathbf{w}(\mathrm{t})$, $\mathbf{z}(0) = \mathbf{0}$ and analyze the variation with time of the mean $\boldsymbol{\mu}_{\mathrm{z}}(\mathrm{t})$ and covariance $\mathbf{C}_{\mathrm{zz}}(\mathrm{t})$ of the integrated noise $\mathbf{z}(\mathrm{t})$:

$$\begin{aligned}
\boldsymbol{\mu}_{\mathrm{z}}(\mathrm{t}) &= \mathrm{E}[\mathbf{z}(\mathrm{t})] = \mathrm{E}\left[\int_0^{\mathrm{t}} \mathbf{w}(\alpha)\,\mathrm{d}\alpha\right] = \int_0^{\mathrm{t}} \mathrm{E}[\mathbf{w}(\alpha)]\,\mathrm{d}\alpha = \mathbf{0} && (74)\\
\mathbf{C}_{\mathrm{zz}}(\mathrm{t}) &= \mathrm{E}\big[\mathbf{z}(\mathrm{t})\,\mathbf{z}^{\mathrm{T}}(\mathrm{t})\big] - \boldsymbol{\mu}_{\mathrm{z}}(\mathrm{t})\,\boldsymbol{\mu}_{\mathrm{z}}^{\mathrm{T}}(\mathrm{t}) = \mathrm{E}\left[\int_0^{\mathrm{t}} \mathbf{w}(\alpha)\,\mathrm{d}\alpha \int_0^{\mathrm{t}} \mathbf{w}^{\mathrm{T}}(\beta)\,\mathrm{d}\beta\right]\\
&= \int_0^{\mathrm{t}}\int_0^{\mathrm{t}} \mathrm{E}\big[\mathbf{w}(\alpha)\,\mathbf{w}^{\mathrm{T}}(\beta)\big]\,\mathrm{d}\alpha\,\mathrm{d}\beta = \mathbf{Q}_{\mathrm{c}} \int_0^{\mathrm{t}}\int_0^{\mathrm{t}} \delta(\alpha - \beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta = \mathbf{Q}_{\mathrm{c}} \int_0^{\mathrm{t}} \mathrm{d}\beta = \mathbf{Q}_{\mathrm{c}}\,\mathrm{t} && (75)
\end{aligned}$$

This expression shows that the mean of an integrated continuous white noise is always zero, but its covariance grows linearly with time. Integrating now the difference equation $\mathbf{z}_{\mathrm{k}} = \mathbf{z}_{\mathrm{k-1}} + \mathbf{w}_{\mathrm{k-1}}$, $\mathbf{z}_0 = \mathbf{0}$, let's also evaluate the variation with time of the mean $\boldsymbol{\mu}_{\mathrm{k}}$ and covariance $\mathbf{C}_{\mathrm{zz,k}}$ of the integrated noise $\mathbf{z}_{\mathrm{k}}$:

$$\begin{aligned}
\boldsymbol{\mu}_{\mathrm{k}} &= \mathrm{E}[\mathbf{z}_{\mathrm{k}}] = \mathrm{E}\left[\sum_{\mathrm{l}=0}^{\mathrm{k-1}} \mathbf{w}_{\mathrm{l}}\right] = \sum_{\mathrm{l}=0}^{\mathrm{k-1}} \mathrm{E}[\mathbf{w}_{\mathrm{l}}] = \mathbf{0} && (76)\\
\mathbf{C}_{\mathrm{zz,k}} &= \mathrm{E}\big[\mathbf{z}_{\mathrm{k}}\,\mathbf{z}_{\mathrm{k}}^{\mathrm{T}}\big] - \boldsymbol{\mu}_{\mathrm{k}}\,\boldsymbol{\mu}_{\mathrm{k}}^{\mathrm{T}} = \mathrm{E}\left[\sum_{\mathrm{l}=0}^{\mathrm{k-1}} \mathbf{w}_{\mathrm{l}} \sum_{\mathrm{m}=0}^{\mathrm{k-1}} \mathbf{w}_{\mathrm{m}}^{\mathrm{T}}\right] = \sum_{\mathrm{l}=0}^{\mathrm{k-1}}\sum_{\mathrm{m}=0}^{\mathrm{k-1}} \mathrm{E}\big[\mathbf{w}_{\mathrm{l}}\,\mathbf{w}_{\mathrm{m}}^{\mathrm{T}}\big]\\
&= \mathbf{Q}_{\mathrm{d}} \sum_{\mathrm{l}=0}^{\mathrm{k-1}}\sum_{\mathrm{m}=0}^{\mathrm{k-1}} \delta_{\mathrm{l-m}} = \mathbf{Q}_{\mathrm{d}} \sum_{\mathrm{l}=0}^{\mathrm{k-1}} 1 = \mathbf{Q}_{\mathrm{d}}\,\mathrm{k} && (77)
\end{aligned}$$

[11] As the input vector $\mathbf{u}(\mathrm{t})$ is known, there is no need to expand around it.

The covariance of the integrated discrete white noise process also grows linearly with time. Considering a sampling period of Δt, $t = k \cdot \Delta t$, a discrete zero mean white noise process can be considered equivalent [6] to a continuous one if their covariances are related by:

$$\mathbf{Q}_d = \mathbf{Q}_c \cdot \Delta t \tag{78}$$

Discretization of Linear Continuous Time Systems

Returning to the main argument, and considering that the state vector needs to be known only at a series of discrete time points, it is possible to discretize the (67) linear continuous time system if $\mathbf{A}(t)$, $\mathbf{B}(t)$, and $\widetilde{\mathbf{u}}(t)$ are considered constant during the integration interval, which starts at $t_{k-1} = (k-1) \cdot \Delta t$ and concludes at $t_k = k \cdot \Delta t$. The introduced discretization errors are higher the farther away this assumption is from reality. Introducing (78), the state system is now discrete and linear [6]:

$$\mathbf{x}_k \approx \mathbf{F}_{k-1}\,\mathbf{x}_{k-1} + \mathbf{G}_{k-1}\,\widetilde{\mathbf{u}}_{k-1} + \widetilde{\mathbf{w}}_{k-1} \tag{79}$$

$$\mathbf{x}_k = \mathbf{x}(t_k) = \mathbf{x}(k\,\Delta t) \tag{80}$$

$$\mathbf{F}_k = \exp(\mathbf{A}_k\,\Delta t) = \exp\big(\mathbf{A}(k\,\Delta t)\ \Delta t\big) \tag{81}$$

$$\begin{aligned}\mathbf{G}_k &= \mathbf{F}_k \int_0^{\Delta t} \exp\big(-\mathbf{A}(\tau)\ \tau\big)\,d\tau\,\mathbf{B}(k\,\Delta t) \\ &= \mathbf{F}_k\,\big[\mathbf{I} - \exp\big(-\mathbf{A}(k\,\Delta t)\,\Delta t\big)\big]\,\mathbf{A}^{-1}(k\,\Delta t)\ \mathbf{B}(k\,\Delta t)\end{aligned} \tag{82}$$

$$\widetilde{\mathbf{u}}_k = \widetilde{\mathbf{u}}(t_k) = \widetilde{\mathbf{u}}(k\,\Delta t) \tag{83}$$

$$\widetilde{\mathbf{w}}_k = \widetilde{\mathbf{w}}(k\,\Delta t) \sim \Big(\mathbf{0},\, \widetilde{\mathbf{Q}}_c(k \cdot \Delta t) \cdot \Delta t\Big) = \Big(\mathbf{0},\, \mathbf{L}_k\,\mathbf{Q}_c\,\mathbf{L}_k^T\,\Delta t\Big) = \Big(\mathbf{0},\, \widetilde{\mathbf{Q}}_{d,k}\Big) \tag{84}$$

$$\mathbf{R}_{\widetilde{w}\widetilde{w},kj} = E\left[\widetilde{\mathbf{w}}_k\,\widetilde{\mathbf{w}}_j^T\right] = \widetilde{\mathbf{Q}}_{d,k}\,\delta_{k-j} \tag{85}$$

Note that both the *system state transition matrix* $\mathbf{F}_k \in \mathbb{R}^{mxm}$ and the *input transition matrix* $\mathbf{G}_k \in \mathbb{R}^{mxm}$ make use of the matrix exponential function, although computing the later is not required, as shown in section 3.3.3.

Mean and Covariance of State Vector

It is possible to evaluate the mean $\boldsymbol{\mu}_{x,k}$ and covariance $\mathbf{C}_{xx,k} = \mathbf{P}_k$ of the state vector given by (79), which provide their variation with time[12]:

$$\boldsymbol{\mu}_{x,k} = E[\mathbf{x}_k] = \mathbf{F}_{k-1}\,\boldsymbol{\mu}_{x,k-1} + \mathbf{G}_{k-1}\,\widetilde{\mathbf{u}}_{k-1} \tag{86}$$

$$\begin{aligned}\mathbf{C}_{xx,k} &= \mathbf{P}_k = E\left[\big(\mathbf{x}_k - \boldsymbol{\mu}_{x,k}\big)\ \big(\mathbf{x}_k - \boldsymbol{\mu}_{x,k}\big)^T\right] \\ &= E\left[\Big(\mathbf{F}_{k-1}\ \big(\mathbf{x}_{k-1} - \boldsymbol{\mu}_{x,k-1}\big) + \widetilde{\mathbf{w}}_{k-1}\Big)\Big(\mathbf{F}_{k-1}\ \big(\mathbf{x}_{k-1} - \boldsymbol{\mu}_{x,k-1}\big) + \widetilde{\mathbf{w}}_{k-1}\Big)^T\right] \\ &= \mathbf{F}_{k-1}\,\mathbf{C}_{xx,k-1}\,\mathbf{F}_{k-1}^T + \widetilde{\mathbf{Q}}_{d,k-1} = \mathbf{F}_{k-1}\,\mathbf{P}_{k-1}\,\mathbf{F}_{k-1}^T + \widetilde{\mathbf{Q}}_{d,k-1}\end{aligned} \tag{87}$$

Based on (79), $\mathbf{x}_k$ is a linear combination of a series of known real vectors $\widetilde{\mathbf{u}}_0, \ldots, \widetilde{\mathbf{u}}_{k-1}$ plus a series of independent random vectors $\mathbf{x}_0, \widetilde{\mathbf{w}}_0, \ldots, \widetilde{\mathbf{w}}_{k-1}$. According to the central limit theorem stated in section 2.1, $\mathbf{x}_k \sim N\left(\boldsymbol{\mu}_{x,k},\, \mathbf{C}_{xx,k}\right) = N\left(\boldsymbol{\mu}_{x,k},\, \mathbf{P}_k\right)$ is a normal or Gaussian random vector completely characterized by its mean and covariance.

The summary of this section is that given a continuous time non linear state space system such as (63), it is possible, with some linearization and discretization errors, to transform it into an equivalent discrete time linear system (79) that can be integrated to obtain the estimated value of the state vector $\mathbf{x}(t)$ at a series of discrete times $t_k = k\,\Delta t$ characterized by its mean $\boldsymbol{\mu}_{x,k}$ (86) and covariance $\mathbf{C}_{xx,k} = \mathbf{P}_k$ (87). Without further assistance, (87) shows that the uncertainty of the results grows with time because of the accumulation of the white noise present in the system [6]. The next section shows how the addition of measurements can solve this problem.

[12] To compute the covariance, note that there is no correlation between $\big(\mathbf{x}_{k-1} - \mu\mathbf{x}_{k-1}\big)$ and $\widetilde{\mathbf{w}}_{k-1}$.

3.3.2 Sampled Observations

Given the sampled data system of section 3.3.1, let's imagine that there exist a series of sensors capable of measuring certain variables related to the state vector at the same time points at which the state system is discretized in section 3.3.1:

$$\mathbf{y}_{\mathrm{k}} = \mathrm{h}\left(\mathbf{x}_{\mathrm{k}},\, \mathbf{v}_{\mathrm{k}},\, \mathrm{t}_{\mathrm{k}}\right) \tag{88}$$

where $\mathbf{y}_{\mathrm{k}} = \mathbf{y}\left(\mathrm{t}_{\mathrm{k}}\right) \in \mathbb{R}^{\mathrm{q}}$ is the *measurement* or *observation vector* provided by the sensors, $\mathbf{x}_{\mathrm{k}} = \mathbf{x}\left(\mathrm{t}_{\mathrm{k}}\right) \in \mathbb{R}^{\mathrm{m}}$ is the state vector, $\mathrm{t}_{\mathrm{k}} = \mathrm{t}\left(\mathrm{k}\,\Delta\mathrm{t}\right)$ is the discrete time at which the measurements are taken, and $\mathbf{v}_{\mathrm{k}} = \mathbf{v}\left(\mathrm{t}_{\mathrm{k}}\right) \in \mathbb{R}^{\mathrm{q}}$ is the *measurement* or *observation noise*, which can be modeled by a zero mean white noise random process[13] of covariance $\mathbf{R}$ (sections 2.1 and 2.4):

$$\begin{aligned} \mathbf{v}_{\mathrm{k}} &\sim (\mathbf{0},\, \mathbf{R}) && (89)\\ \mathbf{R}_{\mathrm{vv,kj}} &= \mathrm{E}\left[\mathbf{v}_{\mathrm{k}}\,\mathbf{v}_{\mathrm{j}}^{\mathrm{T}}\right] = \mathbf{R}\,\delta_{\mathrm{k-j}} && (90) \end{aligned}$$

Let's also assume that the measurement noise and the process noise of section 3.3.1 are orthogonal:

$$\mathbf{R}_{\mathrm{vw,kj}} = \mathrm{E}\left[\mathbf{v}_{\mathrm{k}}\,\mathbf{w}_{\mathrm{j}}^{\mathrm{T}}\right] = \mathbf{0} \tag{91}$$

Linearization of Observations

The discrete observations represented by (88) can be linearized by performing a Taylor expansion around an unknown nominal state $\mathbf{x}_{\mathrm{Nk}} = \mathbf{x}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right)$ and observation noise $\mathbf{v}_{\mathrm{Nk}} = \mathbf{v}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right)$, assuming without loss of generality that $\mathbf{v}_{\mathrm{Nk}} = \mathbf{0}$. If it is not, it can be written as the sum of a zero mean part and a known deterministic part, which can then be included in the nonlinear function h. The expansion is truncated so only the first order terms remain, introducing linearization errors; these are higher the more nonlinear that $\mathrm{h}\left(\mathbf{x}_{\mathrm{k}},\, \mathbf{v}_{\mathrm{k}},\, \mathrm{t}_{\mathrm{k}}\right)$ is with respect to $\mathbf{x}_{\mathrm{k}}$ and $\mathbf{v}_{\mathrm{k}}$, and the farther away that $\mathbf{x}_{\mathrm{k}}$ is from $\mathbf{x}_{\mathrm{Nk}}$ and $\mathbf{v}_{\mathrm{k}}$ from $\mathbf{v}_{\mathrm{Nk}} = \mathbf{0}$ [6].

$$\mathbf{y}_{\mathrm{k}} \approx \mathrm{h}|_{\mathrm{N}} + \left.\frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\right|_{\mathrm{N}} \left(\mathbf{x}_{\mathrm{k}} - \mathbf{x}_{\mathrm{Nk}}\right) + \left.\frac{\partial \mathrm{h}}{\partial \mathbf{v}_{\mathrm{k}}}\right|_{\mathrm{N}} \mathbf{v}_{\mathrm{k}} = \left.\frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\right|_{\mathrm{N}} \mathbf{x}_{\mathrm{k}} + \left(\mathrm{h}|_{\mathrm{N}} - \left.\frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\right|_{\mathrm{N}} \mathbf{x}_{\mathrm{Nk}}\right) + \left.\frac{\partial \mathrm{h}}{\partial \mathbf{v}_{\mathrm{k}}}\right|_{\mathrm{N}} \mathbf{v}_{\mathrm{k}} \tag{92}$$

where $|_{\mathrm{N}}$ stands for evaluation at $\left(\mathbf{x}_{\mathrm{Nk}},\, \mathbf{0},\, \mathrm{t}_{\mathrm{k}}\right)$. The observations system is now discrete time and linear:

$$\begin{aligned} \mathbf{y}_{\mathrm{k}} &\approx \mathbf{H}_{\mathrm{k}}\,\mathbf{x}_{\mathrm{k}} + \mathbf{z}_{\mathrm{k}} + \widetilde{\mathbf{v}}_{\mathrm{k}} && (93)\\ \mathbf{H}_{\mathrm{k}} &= \mathbf{H}\left(\mathrm{t}_{\mathrm{k}}\right) = \frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\left(\mathbf{x}_{\mathrm{Nk}},\, \mathbf{0},\, \mathrm{t}_{\mathrm{k}}\right) && (94)\\ \mathbf{z}_{\mathrm{k}} &= \mathbf{z}\left(\mathrm{t}_{\mathrm{k}}\right) = \mathrm{h}\left(\mathbf{x}_{\mathrm{Nk}},\, \mathbf{0},\, \mathrm{t}_{\mathrm{k}}\right) - \mathbf{H}_{\mathrm{k}}\,\mathbf{x}_{\mathrm{Nk}} && (95)\\ \mathbf{M}_{\mathrm{k}} &= \mathbf{M}\left(\mathrm{t}_{\mathrm{k}}\right) = \frac{\partial \mathrm{h}}{\partial \mathbf{v}_{\mathrm{k}}}\left(\mathbf{x}_{\mathrm{Nk}},\, \mathbf{0},\, \mathrm{t}_{\mathrm{k}}\right) && (96)\\ \widetilde{\mathbf{v}}_{\mathrm{k}} &= \widetilde{\mathbf{v}}\left(\mathrm{t}_{\mathrm{k}}\right) = \mathbf{M}_{\mathrm{k}}\,\mathbf{v}_{\mathrm{k}} \sim \left(\mathbf{0},\, \mathbf{M}_{\mathrm{k}}\,\mathbf{R}\,\mathbf{M}_{\mathrm{k}}^{\mathrm{T}}\right) = \left(\mathbf{0},\, \widetilde{\mathbf{R}}_{\mathrm{k}}\right) && (97)\\ \mathbf{R}_{\widetilde{\mathrm{v}}\widetilde{\mathrm{v}},\mathrm{kj}} &= \mathrm{E}\left[\widetilde{\mathbf{v}}_{\mathrm{k}}\,\widetilde{\mathbf{v}}_{\mathrm{j}}^{\mathrm{T}}\right] = \widetilde{\mathbf{R}}_{\mathrm{k}}\,\delta_{\mathrm{k-j}} && (98) \end{aligned}$$

The above observation system is based on an *output matrix* $\mathbf{H}_{\mathrm{k}} \in \mathbb{R}^{\mathrm{qxm}}$ that is the jacobian of the non linear system with respect to the state vector evaluated at the unknown nominal state, and a vector $\mathbf{z}_{\mathrm{k}} \in \mathbb{R}^{\mathrm{q}}$ that depends exclusively of the nominal state. It also employs a modified observation noise vector $\widetilde{\mathbf{v}}_{\mathrm{k}} \in \mathbb{R}^{\mathrm{q}}$. It is worth noting that computation of $\mathbf{z}_{\mathrm{k}}$ is not necessary to obtain the solution, as shown in section 3.3.3.

Constant State Vector Estimation based on Observations

The objective of this section is to obtain the best possible estimate $\hat{\mathbf{x}}_{\mathrm{k}}$ of a constant[14] state vector $\mathbf{x}$ based on the observations $\mathbf{y}_{\mathrm{k}}$ provided by (93) and the previous estimate $\hat{\mathbf{x}}_{\mathrm{k-1}}$. Let's use an expression like (99), where $\mathbf{K}_{\mathrm{k}} \in \mathbb{R}^{\mathrm{mxq}}$ is called the *gain matrix* and $\mathbf{r}_{\mathrm{k}} \in \mathbb{R}^{\mathrm{q}}$ the *innovations vector*:

$$\hat{\mathbf{x}}_{\mathrm{k}} = \hat{\mathbf{x}}_{\mathrm{k-1}} + \mathbf{K}_{\mathrm{k}}\,\mathbf{r}_{\mathrm{k}} = \hat{\mathbf{x}}_{\mathrm{k-1}} + \mathbf{K}_{\mathrm{k}}\,\left(\mathbf{y}_{\mathrm{k}} - \mathbf{H}_{\mathrm{k}}\,\hat{\mathbf{x}}_{\mathrm{k-1}} - \mathbf{z}_{\mathrm{k}}\right) \tag{99}$$

[13] Note that the measurement noise does not need to be Gaussian.

[14] Note that this is the only section where the state vector $\mathbf{x}_{\mathrm{k}}$ is required to be constant, this is, $\mathbf{x} = \mathbf{x}_0 = \mathbf{x}_{\mathrm{k}}\ \forall\,\mathrm{k}$.

The *estimation error* $\boldsymbol{\varepsilon}_{x,k}$ and its mean can then be computed based on (99) and (93):

$$\begin{aligned} \boldsymbol{\varepsilon}_{x,k} &= \mathbf{x} - \hat{\mathbf{x}}_k = (\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k) \; \boldsymbol{\varepsilon}_{x,k-1} - \mathbf{K}_k \, \widetilde{\mathbf{v}}_k & (100) \\ \boldsymbol{\mu}_{\varepsilon x,k} &= \mathrm{E}\left[\boldsymbol{\varepsilon}_{x,k}\right] = (\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k) \; \boldsymbol{\mu}_{\varepsilon x,k-1} & (101) \end{aligned}$$

As the linearized discrete noise $\widetilde{\mathbf{v}}_k$ is zero mean per (97), (99) is called an *unbiased estimator* [6], because if the initial estimate $\hat{\mathbf{x}}_0$ is set equal to the expected value of the state vector $\left(\hat{\mathbf{x}}_0 = \boldsymbol{\mu}_x \rightarrow \boldsymbol{\mu}_{\varepsilon x,0} = \mathbf{0}\right)$, then $\boldsymbol{\mu}_{\varepsilon x,k} = \mathrm{E}\left[\varepsilon_{x,k}\right] = \mathbf{0} \; \forall \, k$, this is, the expected value of $\hat{\mathbf{x}}_k$ is equal to $\boldsymbol{\mu}_x = \mathrm{E}\left[\mathbf{x}\right]$ for all t_k. This is regardless of the value of the gain matrix $\mathbf{K}_k$.

Let's follow a similar process to compute the covariance of the estimation error $\mathbf{C}_{xx,k} = \mathbf{P}_k$. To do so, note that the observation noise is independent from the estimation error $\left(\mathrm{E}\left[\widetilde{\mathbf{v}}_k \, \boldsymbol{\varepsilon}_{x,k-1}^T\right] = \mathbf{0}\right)$:

$$\mathbf{C}_{xx,k} = \mathbf{P}_k = \mathrm{E}\left[\boldsymbol{\varepsilon}_{x,k} \, \boldsymbol{\varepsilon}_{x,k}^T\right] - \boldsymbol{\mu}_{\varepsilon x,k} \, \boldsymbol{\mu}_{\varepsilon x,k}^T = (\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k) \; \mathbf{P}_{k-1} \; (\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k)^T + \mathbf{K}_k \, \widetilde{\mathbf{R}}_k \, \mathbf{K}_k^T \tag{102}$$

This expression guarantees that $\mathbf{C}_{xx,k} = \mathbf{P}_k$ is positive definite (as all covariance matrices) given that so are $\mathbf{P}_{k-1}$ and $\widetilde{\mathbf{R}}_k$. Let's use the minimization of the sum of the variances of the estimation errors as the criterion to obtain the gain matrix $\mathbf{K}_k$, and hence fill up (99) and (102) to obtain the estimation of the state vector as well as the covariance of its error. That way, the estimation error is not only zero mean but it is also consistently as close as possible to zero [6].

$$\begin{aligned} \mathbf{J}_k &= \mathrm{E}\left[\left(x_1 - \hat{x}_{1,k}\right)^2 + \ldots + \left(x_m - \hat{x}_{m,k}\right)^2\right] = \mathrm{E}\left[\varepsilon_{x1,k}^2 + \ldots + \varepsilon_{xm,k}^2\right] & \\ &= \mathrm{E}\left[\boldsymbol{\varepsilon}_{x,k}^T \, \boldsymbol{\varepsilon}_{x,k}\right] = \mathrm{E}\left[\mathrm{Tr}\left(\boldsymbol{\varepsilon}_{x,k} \, \boldsymbol{\varepsilon}_{x,k}^T\right)\right] = \mathrm{Tr} \, \mathbf{C}_{xx,k} = \mathrm{Tr} \, \mathbf{P}_k & (103) \\ \frac{\partial \mathbf{J}_k}{\partial \mathbf{K}_k} &= 2 \; (\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k) \; \mathbf{P}_{k-1} \; \left(-\mathbf{H}_k^T\right) + 2 \, \mathbf{K}_k \, \widetilde{\mathbf{R}}_k & (104) \end{aligned}$$

where Tr () stands for trace of a matrix, and some not so obvious matrix algebra properties have been employed. Setting the (104) derivative to zero provides the optimum gain matrix:

$$\mathbf{K}_k = \mathbf{P}_{k-1} \, \mathbf{H}_k^T \left(\mathbf{H}_k \, \mathbf{P}_{k-1} \, \mathbf{H}_k^T + \widetilde{\mathbf{R}}_k\right)^{-1} \tag{105}$$

3.3.3 The Extended Kalman Filter

Provided with discrete time and linear state dynamics (79) and observations (93), the goal of state estimation is to obtain the best possible estimate of the state vector $\mathbf{x}_k = \mathbf{x}\left(t_k\right)$ based on the knowledge of the system provided by the state dynamics and the availability of observations [6]. At a given time $t_k = k \, \Delta t$, the *a priori estimation* $\hat{\mathbf{x}}_k^-$ is defined as the estimation of $\mathbf{x}_k$, this is, the estimation of the state vector at time t_k, making use of all measurements taken before t_k but not including those at t_k. The *a posteriori estimation* $\hat{\mathbf{x}}_k^+$ is defined as the estimation of $\mathbf{x}_k$ that makes use of all measurements up and including t_k. In the same way, it is possible to define the *a priori* and *a posteriori covariances* of the estimation error $\mathbf{P}_k^-$ and $\mathbf{P}_k^+$:

$$\begin{aligned} \mathbf{P}_k^- &= \mathrm{E}\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right) \; \left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right)^T\right] - \mathrm{E}\left[\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right] \, \mathrm{E}\left[\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right]^T & (106) \\ \mathbf{P}_k^+ &= \mathrm{E}\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right) \; \left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right)^T\right] - \mathrm{E}\left[\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right] \, \mathrm{E}\left[\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right]^T & (107) \end{aligned}$$

The process starts with an initial estimation of the state vector $\hat{\mathbf{x}}_0^+$ before any measurements are available (they start at $k = 1$). Since no measurements are available, it is reasonable to form $\hat{\mathbf{x}}_0^+$ as the expected value of the initial state $\mathbf{x}_0$ [6]:

$$\hat{\mathbf{x}}_0^+ = \boldsymbol{\mu}_{x,0} = \mathrm{E}\left[\mathbf{x}_0\right] = \mathrm{E}\big[\mathbf{x}\left(t_0\right)\big] \tag{108}$$

The covariance of the initial estimation error $\mathbf{P}_0^+$ is also required, representing the uncertainty in the initial estimation $\hat{\mathbf{x}}_0^+$ [15]:

$$\mathbf{P}_0^+ = \mathrm{E}\left[\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right)\,\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right)^{\mathrm{T}}\right] - \mathrm{E}\left[\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right]\,\mathrm{E}\left[\mathbf{x}_0 - \hat{\mathbf{x}}_0^+\right]^{\mathrm{T}} = \mathrm{E}\left[\left(\mathbf{x}_0 - \boldsymbol{\mu}_{\mathrm{x},0}\right)\,\left(\mathbf{x}_0 - \boldsymbol{\mu}_{\mathrm{x},0}\right)^{\mathrm{T}}\right] \tag{109}$$

Time Update and Measurement Update Equations

The next step is to propagate the state estimation without the use of any observations from $\hat{\mathbf{x}}_0^+$ to $\hat{\mathbf{x}}_1^-$, with the objective of obtaining an estimation that coincides with the state vector mean, this is, $\hat{\mathbf{x}}_1^- = \boldsymbol{\mu}_{\mathrm{x},1} = \mathrm{E}\left[\mathbf{x}_1\right]$. Recalling the evolution of the state vector expected value provided by (86), and extending the same reasoning to all steps, it makes sense intuitively to propagate the state estimate the same way that the mean of the state propagates [6]. Hence, the time propagation for the state estimate results in:

$$\hat{\mathbf{x}}_{\mathrm{k}}^- = \mathbf{F}_{\mathrm{k}-1}\,\hat{\mathbf{x}}_{\mathrm{k}-1}^+ + \mathbf{G}_{\mathrm{k}-1}\,\widetilde{\mathbf{u}}_{\mathrm{k}-1} \tag{110}$$

A similar reasoning is employed for the propagation of the covariance of the estimation error in the absence of observations. Recalling the evolution of the state vector covariance provided by (87) and extending the same reasoning to all steps, the time propagation of the covariance results in [6]:

$$\mathbf{P}_{\mathrm{k}}^- = \mathbf{F}_{\mathrm{k}-1}\,\mathbf{P}_{\mathrm{k}-1}^+\,\mathbf{F}_{\mathrm{k}-1}^{\mathrm{T}} + \widetilde{\mathbf{Q}}_{\mathrm{d,k}-1} \tag{111}$$

The above equations are called the *time update equations*. Once the a priori estimation and error covariance have been computed, it is possible to update them with the information contained in the observation. This is done with the expressions derived in section 3.3.2, replacing $\hat{\mathbf{x}}_{\mathrm{k}-1}$ with $\hat{\mathbf{x}}_{\mathrm{k}}^-$, $\hat{\mathbf{x}}_{\mathrm{k}}$ with $\hat{\mathbf{x}}_{\mathrm{k}}^+$, $\mathbf{P}_{\mathrm{k}-1}$ with $\mathbf{P}_{\mathrm{k}}^-$, and $\mathbf{P}_{\mathrm{k}}$ with $\mathbf{P}_{\mathrm{k}}^+$ [6]. These are called the *measurement update equations*:

$$\begin{aligned}
\mathbf{K}_{\mathrm{k}} &= \mathbf{P}_{\mathrm{k}}^-\,\mathbf{H}_{\mathrm{k}}^{\mathrm{T}}\left(\mathbf{H}_{\mathrm{k}}\,\mathbf{P}_{\mathrm{k}}^-\,\mathbf{H}_{\mathrm{k}}^{\mathrm{T}} + \widetilde{\mathbf{R}}_{\mathrm{k}}\right)^{-1} && (112)\\
\hat{\mathbf{x}}_{\mathrm{k}}^+ &= \hat{\mathbf{x}}_{\mathrm{k}}^- + \mathbf{K}_{\mathrm{k}}\,\mathbf{r}_{\mathrm{k}} = \hat{\mathbf{x}}_{\mathrm{k}}^- + \mathbf{K}_{\mathrm{k}}\,\left(\mathbf{y}_{\mathrm{k}} - \mathbf{H}_{\mathrm{k}}\,\hat{\mathbf{x}}_{\mathrm{k}}^- - \mathbf{z}_{\mathrm{k}}\right) && (113)\\
\mathbf{P}_{\mathrm{k}}^+ &= \left(\mathbf{I} - \mathbf{K}_{\mathrm{k}}\,\mathbf{H}_{\mathrm{k}}\right)\,\mathbf{P}_{\mathrm{k}}^-\,\left(\mathbf{I} - \mathbf{K}_{\mathrm{k}}\,\mathbf{H}_{\mathrm{k}}\right)^{\mathrm{T}} + \mathbf{K}_{\mathrm{k}}\,\widetilde{\mathbf{R}}_{\mathrm{k}}\,\mathbf{K}_{\mathrm{k}}^{\mathrm{T}} && (114)
\end{aligned}$$

Introduction of the Nominal Trajectory

The time and measurement update equations developed in the previous section provide the means to compute the variation with time of the estimated state vector $\left(\hat{\mathbf{x}}_{\mathrm{k}}^-,\,\hat{\mathbf{x}}_{\mathrm{k}}^+\right)$ as well as that of the covariance of the estimation errors $\left(\mathbf{P}_{\mathrm{k}}^-,\,\mathbf{P}_{\mathrm{k}}^+\right)$. However, to do so, it is necessary to define what is the nominal point $\mathbf{x}_{\mathrm{Nk}} = \mathbf{x}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right),\, \mathbf{w}_{\mathrm{Nk}} = \mathbf{w}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right) = \mathbf{0},\, \mathbf{v}_{\mathrm{Nk}} = \mathbf{v}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right) = \mathbf{0}$ around which the dynamics system is linearized in section 3.3.1 and the observations in section 3.3.2.

The *extended Kalman filter* (`EKF`) provides a solution to this problem that is simple but not too intuitive. The `EKF` considers its own a priori state estimate as the nominal trajectory, this is, the nonlinear state system and observations are linearized around the `EKF` estimate, and simultaneously that same estimate depends on the linearized system [6]:

$$\mathbf{x}_{\mathrm{Nk}} = \mathbf{x}_{\mathrm{N}}\left(\mathrm{t}_{\mathrm{k}}\right) = \hat{\mathbf{x}}_{\mathrm{k}}^- \tag{115}$$

This assumption can be introduced into the expressions for the observations output matrix $\mathbf{H}_{\mathrm{k}}$ and observations input vector $\mathbf{z}_{\mathrm{k}}$ provided by (94) and (95), with the results introduced into the state estimation measurement update equation (113):

$$\begin{aligned}
\hat{\mathbf{x}}_{\mathrm{k}}^+ &= \hat{\mathbf{x}}_{\mathrm{k}}^- + \mathbf{K}_{\mathrm{k}}\left(\mathbf{y}_{\mathrm{k}} - \frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\left(\hat{\mathbf{x}}_{\mathrm{k}}^-,\,\mathbf{0},\,\mathrm{t}_{\mathrm{k}}\right)\,\hat{\mathbf{x}}_{\mathrm{k}}^- - \mathrm{h}\left(\hat{\mathbf{x}}_{\mathrm{k}}^-,\,\mathbf{0},\,\mathrm{t}_{\mathrm{k}}\right) + \frac{\partial \mathrm{h}}{\partial \mathbf{x}_{\mathrm{k}}}\left(\hat{\mathbf{x}}_{\mathrm{k}}^-,\,\mathbf{0},\,\mathrm{t}_{\mathrm{k}}\right)\,\hat{\mathbf{x}}_{\mathrm{k}}^-\right)\\
&= \hat{\mathbf{x}}_{\mathrm{k}}^- + \mathbf{K}_{\mathrm{k}}\left[\mathbf{y}_{\mathrm{k}} - \mathrm{h}\left(\hat{\mathbf{x}}_{\mathrm{k}}^-,\,\mathbf{0},\,\mathrm{t}_{\mathrm{k}}\right)\right] && (116)
\end{aligned}$$

[15] If the initial state is known with exactitude, use $\mathbf{P}_0^+ = 0$. Otherwise, use higher values the less confidence the user has in the accuracy of $\hat{\mathbf{x}}_0^+$.

Note that, in contrast with (113), it is no longer necessary to compute the observations input vector $\mathbf{z}_k$. In order to diminish the state system linearization errors described in section 3.3.1, it is also possible to replace the state estimate time update equation (110) with a zeroth order forward integration of the continuous time state system (63), with the time derivative evaluated at $\hat{\mathbf{x}}_{k-1}^{+}$:

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_{k-1}^+ + \Delta t \cdot f\left(\hat{\mathbf{x}}_{k-1}^+, \mathbf{u}_{k-1}, \mathbf{0}, t_{k-1}\right) \tag{117}$$

An extra benefit of this approach compared with (110) is that it is no longer necessary to perform the expensive computations required to evaluate $\mathbf{G}_{k-1}$ (82).

EKF Summary

Given a continuous time nonlinear state system (63) with process noise provided by (64) and (65), together with a series of discrete time nonlinear observations (88) with measurement noise given by (89) and (90), and considering no correlation between both noises (91), it is possible to compute estimations of the state at the same time points at which the observations are provided, in such a way that their errors (difference with respect to the true state) are zero mean and have a covariance that is also computed by means of the following equations:

$$\hat{\mathbf{x}}_0^+ = \boldsymbol{\mu}_{x,0} = E\left[\mathbf{x}_0\right] \tag{118}$$

$$\mathbf{P}_0^+ = E\left[\left(\mathbf{x}_0 - \boldsymbol{\mu}_{x,0}\right)\left(\mathbf{x}_0 - \boldsymbol{\mu}_{x,0}\right)^T\right] \tag{119}$$

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_{k-1}^+ + \Delta t \cdot f\left(\hat{\mathbf{x}}_{k-1}^+, \mathbf{u}_{k-1}, \mathbf{0}, t_{k-1}\right) \tag{120}$$

$$\mathbf{P}_k^- = \mathbf{F}_{k-1}\,\mathbf{P}_{k-1}^+\,\mathbf{F}_{k-1}^T + \widetilde{\mathbf{Q}}_{d,k-1} \tag{121}$$

$$\mathbf{K}_k = \mathbf{P}_k^-\,\mathbf{H}_k^T\left(\mathbf{H}_k\,\mathbf{P}_k^-\,\mathbf{H}_k^T + \widetilde{\mathbf{R}}_k\right)^{-1} \tag{122}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left[\mathbf{y}_k - h\left(\hat{\mathbf{x}}_k^-, \mathbf{0}, t_k\right)\right] \tag{123}$$

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\mathbf{P}_k^-\left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)^T + \mathbf{K}_k\,\widetilde{\mathbf{R}}_k\,\mathbf{K}_k^T \tag{124}$$

In the discussion that follows, $\hat{\mathbf{x}}_k$ is employed to refer to both the a priori and a posteriori state vector estimations $\left(\hat{\mathbf{x}}_k^-, \hat{\mathbf{x}}_k^+\right)$, and $\boldsymbol{\varepsilon}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ for the state estimation errors. The above equations show that $\hat{\mathbf{x}}_k$ is a linear combination of a random vector $\mathbf{x}_0^+$ plus a series of random processes $\widetilde{\mathbf{u}}_k$, so it is itself a random process, and so is $\boldsymbol{\varepsilon}_k$.

Let's leave aside for the time being the linearization and discretization errors of sections 3.3.1 and 3.3.1, and focus on a problem composed by a discrete time linear state system (79) and discrete time linear observations (93). Provided with any user defined positive definite weighting matrix $\mathbf{S}_k$, it can be proved that the solution provided verifies (125), this is, results in a state estimation that always minimizes the weighted sum of squared estimation errors [6], as long as the process and observations noises are Gaussian zero mean uncorrelated white noise processes. If they are not Gaussian, then $\hat{\mathbf{x}}_k$ provides the best linear (in the sense of the previous paragraph) solution to the (125) minimization, although there may be a better nonlinear solution.

$$\hat{\mathbf{x}}_k = \arg\min E\left[\boldsymbol{\varepsilon}_k^T\,\mathbf{S}_k\,\boldsymbol{\varepsilon}_k\right] \tag{125}$$

The errors induced by the discretization of the linear continuous time dynamics system in section 3.3.1 generally do not result in significant errors as modern systems are capable of running the estimation algorithms at elevated frequencies. The system matrix $\mathbf{A}(t)$ and input vector $\widetilde{\mathbf{u}}(t)$ in the continuous time system (67) generally do not vary much during the integration interval, and hence the discretization errors are small. The linearization errors of sections 3.3.1 and 3.3.2 are a different story, and can induce the `EKF` to provide unreliable estimates or even to diverge in case the nonlinearities are severe [6].

4 Introduction to Lie Algebra

This section begins with some basic abstract and linear algebra concepts in sections 4.1 and 4.2, and then introduces Lie algebra in section 4.3, followed by the derivation of some useful Lie jacobians in section 4.4. Its application to the discrete integration of states is discussed in section 4.5, to gradient descent optimization in section 4.6, and to state estimation in section 4.7. The contents of this section are generic to any Lie group without making further mention to rigid bodies. It is only in sections 5 and 6 where the Lie theory concepts are applied first to rotational motion and then to the more generic rigid body motion.

4.1 Algebraic Structures, Maps, and Metric Spaces

In algebra, a *set* is a well defined collection of objects, named elements, while an *operation* $*$ is a uniquely defined rule that assigns to each ordered pair of elements exactly a third element $\{* : \mathbb{A} \times \mathbb{B} \to \mathbb{C} \mid \mathrm{a} * \mathrm{b} = \mathrm{c} \in \mathbb{C}, \forall\, \mathrm{a} \in \mathbb{A}, \forall\, \mathrm{b} \in \mathbb{B}\}$ [15]. Although an operation may involve up to three different sets $(\mathbb{A}, \mathbb{B}, \mathbb{C})$, often two or even the three of them coincide. A set $\mathbb{A}$ is a *subset* of a set $\mathbb{B}$ if all elements of $\mathbb{A}$ are also elements of $\mathbb{B}$. An *algebraic structure* is a combination of a set and one or multiple operations that complies with certain axioms.

A set $\mathbb{A}$ has *group* structure under operation $\{* : \mathbb{A} \times \mathbb{A} \to \mathbb{A}\}$ if it complies with the following four axioms $\forall\, \mathrm{a}, \mathrm{b}, \mathrm{c} \in \mathbb{A}$ [15]:

1. Closure: $\mathrm{a} * \mathrm{b} \in \mathbb{A}$
2. Associativity: $(\mathrm{a} * \mathrm{b}) * \mathrm{c} = \mathrm{a} * (\mathrm{b} * \mathrm{c})$
3. Identity: $\exists\, \mathrm{e} \in \mathbb{A} \mid \mathrm{e} * \mathrm{a} = \mathrm{a} * \mathrm{e} = \mathrm{a}$
4. Inverse: $\exists\, \mathrm{f} \in \mathbb{A} \mid \mathrm{f} * \mathrm{a} = \mathrm{a} * \mathrm{f} = \mathrm{e}$

An *abelian group* is that which in addition also complies with commutativity $\{\mathrm{a} * \mathrm{b} = \mathrm{b} * \mathrm{a}\}$. A set $\mathbb{A}$ has *ring* structure under two operations, usually named addition $\{+ : \mathbb{A} \times \mathbb{A} \to \mathbb{A}\}$ and multiplication $\{\cdot : \mathbb{A} \times \mathbb{A} \to \mathbb{A}\}$, if, in addition to being an abelian group under addition, complies with the following four axioms $\forall\, \mathrm{a}, \mathrm{b}, \mathrm{c} \in \mathbb{A}$ [15]:

1. Closure of $\cdot$: $\mathrm{a} \cdot \mathrm{b} \in \mathbb{A}$
2. Associativity of $\cdot$: $(\mathrm{a} \cdot \mathrm{b}) \cdot \mathrm{c} = \mathrm{a} \cdot (\mathrm{b} \cdot \mathrm{c})$
3. Distributivity of $\cdot$ with respect to $+$: $\mathrm{a} \cdot (\mathrm{b} + \mathrm{c}) = \mathrm{a} \cdot \mathrm{b} + \mathrm{a} \cdot \mathrm{c}, \ (\mathrm{a} + \mathrm{b}) \cdot \mathrm{c} = \mathrm{a} \cdot \mathrm{c} + \mathrm{b} \cdot \mathrm{c}$
4. Identity of $\cdot$: $\exists\, 1 \in \mathbb{A} \mid 1 \cdot \mathrm{a} = \mathrm{a} \cdot 1 = \mathrm{a}$

An *abelian ring* is that which in addition also complies with commutativity over multiplication $\{\mathrm{a} \cdot \mathrm{b} = \mathrm{b} \cdot \mathrm{a}\}$. Note that by convention, the identity and inverse of addition are denoted 0 and $-\mathrm{a}$, respectively, while those of multiplication are denoted 1 and a^{-1}. A set $\mathbb{A}$ has *field* structure under operations $+$ and $\cdot$ if $\mathbb{A}$ is an abelian group under $+$ and $\mathbb{A} - \{0\}$ (the set $\mathbb{A}$ without the additive identity 0) is an abelian group under $\cdot$ [15]. In an *ordered field*, the implementation of the addition and multiplication operations enables determining if one element is greater, equal, or lower than a second element. The set of real numbers $\mathbb{R}$ endowed with the operations of addition $+$ and multiplication $\cdot$ forms an ordered field, known as the field of real numbers $\langle \mathbb{R}, +, \cdot \rangle$, nearly always abbreviated to simply $\mathbb{R}$.

A *topological space* is an ordered pair $(\mathbb{A}, \tau)$, where $\mathbb{A}$ is a set and τ is a collection of subsets of $\mathbb{A}$, satisfying the following axioms [16]:

1. The empty set and $\mathbb{A}$ itself belong to τ.
2. Any arbitrary (finite or infinite) union of members of τ still belongs to τ.
3. The intersection of any finite number of members of τ still belongs to τ.

The elements of τ are called open sets and the collection τ is called a topology on $\mathbb{A}$. Topological spaces comprise the most general notion of a mathematical space; all other spaces defined below are specializations with extra structure or constraints.

A *vector space* (*linear space*) over a field $\mathbb{F}$ is a set $\mathbb{V}$ together with two operations, addition $\{+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}\}$ and scalar multiplication $\{\cdot : \mathbb{F} \times \mathbb{V} \to \mathbb{V}\}$ that, in addition of $\mathbb{V}$ being an abelian group under $+$, satisfies the following axioms $\forall\ \mathrm{u}, \mathrm{v} \in \mathbb{V}$ and $\forall\ \mathrm{a}, \mathrm{b} \in \mathbb{F}$ [17, 18]. Elements of $\mathbb{F}$ are called scalars, while those of $\mathbb{V}$ vectors.

1. Closure of $\cdot$: $\mathrm{a} \cdot \mathrm{u} \in \mathbb{V}$
2. Compatibility of $\cdot$ with field $\cdot$: $\mathrm{a} \cdot (\mathrm{b} \cdot \mathrm{v}) = (\mathrm{a} \cdot \mathrm{b}) \cdot \mathrm{v}$
3. Identity of $\cdot$: $1 \cdot \mathrm{v} = \mathrm{v}$, where 1 denotes the field $\cdot$ identity.
4. Distributivity of $\cdot$ with respect to $+$: $\mathrm{a} \cdot (\mathrm{u} + \mathrm{v}) = \mathrm{a} \cdot \mathrm{u} + \mathrm{a} \cdot \mathrm{v}$
5. Distributivity of $\cdot$ with respect to field $+$: $(\mathrm{a} + \mathrm{b}) \cdot \mathrm{v} = \mathrm{a} \cdot \mathrm{v} + \mathrm{b} \cdot \mathrm{v}$

A *map* or *morphism* is a rule that to every element in a set $\mathbb{A}$ assigns a unique element in a different set $\mathbb{B}$ $\{\mathrm{f} : \mathbb{A} \to \mathbb{B} \mid \mathrm{f}(\mathrm{a}) = \mathrm{b} \in \mathbb{B}, \forall\ \mathrm{a} \in \mathbb{A}\}$ [15]. A map is *injective* if each element in $\mathbb{B}$ is the image or map output of no more than one element of $\mathbb{A}$, *surjective* if each element in $\mathbb{B}$ is the image of at least one element of $\mathbb{A}$, and *bijective* is the map is simultaneously injective and surjective.

A *homomorphism* is a structure preserving map between two algebraic structures of the same type (groups, rings, fields, vector spaces, etc.) [15]. Note that neither the sets nor the operations of the structures need to coincide, and that a homomorphism preserves every operation contained in the algebraic structures. In the case of groups, considering a homomorphism $\{\mathrm{f} : \langle \mathbb{A},\ + \rangle \to \langle \mathbb{B},\ \cdot \rangle\}$, it complies with all group axioms $\forall\ \mathrm{a}, \mathrm{b}, \mathrm{c} \in \mathbb{A}$:

1. Closure: $\mathrm{f}(\mathrm{a} + \mathrm{b}) = \mathrm{f}(\mathrm{a}) \cdot \mathrm{f}(\mathrm{b})$
2. Associativity: $\mathrm{f}\big((\mathrm{a} + \mathrm{b}) + \mathrm{c}\big) = \mathrm{f}\big(\mathrm{a} + (\mathrm{b} + \mathrm{c})\big) = \big(\mathrm{f}(\mathrm{a}) \cdot \mathrm{f}(\mathrm{b})\big) \cdot \mathrm{f}(\mathrm{c}) = \mathrm{f}(\mathrm{a}) \cdot \big(\mathrm{f}(\mathrm{b}) \cdot \mathrm{f}(\mathrm{c})\big)$
3. Identity: $\mathrm{f}(0) = 1 \to \mathrm{f}(\mathrm{a}) = \mathrm{f}(\mathrm{a} + 0) = \mathrm{f}(0 + \mathrm{a}) = \mathrm{f}(\mathrm{a}) \cdot \mathrm{f}(0) = \mathrm{f}(0) \cdot \mathrm{f}(\mathrm{a}) = \mathrm{f}(\mathrm{a}) \cdot 1 = 1 \cdot \mathrm{f}(\mathrm{a})$
4. Inverse: $\mathrm{f}(-\mathrm{a}) = \mathrm{f}^{-1}(\mathrm{a}) \to \mathrm{f}(0) = \mathrm{f}(\mathrm{a} - \mathrm{a}) = \mathrm{f}(\mathrm{a}) \cdot \mathrm{f}(-\mathrm{a}) = \mathrm{f}(\mathrm{a}) \cdot \mathrm{f}^{-1}(\mathrm{a}) = \mathrm{f}^{-1}(\mathrm{a}) \cdot \mathrm{f}(\mathrm{a}) = 1$

Homomorphisms for other algebraic structures are defined similarly. A bijective homomorphism is known as an *isomorphism* [15]. A *metric* is an operation between two elements of the same set onto a field $\{\mathrm{d} : \mathbb{V} \times \mathbb{V} \to \mathbb{F}\}$. It defines the concept of distance between any two members of the set and complies with the following axioms $\forall\ \mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{V}$:

1. Identity of indiscernibles: $\mathrm{d}(\mathrm{u}, \mathrm{v}) = 0 \Leftrightarrow \mathrm{u} = \mathrm{v}$
2. Symmetry: $\mathrm{d}(\mathrm{u}, \mathrm{v}) = \mathrm{d}(\mathrm{v}, \mathrm{u})$
3. Subadditivity: $\mathrm{d}(\mathrm{u}, \mathrm{w}) \leq \mathrm{d}(\mathrm{u}, \mathrm{v}) + \mathrm{d}(\mathrm{v}, \mathrm{w})$

Based on these three axioms, it is straightforward to prove that $\mathrm{d}(\mathrm{u}, \mathrm{v}) \geq 0\ \forall\ \mathrm{u}, \mathrm{v} \in \mathbb{V}$. The most common metric is the *inner product* $\{\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{F}\}$, usually associated to a vector space, which satisfies the following three axioms $\forall\ \mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{V}$ and $\forall\ \mathrm{a}, \mathrm{b} \in \mathbb{F}$ [17]:

1. Commutativity: $\langle \mathrm{u}, \mathrm{v} \rangle = \langle \mathrm{v}, \mathrm{u} \rangle$
2. Linearity with respect to $+$ and $\cdot$: $\langle \mathrm{u}, \mathrm{a} \cdot \mathrm{v} + \mathrm{b} \cdot \mathrm{w} \rangle = \mathrm{a} \cdot \langle \mathrm{u}, \mathrm{v} \rangle + \mathrm{b} \cdot \langle \mathrm{u}, \mathrm{w} \rangle$
3. Positive definiteness: $\langle \mathrm{v}, \mathrm{v} \rangle \geq 0$ and $\langle \mathrm{v}, \mathrm{v} \rangle = 0 \Leftrightarrow \mathrm{v} = 0$

A *metric space* is a combination of a set with a metric on that same set [19], while an *inner product space* restricts the definition to the case of a vector space $\mathbb{V}$ over a field $\mathbb{F}$ endowed with an inner product metric over the same field $\mathbb{F}$ [20]. An *Euclidean space* is a finite-dimensional inner product space over the field of the real numbers $\mathbb{R}$ [21].

A *group action* on a space is a group homomorphism of a given group $\langle \mathbb{A}, * \rangle$ into the group of transformations of the space $\{\mathrm{g}() : \mathbb{A} \times \mathbb{V} \to \mathbb{V} \mid \mathrm{g}_{\mathrm{a}}(\mathrm{u}) = \mathrm{v} \in \mathbb{V}, \forall\ \mathrm{a} \in \mathbb{A}, \forall\ \mathrm{u} \in \mathbb{V}\}$, and needs to verify two axioms $\forall\ \mathrm{a}, \mathrm{b} \in \mathbb{A}, \forall\ \mathrm{u} \in \mathbb{V}$:

1. Identity: $g_e(u) = u$, where e is the identity of $\mathbb{G}$.
2. Compatibility: $g_{a*b}(u) = g_a\big(g_b(u)\big)$

Returning to the case of inner product spaces, two vectors are *orthogonal* if their inner product is zero, while the length or *norm* of a vector is $\|v\| = \sqrt{\langle v, v\rangle}$. A unit vector is that whose norm is one. An inner product space can also be endowed with an additional operation, the *cross product* $\{\times : \mathbb{V} \times \mathbb{V} \to \mathbb{V}\}$, which complies with the following axioms [17]:

1. Anti commutativity: $u \times v = -v \times u$
2. Compatibility with $\cdot$: $(a \cdot u) \times v = u \times (a \cdot v) = a \cdot (u \times v)$
3. Distributivity with $+$: $(u + v) \times w = (u \times w) + (v \times w)$

It can also be quickly derived that $\langle u \times v, u\rangle = \langle u \times v, v\rangle = 0$.

4.2 Points, Vectors, and Axes

Note that in the abstract discussion above it has not yet been defined what a vector is. This section focuses on the three-dimensional Euclidean space $\mathbb{E}^3$ [22], which can be represented by a cartesian frame, where every *point* $\mathbf{p} \in \mathbb{E}^3$ can be identified by its three coordinates $\mathbf{p} = [p_1, p_2, p_3]^T \in \mathbb{R}^3$. A *vector* in $\mathbb{E}^3$ is defined by a pair of points $\mathbf{p}, \mathbf{q} \in \mathbb{E}^3$ with a directed arrow connecting $\mathbf{p}$ to $\mathbf{q}$, where the vector $\mathbf{v} = [v_1, v_2, v_3]^T \in \mathbb{R}^3$ is a triplet of numbers, each one being the difference between the corresponding coordinates of the two points $\mathbf{q}$ and $\mathbf{p}$ ($\mathbf{v} = \mathbf{q} - \mathbf{p} \in \mathbb{R}^3$). Although they share notation, points and vectors are different geometric objects. A *free vector* is one that does not depend on its starting or base point.

The set of all free vectors in $\mathbb{R}^3$ form an inner product space with cross product over the field of real numbers $\mathbb{R}$, with both products defined as follows by making use of matrix notation:

$$\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 \tag{126}$$

$$\mathbf{u} \times \mathbf{v} = \widehat{\mathbf{u}}\, \mathbf{v} = \begin{bmatrix} 0 & -u_3 & +u_2 \\ +u_3 & 0 & -u_1 \\ -u_2 & +u_1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} = -\mathbf{v} \times \mathbf{u} = -\widehat{\mathbf{v}}\, \mathbf{u} \tag{127}$$

where $\mathbf{v}^T$ is the transpose of $\mathbf{v}$ and $\widehat{\mathbf{v}}$ its skew-symmetric form[16]. The inner product or euclidean metric can measure distances and angles, while the cross product defines orientation.

Any vector $\mathbf{v}$ in $\mathbb{R}^3$ can be written as $\mathbf{v} = v_1\, \mathbf{e}_1 + v_2\, \mathbf{e}_2 + v_3\, \mathbf{e}_3$, where $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the three linearly independent basis vectors and v_1, v_2, v_3 the coordinates or components of $\mathbf{v}$ with respect to that basis [23]. The *basis* is called orthogonal if $\mathbf{e}_i^T\, \mathbf{e}_j = 0$ when $i \neq j$, orthonormal if additionally $\mathbf{e}_i^T\, \mathbf{e}_j = 1$ when $i = j$, and right handed if additionally $\epsilon_{ijk} = \mathbf{e}_i^T\, \widehat{\mathbf{e}}_j\, \mathbf{e}_k$ is 1 for ϵ_{123}, ϵ_{231}, and ϵ_{312}, -1 for ϵ_{132}, ϵ_{213}, and ϵ_{321}, and 0 in all other cases [17].

An *axis* or *line* is defined by its direction $\mathbf{n}$ (provided by a free vector) and a point $\mathbf{p}$ that it passes through. Its coordinates are $(\mathbf{n}, \mathbf{m}) \in \mathbb{R}^6$, where $\mathbf{m} = \widehat{\mathbf{p}}\, \mathbf{n}$ is called the *moment* of the line. The coordinates $(\mathbf{n}, \mathbf{m})$ are independent of $\mathbf{p}$. The moment $\mathbf{m}$ is normal to the plane through the line and the origin with norm equal to the distance from the line to the origin. The point belonging to the line that is closest to the origin responds to $\mathbf{p}_\perp = \widehat{\mathbf{n}}\, \mathbf{m}$. A line has four degrees of freedom and hence two redundancies, provided by $\mathbf{n}$ being a direction and hence a unit vector ($\|\mathbf{n}\| = 1$) and $\mathbf{n}$ being orthogonal to $\mathbf{m}$ by definition ($\mathbf{n}^T\, \mathbf{m} = 0$) [24].

It is important to remark that although this is the formal definition of an axis, it can also be informally considered that an axis is synonymous with just a direction $\mathbf{n}$ with two degrees of freedom ($\|\mathbf{n}\| = 1$), passing through the origin ($\mathbf{p} = \mathbf{0} \to \mathbf{m} = \mathbf{0}$). The reason is that in most occasions it is convenient to consider that a rigid body rotates about the origin of the reference frame representing it.

[16] An skew-symmetric matrix is one whose negative equals its transpose.

4.3 Lie Groups and Lie Algebras

A *manifold* is a topological space that locally resembles Euclidean space near each element, so each element of an m dimensional manifold has a neighborhood that is homeomorphic[17] to the m dimensional Euclidean space [25]. Manifolds, which are embedded in spaces of higher dimension, are curved, smooth (hyper) surfaces with no edges or spikes; they are defined by the constraints imposed on the state [2], this is, the state vector is restricted to moving within the manifold.

A *Lie group* $\langle\mathcal{G},\circ\rangle$ is a smooth manifold whose elements satisfy the group axioms. They combine the local properties of smooth manifolds, enabling the use of calculus, with the global properties of groups, allowing the nonlinear composition of distant objects [2]. Elements of $\mathcal{G}$ are denoted with $\mathcal{X}$, the identity with $\mathcal{E}$, and the inverse with $\mathcal{X}^{-1}$. As in any other group, Lie groups are capable of transforming elements of other sets by means of their actions. In this sense, the group operation $\{\circ : \mathcal{G}\times\mathcal{G}\rightarrow\mathcal{G}\}$, generally called *composition*, can be considered as an action of the group on itself.

If $\mathcal{X}\left(\mathrm{t}\right)$ is an element or point of the Lie group moving on the manifold, its derivative with time belongs to the space tangent to $\mathcal{G}$ at $\mathcal{X}$, denoted by $\mathrm{T}_{\mathcal{X}}\mathcal{G}$. There exists a unique tangent space at each point, but the structure of such tangent spaces is the same everywhere [2]. The *tangent space* at a point is a real vector space of the same dimension as the manifold that intuitively contains all the possible directions in which one can tangentially pass through the point.

The *Lie algebra* $\mathfrak{m}$ is defined as the tangent space at the identity $\{\mathfrak{m}=\mathrm{T}_{\mathcal{E}}\mathcal{G}\}$, and it is a vector space whose elements can be identified with vectors in $\mathbb{R}^{\mathrm{m}}$, with m being the number of degrees of freedom of the Lie group $\mathcal{G}$ [2]. Elements of the tangent space are usually denoted $\mathbf{v}$ when referring to velocities and $\boldsymbol{\tau}=\mathbf{v}\cdot\mathrm{t}$ for more general elements. Lie algebras can be defined at any manifold point $\mathcal{X}$, establishing local coordinates for $\mathrm{T}_{\mathcal{X}}\mathcal{G}$, and its elements are denoted by the $<\cdot^{\wedge}>$ symbol, such as $\mathbf{v}^{\mathcal{X}\wedge}\in\mathrm{T}_{\mathcal{X}}\mathcal{G}$ or $\boldsymbol{\tau}^{\mathcal{E}\wedge}\in\mathrm{T}_{\mathcal{E}}\mathcal{G}$.

As all tangent spaces or Lie algebras have the same structure no matter the position of the element $\mathcal{X}$ within the Lie group $\mathcal{G}$, it can be considered with no loss of generality[18] that all group actions at $\mathcal{X}\in\mathcal{G}$, noted as $\mathrm{g}_{\mathcal{X}}()$, transform elements viewed in the *local* or *body* frame represented by $\mathcal{X}$ into the *global* or *space* frame represented by $\mathcal{E}$ [2]. The opposite is true for the inverse operations noted as $\mathrm{g}_{\mathcal{X}}^{-1}()=\mathrm{g}_{\mathcal{X}^{-1}}()$. When the group action is the composition $\circ$ itself, elements to the right of $\mathcal{X}$ belong to the body frame, while those on the left are viewed on the global frame.

4.3.1 Lie Algebra Velocities, Hat and Vee Operators

The structure of the Lie algebra can be obtained by time derivating the group inverse constraint [2]. Every Lie group employed in this article (refer to sections 5 and 6) has a $\circ$ composition operator realized by some type of multiplication. If this is the case, the group inverse constraint responds to $\mathcal{X}\circ\mathcal{X}^{-1}=\mathcal{X}^{-1}\circ\mathcal{X}=\mathcal{E}$, and its derivation with time leads to the Lie algebra *velocities* viewed in either the body or local frames[19]:

$$\dot{\mathcal{X}}\circ\mathcal{X}^{-1}+\mathcal{X}\circ\dot{\mathcal{X}^{-1}}=0 \quad\rightarrow\quad \mathbf{v}^{\mathcal{E}\wedge}=\dot{\mathcal{X}}\circ\mathcal{X}^{-1}=-\mathcal{X}\circ\dot{\mathcal{X}^{-1}} \tag{128}$$

$$\mathcal{X}^{-1}\circ\dot{\mathcal{X}}+\dot{\mathcal{X}^{-1}}\circ\mathcal{X}=0 \quad\rightarrow\quad \mathbf{v}^{\mathcal{X}\wedge}=\mathcal{X}^{-1}\circ\dot{\mathcal{X}}=-\dot{\mathcal{X}^{-1}}\circ\mathcal{X} \tag{129}$$

The $\boldsymbol{\tau}^{\wedge}$ or $\mathbf{v}^{\wedge}$ elements of the Lie algebra hence do not have trivial structures but can always be expressed as linear combinations of some base elements $\mathbf{e}_{\mathrm{i}}$, which are called the *generators* of $\mathfrak{m}$ [2]. It is generally more convenient to manipulate them as vectors $\boldsymbol{\tau}\in\mathbb{R}^{\mathrm{m}}$, as they can then be grouped together in larger state vectors and operated by means of linear algebra. The isomorphisms that linearly convert between them are called *hat* $\{\cdot^{\wedge}:\mathbb{R}^{\mathrm{m}}\rightarrow\mathfrak{m}\mid\boldsymbol{\tau}\rightarrow\boldsymbol{\tau}^{\wedge}\}$ and *vee* $\left\{\cdot^{\vee}:\mathfrak{m}\rightarrow\mathbb{R}^{\mathrm{m}}\mid\left(\boldsymbol{\tau}^{\wedge}\right)^{\vee}\rightarrow\boldsymbol{\tau}\right\}$.

[17] A homeomorphism or topological isomorphism is a continuous function between topological spaces that has a continuous inverse function.

[18] This is just a convention, and the opposite one is employed in some texts.

[19] Note that $\dot{\mathcal{X}^{-1}}$ represents the time derivative of the inverse, not the inverse of the time derivative.

4.3.2 Exponential and Logarithmic Maps, Plus and Minus Operators

The *exponential map* $\{\exp() : \mathfrak{m} \to \mathcal{G} \mid \mathcal{X} = \exp(\boldsymbol{\tau}^\wedge)\}$ wraps the tangent element around the manifold following the geodesic or minimum distance line, effectively converting elements of the Lie algebra into those of the manifold or Lie group. The unwrapping or inverse operation is the *logarithmic map* $\{\log() : \mathcal{G} \to \mathfrak{m} \mid \boldsymbol{\tau}^\wedge = \log(\mathcal{X})\}$. The hat and vee operators can be incorporated into these maps, resulting in the *capitalized maps* $\{\mathrm{Exp}() : \mathbb{R}^{\mathrm{m}} \to \mathcal{G} \mid \mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau})\}$ and $\{\mathrm{Log}() : \mathcal{G} \to \mathbb{R}^{\mathrm{m}} \mid \boldsymbol{\tau} = \mathrm{Log}(\mathcal{X})\}$. Note that the exponential map complies with the following properties $\forall\ \mathrm{t} \in \mathbb{R}$ [2]:

$$\exp(\mathrm{t}\,\boldsymbol{\tau}^\wedge) = \exp(\boldsymbol{\tau}^\wedge)^{\mathrm{t}} \tag{130}$$

$$\exp\left(\mathcal{X} \circ \boldsymbol{\tau}^\wedge \circ \mathcal{X}^{-1}\right) = \mathcal{X} \circ \exp(\boldsymbol{\tau}^\wedge) \circ \mathcal{X}^{-1} \tag{131}$$

The *plus* and *minus operators* enable operating with increments of the nonlinear manifold expressed in the corresponding linear tangent vector space [2]. As the Lie group $\mathcal{G}$ is not abelian, there exist right $\oplus$ and $\ominus$ operators as well as left $\boxplus$ and $\boxminus$ ones. Note that the addition of (usually) small perturbations $\Delta\boldsymbol{\tau}$ to a given manifold $\mathcal{X}$ result in a perturbed manifold $\mathcal{Y}$:

$$\mathcal{Y} = \mathcal{X} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}} = \mathcal{X} \circ \mathrm{Exp}\left(\Delta\boldsymbol{\tau}^{\mathcal{X}}\right) \in \mathcal{G} \tag{132}$$

$$\Delta\boldsymbol{\tau}^{\mathcal{X}} = \mathcal{Y} \ominus \mathcal{X} = \mathrm{Log}\left(\mathcal{X}^{-1} \circ \mathcal{Y}\right) \in \mathrm{T}_{\mathcal{X}}\mathcal{G} \tag{133}$$

$$\mathcal{Y} = \Delta\boldsymbol{\tau}^{\mathcal{E}} \boxplus \mathcal{X} = Exp\left(\Delta\boldsymbol{\tau}^{\mathcal{E}}\right) \circ \mathcal{X} \in \mathcal{G} \tag{134}$$

$$\Delta\boldsymbol{\tau}^{\mathcal{E}} = \mathcal{Y} \boxminus \mathcal{X} = \mathrm{Log}\left(\mathcal{Y} \circ \mathcal{X}^{-1}\right) \in \mathrm{T}_{\mathcal{E}}\mathcal{G} \tag{135}$$

4.3.3 Adjoint Action

The vectors or elements of the tangent space at $\mathcal{X}$ can be transformed to the tangent space at the identity $\mathcal{E}$ by means of the *adjoint* $\{\mathbf{Ad}() : \mathcal{G} \times \mathfrak{m} \to \mathfrak{m}\}$ [2]. The adjoint is hence an action of the Lie group that operates on its own Lie algebra. The adjoint action can be obtained by the equivalence of the perturbed state $\mathcal{Y}$ in (132, 134) by means of (131):

$$\boldsymbol{\tau}^{\mathcal{E}\wedge} = \mathbf{Ad}_{\mathcal{X}}\left(\boldsymbol{\tau}^{\mathcal{X}\wedge}\right) = \mathcal{X} \circ \boldsymbol{\tau}^{\mathcal{X}\wedge} \circ \mathcal{X}^{-1} \tag{136}$$

The adjoint action is a linear homomorphism, and hence complies with the following expressions $\forall\ \mathcal{X},\ \mathcal{Y} \in \mathcal{G}$, $\forall\ \mathrm{a}, \mathrm{b} \in \mathbb{R}$, $\forall\ \boldsymbol{\tau}^{\mathcal{X}\wedge}, \boldsymbol{\sigma}^{\mathcal{X}\wedge} \in \mathrm{T}_{\mathcal{X}}\mathcal{G}$, and $\forall\ \boldsymbol{\tau}^{\mathcal{Y}\wedge} \in \mathrm{T}_{\mathcal{Y}}\mathcal{G}$:

$$\mathbf{Ad}_{\mathcal{X}}\left(\mathrm{a}\,\boldsymbol{\tau}^{\mathcal{X}\wedge} + \mathrm{b}\,\boldsymbol{\sigma}^{\mathcal{X}\wedge}\right) = \mathrm{a}\,\mathbf{Ad}_{\mathcal{X}}\left(\boldsymbol{\tau}^{\mathcal{X}\wedge}\right) + \mathrm{b}\,\mathbf{Ad}_{\mathcal{X}}\left(\boldsymbol{\sigma}^{\mathcal{X}\wedge}\right) \tag{137}$$

$$\mathbf{Ad}_{\mathcal{X}}\left(\mathbf{Ad}_{\mathcal{Y}}\left(\boldsymbol{\tau}^{\mathcal{Y}\wedge}\right)\right) = \mathbf{Ad}_{\mathcal{X}\circ\mathcal{Y}}\left(\boldsymbol{\tau}^{\mathcal{Y}\wedge}\right) \tag{138}$$

As the adjoint is a linear transform, it is always possible to obtain an equivalent matrix operator, the *adjoint matrix* $\left\{\mathbf{Ad} : \mathcal{G} \times \mathbb{R}^{\mathrm{m}} \to \mathbb{R}^{\mathrm{m}} \mid \mathbf{Ad}_{\mathcal{X}} \cdot \boldsymbol{\tau} = \left(\mathcal{X}\,\boldsymbol{\tau}^\wedge\,\mathcal{X}^{-1}\right)^\vee,\ \mathbf{Ad}_{\mathcal{X}} \in \mathbb{R}^{\mathrm{mxm}}\right\}$, that maps the cartesian tangent space vectors instead of the Lie algebra elements [26]. Both maps share the same symbols but are easily distinguished by context.

$$\boldsymbol{\tau}^{\mathcal{E}} = \mathbf{Ad}_{\mathcal{X}} \cdot \boldsymbol{\tau}^{\mathcal{X}} = \left(\mathcal{X}\,\boldsymbol{\tau}^{\mathcal{X}\wedge}\,\mathcal{X}^{-1}\right)^\vee \tag{139}$$

The adjoint matrix complies with the following properties:

$$\mathcal{X} \oplus \boldsymbol{\tau}^{\mathcal{X}} = \left(\mathbf{Ad}_{\mathcal{X}} \cdot \boldsymbol{\tau}^{\mathcal{X}}\right) \boxplus \mathcal{X} \tag{140}$$

$$\mathbf{Ad}_{\mathcal{X}^{-1}} = \mathbf{Ad}_{\mathcal{X}}^{-1} \tag{141}$$

$$\mathbf{Ad}_{\mathcal{X}\circ\mathcal{Y}} = \mathbf{Ad}_{\mathcal{X}}\,\mathbf{Ad}_{\mathcal{Y}} \tag{142}$$

4.3.4 Right and Left Lie Group Derivatives

Given a function $\{\mathrm{f} : \mathcal{G} \to \mathcal{H} \mid \mathcal{Y} = \mathrm{f}(\mathcal{X}) \in \mathcal{H}, \forall \mathcal{X} \in \mathcal{G}\}$ that maps together two manifolds or Lie groups of dimensions m and n respectively, it is possible to make use of the plus and minus operators to establish right and left derivatives (jacobians) that linearly map their respective Lie algebras or tangent spaces, either locally $\left(\mathrm{T}_{\mathcal{X}}\mathcal{G} \rightarrow \mathrm{T}_{\mathrm{f}(\mathcal{X})}\mathcal{H}\right)$ if employing the right $\oplus$ and $\ominus$ operators, or globally $(\mathrm{T}_{\mathcal{E}}\mathcal{G} \rightarrow \mathrm{T}_{\mathcal{E}}\mathcal{H})$ when using the left $\boxplus$ and $\boxminus$ operators [2]. As a tangent space can always be identified to a Euclidean space of the same dimension, this enables the application of the concepts of random vectors, stochastic processes, and their correlation (section 2) to Lie algebras, leading to the section 4.3.5 expressions for Lie covariances. In addition, these derivatives constitute the basis for the construction of the Lie group jacobians in section 4.4, which in turn are indispensable for the establishment of rigorous solutions for the gradient descent minimization (optimization) and state estimation in Lie groups, as described in sections 4.6 and 4.7, respectively.

The *right jacobian* of $\mathrm{f}(\mathcal{X})$ is defined as the derivative of $\mathrm{f}(\mathcal{X})$ with respect to $\mathcal{X}$ when the increments are viewed in their respective local tangent spaces, this is, tangent respectively at $\mathcal{X} \in \mathcal{G}$ and $\mathrm{f}(\mathcal{X}) \in \mathcal{H}$ [2]. The *left jacobian* of $\mathrm{f}(\mathcal{X})$ is defined similarly, but with the increments viewed in the global tangent spaces for $\mathcal{G}$ and $\mathcal{H}$ respectively:

$$\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}(\mathcal{X})} = \lim_{\Delta\boldsymbol{\tau}^{\mathcal{X}} \to \mathbf{0}} \frac{\mathrm{f}\left(\mathcal{X} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}}\right) \ominus \mathrm{f}(\mathcal{X})}{\Delta\boldsymbol{\tau}^{\mathcal{X}}} = \lim_{\Delta\boldsymbol{\tau}^{\mathcal{X}} \to \mathbf{0}} \frac{\mathrm{Log}\left[\mathrm{f}^{-1}(\mathcal{X}) \circ \mathrm{f}\Big(\mathcal{X} \circ \mathrm{Exp}\left(\Delta\boldsymbol{\tau}^{\mathcal{X}}\right)\Big)\right]}{\Delta\boldsymbol{\tau}^{\mathcal{X}}} \in \mathbb{R}^{\mathrm{nxm}} \tag{143}$$

$$\mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}(\mathcal{X})} = \lim_{\Delta\boldsymbol{\tau}^{\mathcal{E}} \to \mathbf{0}} \frac{\mathrm{f}\left(\Delta\boldsymbol{\tau}^{\mathcal{E}} \boxplus \mathcal{X}\right) \boxminus \mathrm{f}(\mathcal{X})}{\Delta\boldsymbol{\tau}^{\mathcal{E}}} = \lim_{\Delta\boldsymbol{\tau}^{\mathcal{E}} \to \mathbf{0}} \frac{\mathrm{Log}\left[\mathrm{f}\Big(\mathrm{Exp}\left(\Delta\boldsymbol{\tau}^{\mathcal{E}} \circ \mathcal{X}\right)\Big) \circ \mathrm{f}^{-1}(\mathcal{X})\right]}{\Delta\boldsymbol{\tau}^{\mathcal{E}}} \in \mathbb{R}^{\mathrm{nxm}} \tag{144}$$

The following first order Taylor expansions can then be directly established:

$$\mathrm{f}\left(\mathcal{X} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}}\right) \approx \mathrm{f}(\mathcal{X}) \oplus \left[\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}(\mathcal{X})}\,\Delta\boldsymbol{\tau}^{\mathcal{X}}\right] = \mathcal{Y} \oplus \Delta\boldsymbol{\tau}^{\mathcal{Y}} \quad \in \mathcal{H} \tag{145}$$

$$\mathrm{f}\left(\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{G}}} \boxplus \mathcal{X}\right) \approx \left[\mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}(\mathcal{X})}\,\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{G}}}\right] \boxplus \mathrm{f}(\mathcal{X}) = \Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{H}}} \boxplus \mathcal{Y} \quad \in \mathcal{H} \tag{146}$$

Note that the $\oplus$ or $\boxplus$ symbols that appear as jacobian subindexes in (143) and (144) indicate that the domain $\mathcal{G}$ is indeed a Lie group, and should be replaced by a standard $+$ operator if this is not the case and the function f domain is in fact a real or Euclidean space. If this is the case, the expressions $\left(\mathcal{X} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}}\right)$ and $\left(\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{G}}} \boxplus \mathcal{X}\right)$ within the equations (143) through (146) shall both be replaced by $(\mathcal{X} + \Delta\boldsymbol{\tau})$.

Similarly, the $\oplus$ and $\boxplus$ symbols that appear as jacobian superindexes indicate that the function f image or codomain $\mathcal{H}$ is also a Lie group, and should otherwise be replaced by $+$ if the codomain is a Euclidean space. If this is the case, the $\ominus$ and $\boxminus$ operators within (143) and (144) shall be replaced by the standard $-$ operator, and the $\mathrm{f}(\mathcal{X}) \oplus$ and $\boxplus \mathrm{f}(\mathcal{X})$ expressions within (145) and (146) shall both be replaced by $\mathrm{f}(\mathcal{X}) +$.

Equations (145) and (146) lead to the following expressions for the function f propagation of the tangent spaces:

$$\Delta\boldsymbol{\tau}^{\mathcal{Y}} = \Delta\boldsymbol{\tau}^{\mathrm{f}(\mathcal{X})} = \mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}(\mathcal{X})}\,\Delta\boldsymbol{\tau}^{\mathcal{X}} \tag{147}$$

$$\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{H}}} = \mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}(\mathcal{X})}\,\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{G}}} \tag{148}$$

$$\tag{149}$$

In addition, (140) enables establishing a relationship between the right and left jacobians of $\mathrm{f}(\mathcal{X})$:

$$\mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}(\mathcal{X})} = \mathbf{Ad}_{\mathrm{f}(\mathcal{X})}\,\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}(\mathcal{X})}\,\mathbf{Ad}_{\mathcal{X}}^{-1} \tag{150}$$

4.3.5 Lie Groups Uncertainty and Covariance

As the $\oplus$ and $\ominus$ operators can be employed to define perturbations in the local tangent space $\left(\mathrm{T}_{\boldsymbol{\mu}_{\mathcal{X}}}\mathcal{G}\right)$ around a nominal or expected point $\mathrm{E}[\mathcal{X}] = \boldsymbol{\mu}_{\mathcal{X}} \in \mathcal{G}$, it is possible to employ a *local autocovariance* definition similar to the one used for Euclidean spaces (24), leading to the definition of stochastic variables (vectors) on manifolds $\mathcal{X} \sim \left(\boldsymbol{\mu}_{\mathcal{X}},\ \mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}}\right)$ [2]:

$$\mathcal{X} = \boldsymbol{\mu}_{\mathcal{X}} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}} \quad \in \mathcal{G} \tag{151}$$

$$\Delta\boldsymbol{\tau}^{\mathcal{X}} = \mathcal{X} \ominus \boldsymbol{\mu}_{\mathcal{X}} \quad \in \mathrm{T}_{\boldsymbol{\mu}_{\mathcal{X}}}\mathcal{G} \tag{152}$$

$$\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}} = \mathrm{E}\left[\Delta\boldsymbol{\tau}^{\mathcal{X}}\,\Delta\boldsymbol{\tau}^{\mathcal{X},\mathrm{T}}\right] = \mathrm{E}\left[\left(\mathcal{X} \ominus \boldsymbol{\mu}_{\mathcal{X}}\right)\left(\mathcal{X} \ominus \boldsymbol{\mu}_{\mathcal{X}}\right)^{\mathrm{T}}\right] \quad \in \mathbb{R}^{\mathrm{mxm}} \tag{153}$$

The different types of correlation matrices for stochastic processes introduced in section 2.3 can also be defined accordingly. Note that although the notation of $\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}}$ refers to the covariance of the manifold or Lie group $\mathcal{X} \in \mathcal{G}$, the definition in fact refers to the covariance of the nominal point local tangent space $\Delta\boldsymbol{\tau}^{\mathcal{X}} \in \mathrm{T}_{\boldsymbol{\mu}\mathcal{X}}\mathcal{G}$, with its dimension matching the number of degrees of freedom of the manifold.

A similar process employing the $\boxplus$ and $\boxminus$ operators leads to the *global autocovariance*:

$$\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{E}} = \mathrm{E}\left[\Delta\boldsymbol{\tau}^{\mathcal{E}}\,\Delta\boldsymbol{\tau}^{\mathcal{E},\mathrm{T}}\right] = \mathrm{E}\left[\left(\mathcal{X} \boxminus \boldsymbol{\mu}_{\mathcal{X}}\right)\left(\mathcal{X} \boxminus \boldsymbol{\mu}_{\mathcal{X}}\right)^{\mathrm{T}}\right] \quad \in \mathbb{R}^{\mathrm{mxm}} \tag{154}$$

$$\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{E}} = \mathrm{E}\left[\mathbf{Ad}_{\mathcal{X}}\,\Delta\boldsymbol{\tau}^{\mathcal{X}}\,\Delta\boldsymbol{\tau}^{\mathcal{X},\mathrm{T}}\,\mathbf{Ad}_{\mathcal{X}}^{\mathrm{T}}\right] = \mathbf{Ad}_{\mathcal{X}}\;\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}}\;\mathbf{Ad}_{\mathcal{X}}^{\mathrm{T}} \tag{155}$$

$$\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}} = \mathrm{E}\left[\mathbf{Ad}_{\mathcal{X}}^{-1}\,\Delta\boldsymbol{\tau}^{\mathcal{E}}\,\Delta\boldsymbol{\tau}^{\mathcal{E},\mathrm{T}}\,\mathbf{Ad}_{\mathcal{X}}^{-\mathrm{T}}\right] = \mathbf{Ad}_{\mathcal{X}}^{-1}\;\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{E}}\;\mathbf{Ad}_{\mathcal{X}}^{-\mathrm{T}} \tag{156}$$

Given a function $\{\mathrm{f} : \mathcal{G} \rightarrow \mathcal{H} \mid \mathcal{Y} = \mathrm{f}\left(\mathcal{X}\right) \in \mathcal{H},\, \forall \mathcal{X} \in \mathcal{G}\}$ as that introduced in section 4.3.4, which maps together two manifolds or Lie groups of dimensions m and n respectively, it is possible to propagate the covariances $\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}}$ and $\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{E}}$ from the domain manifold $\mathcal{G}$ to the image one $\mathcal{H}$:

$$\begin{aligned}
\mathbf{C}_{\mathcal{Y}\mathcal{Y}}^{\mathcal{Y}} &= \mathrm{E}\left[\Delta\boldsymbol{\tau}^{\mathcal{Y}}\,\Delta\boldsymbol{\tau}^{\mathcal{Y},\mathrm{T}}\right] = \mathrm{E}\left[\Delta\boldsymbol{\tau}^{\mathrm{f}(\mathcal{X})}\,\Delta\boldsymbol{\tau}^{\mathrm{f}(\mathcal{X}),\mathrm{T}}\right] \\
&= \mathrm{E}\left[\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}\left(\mathcal{X}\right)}\,\Delta\boldsymbol{\tau}^{\mathcal{X}}\,\Delta\boldsymbol{\tau}^{\mathcal{X},\mathrm{T}}\,\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}\left(\mathcal{X}\right),\mathrm{T}}\right] = \mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}\left(\mathcal{X}\right)}\;\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{X}}\;\mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\mathrm{f}\left(\mathcal{X}\right),\mathrm{T}} \quad \in \mathbb{R}^{\mathrm{nxn}}
\end{aligned} \tag{157}$$

$$\mathbf{C}_{\mathcal{Y}\mathcal{Y}}^{\mathcal{E}} = \mathrm{E}\left[\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{H}}}\,\Delta\boldsymbol{\tau}^{\mathcal{E}_{\mathcal{H}},\mathrm{T}}\right] = \mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}\left(\mathcal{X}\right)}\;\mathbf{C}_{\mathcal{X}\mathcal{X}}^{\mathcal{E}}\;\mathbf{J}_{\boxplus\,\mathcal{X}}^{\boxplus\,\mathrm{f}\left(\mathcal{X}\right),\mathrm{T}} \quad \in \mathbb{R}^{\mathrm{nxn}} \tag{158}$$

The establishment and propagation of covariance matrices of the proper dimensions is key for the application of state estimation techniques such as Kalman filtering when some of the state vector components belong to Lie manifolds and their tangent spaces, as described in section 4.7.

4.4 Euclidean and Lie Jacobians

The definition of the proper derivative matrices or jacobians is indispensable for all calculus techniques that rely on linearization, such as optimization by means of the gradient descent method (section 4.6) or state estimation through Kalman filtering (section 4.7). Given a function $\{\mathrm{f}\left(\mathbf{x}\right) : \mathbb{R}^{\mathrm{m}} \rightarrow \mathbb{R}^{\mathrm{n}}\}$, its (Euclidean) jacobian $\mathbf{J}_{+\,\mathbf{x}}^{+\,\mathrm{f}(\mathbf{x})} \in \mathbb{R}^{\mathrm{nxm}}$ stacks the partial derivatives of each component of the output space with respect to those of the input space:

$$\mathrm{J}_{+\,\mathbf{x},\mathrm{ij}}^{+\,\mathrm{f}(\mathbf{x})} = \lim_{\Delta\mathrm{x_j} \rightarrow 0} \frac{\mathrm{f_i}\left(\mathrm{x_j} + \Delta\mathrm{x_j}\right) - \mathrm{f_i}\left(\mathrm{x_j}\right)}{\Delta\mathrm{x_j}} \tag{159}$$

This section relies on the right and left Lie group derivatives introduced in section 4.3.4 to properly define jacobians when either the input or output spaces (or both) are not Euclidean but Lie groups. The various jacobians listed in table 1 have been obtained by means of the expressions that appear on section 4.3 together with the chain rule, and include instances in which both the domain and codomain of the $\mathrm{f}\left(\mathcal{X}\right)$ function are either Euclidean or Lie groups. Some are generic, while others, which rely on group actions, depend on the specific set on which the action is applied and hence can only be established for a specific Lie group, such as rotational or rigid body motions (sections 5 and 6).

There are two constructions that appear repeatedly within table 1. The first is the adjoint matrix (139), which maps the local and global tangent spaces at a given point on a manifold or Lie group, while the second are the right and left jacobians of the capitalized exponential function, also known as simply the *right jacobian* $\mathrm{J_R}\left(\boldsymbol{\tau}\right)$ and the *left jacobian* $\mathrm{J_L}\left(\boldsymbol{\tau}\right)$. These compare variations in the tangent space of the output $\mathrm{Exp}\left(\boldsymbol{\tau}\right)$ map (locally for $\mathrm{J_R}$ and globally for $\mathrm{J_L}$) with (Euclidean) variations in the input argument $\boldsymbol{\tau}$, and are obtained in sections

5.12 and 6.13 for the specific cases of rotational and rigid body motions, respectively.

$$
\begin{aligned}
\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})} &= \mathbf{J}_{\mathrm{L}}(\boldsymbol{\tau})\,\mathbf{J}_{\mathrm{R}}^{-1}(\boldsymbol{\tau}) && (160)\\
\mathbf{J}_{\mathrm{R}}(-\boldsymbol{\tau}) &= \mathbf{J}_{\mathrm{L}}(\boldsymbol{\tau}) && (161)
\end{aligned}
$$

Being located in the tangent spaces, the right and left jacobians can be related by means of the adjoint, resulting in (160). Expression (161) in turn can be obtained by means of the chain rule.

Jacobian	Result	Taylor Expansion			
$\mathbf{J}^{\oplus\,\mathcal{X}^{-1}}_{\oplus\,\mathcal{X}}$	$-\mathbf{Ad}_{\mathcal{X}}$	$(\mathcal{X}\oplus\Delta\boldsymbol{\tau})^{-1}$	$\approx$	$\mathcal{X}^{-1}\oplus\left[\mathbf{J}^{\oplus\,\mathcal{X}^{-1}}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathcal{G}$
$\mathbf{J}^{\boxplus\,\mathcal{X}^{-1}}_{\boxplus\,\mathcal{X}}$	$-\mathbf{Ad}_{\mathcal{X}}^{-1}$	$(\Delta\boldsymbol{\tau}\boxplus\mathcal{X})^{-1}$	$\approx$	$\left[\mathbf{J}^{\boxplus\,\mathcal{X}^{-1}}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]\boxplus\mathcal{X}^{-1}$	$\in\mathcal{G}$
$\mathbf{J}^{\oplus\,\mathcal{X}\circ\mathcal{Y}}_{\oplus\,\mathcal{X}}$	$\mathbf{Ad}_{\mathcal{Y}}^{-1}$	$(\mathcal{X}\oplus\Delta\boldsymbol{\tau})\circ\mathcal{Y}$	$\approx$	$(\mathcal{X}\circ\mathcal{Y})\oplus\left[\mathbf{J}^{\oplus\,\mathcal{X}\circ\mathcal{Y}}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathcal{G}$
$\mathbf{J}^{\boxplus\,\mathcal{X}\circ\mathcal{Y}}_{\boxplus\,\mathcal{X}}$	$\mathbf{I}_{\mathrm{mxm}}$	$(\Delta\boldsymbol{\tau}\boxplus\mathcal{X})\circ\mathcal{Y}$	$\approx$	$\left[\mathbf{J}^{\boxplus\,\mathcal{X}\circ\mathcal{Y}}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]\boxplus(\mathcal{X}\circ\mathcal{Y})$	$\in\mathcal{G}$
$\mathbf{J}^{\oplus\,\mathcal{X}\circ\mathcal{Y}}_{\oplus\,\mathcal{Y}}$	$\mathbf{I}_{\mathrm{mxm}}$	$\mathcal{X}\circ(\mathcal{Y}\oplus\Delta\boldsymbol{\tau})$	$\approx$	$(\mathcal{X}\circ\mathcal{Y})\oplus\left[\mathbf{J}^{\oplus\,\mathcal{X}\circ\mathcal{Y}}_{\oplus\,\mathcal{Y}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathcal{G}$
$\mathbf{J}^{\boxplus\,\mathcal{X}\circ\mathcal{Y}}_{\boxplus\,\mathcal{Y}}$	$\mathbf{Ad}_{\mathcal{X}}$	$\mathcal{X}\circ(\Delta\boldsymbol{\tau}\boxplus\mathcal{Y})$	$\approx$	$\left[\mathbf{J}^{\boxplus\,\mathcal{X}\circ\mathcal{Y}}_{\boxplus\,\mathcal{Y}}\,\Delta\boldsymbol{\tau}\right]\boxplus(\mathcal{X}\circ\mathcal{Y})$	$\in\mathcal{G}$
$\mathbf{J}^{\oplus\,\mathrm{Exp}(\boldsymbol{\tau})}_{+\,\boldsymbol{\tau}}$	$\mathrm{J}_{\mathrm{R}}(\boldsymbol{\tau})$[20]	$\mathrm{Exp}(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})$	$\approx$	$\mathrm{Exp}(\boldsymbol{\tau})\oplus[\mathrm{J}_{\mathrm{R}}(\boldsymbol{\tau})\,\Delta\boldsymbol{\tau}]$	$\in\mathcal{G}$
$\mathrm{J}_{\mathrm{R}}^{-1}(\boldsymbol{\tau})$		$\boldsymbol{\tau}+\mathbf{J}_{\mathrm{R}}^{-1}(\boldsymbol{\tau})\,\Delta\boldsymbol{\tau}$	$\approx$[21]	$\mathrm{Log}\Big(\mathrm{Exp}(\boldsymbol{\tau})\oplus\Delta\boldsymbol{\tau}\Big)$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{\boxplus\,\mathrm{Exp}(\boldsymbol{\tau})}_{+\,\boldsymbol{\tau}}$	$\mathrm{J}_{\mathrm{L}}(\boldsymbol{\tau})$[22]	$\mathrm{Exp}(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})$	$\approx$	$[\mathrm{J}_{\mathrm{L}}(\boldsymbol{\tau})\,\Delta\boldsymbol{\tau}]\boxplus\mathrm{Exp}(\boldsymbol{\tau})$	$\in\mathcal{G}$
$\mathrm{J}_{\mathrm{L}}^{-1}(\boldsymbol{\tau})$		$\boldsymbol{\tau}+\mathbf{J}_{\mathrm{L}}^{-1}(\boldsymbol{\tau})\,\Delta\boldsymbol{\tau}$	$\approx$[23]	$\mathrm{Log}\Big(\Delta\boldsymbol{\tau}\boxplus\mathrm{Exp}(\boldsymbol{\tau})\Big)$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathrm{Log}(\mathcal{X})}_{\oplus\,\mathcal{X}}$	$\mathrm{J}_{\mathrm{R}}^{-1}\big(\mathrm{Log}(\mathcal{X})\big)$	$\mathrm{Log}(\mathcal{X}\oplus\Delta\boldsymbol{\tau})$	$\approx$	$\mathrm{Log}(\mathcal{X})+\left[\mathbf{J}^{+\,\mathrm{Log}(\mathcal{X})}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathrm{Log}(\mathcal{X})}_{\boxplus\,\mathcal{X}}$	$\mathrm{J}_{\mathrm{L}}^{-1}\big(\mathrm{Log}(\mathcal{X})\big)$	$\mathrm{Log}(\Delta\boldsymbol{\tau}\boxplus\mathcal{X})$	$\approx$	$\mathrm{Log}(\mathcal{X})+\left[\mathbf{J}^{+\,\mathrm{Log}(\mathcal{X})}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{\oplus\,\mathcal{X}\oplus\boldsymbol{\tau}}_{\oplus\,\mathcal{X}}$	$\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}^{-1}$	$(\mathcal{X}\oplus\Delta\boldsymbol{\tau})\oplus\boldsymbol{\tau}$	$\approx$	$(\mathcal{X}\oplus\boldsymbol{\tau})\oplus\left[\mathbf{J}^{\oplus\,\mathcal{X}\oplus\boldsymbol{\tau}}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathcal{G}$
$\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{X}}_{\boxplus\,\mathcal{X}}$	$\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}$	$\boldsymbol{\tau}\boxplus(\Delta\boldsymbol{\tau}\boxplus\mathcal{X})$	$\approx$	$\left[\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{X}}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]\boxplus(\boldsymbol{\tau}\boxplus\mathcal{X})$	$\in\mathcal{G}$
$\mathbf{J}^{\oplus\,\mathcal{X}\oplus\boldsymbol{\tau}}_{+\,\boldsymbol{\tau}}$	$\mathrm{J}_{\mathrm{R}}(\boldsymbol{\tau})$	$\mathcal{X}\oplus(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})$	$\approx$	$(\mathcal{X}\oplus\boldsymbol{\tau})\oplus\left[\mathbf{J}^{\oplus\,\mathcal{X}\oplus\boldsymbol{\tau}}_{+\,\boldsymbol{\tau}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathcal{G}$
$\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{X}}_{+\,\boldsymbol{\tau}}$	$\mathrm{J}_{\mathrm{L}}(\boldsymbol{\tau})$	$(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})\boxplus\mathcal{X}$	$\approx$	$\left[\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{X}}_{+\,\boldsymbol{\tau}}\,\Delta\boldsymbol{\tau}\right]\boxplus(\boldsymbol{\tau}\boxplus\mathcal{X})$	$\in\mathcal{G}$
$\mathbf{J}^{+\,\mathcal{Y}\ominus\mathcal{X}}_{\oplus\,\mathcal{X}}$	$-\mathrm{J}_{\mathrm{L}}^{-1}(\mathcal{Y}\ominus\mathcal{X})$	$\mathcal{Y}\ominus(\mathcal{X}\oplus\Delta\boldsymbol{\tau})$	$\approx$	$(\mathcal{Y}\ominus\mathcal{X})+\left[\mathbf{J}^{+\,\mathcal{Y}\ominus\mathcal{X}}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathcal{Y}\boxminus\mathcal{X}}_{\boxplus\,\mathcal{X}}$	$-\mathrm{J}_{\mathrm{R}}^{-1}(\mathcal{Y}\boxminus\mathcal{X})$	$\mathcal{Y}\boxminus(\Delta\boldsymbol{\tau}\boxplus\mathcal{X})$	$\approx$	$(\mathcal{Y}\boxminus\mathcal{X})+\left[\mathbf{J}^{+\,\mathcal{Y}\boxminus\mathcal{X}}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathcal{Y}\ominus\mathcal{X}}_{\oplus\,\mathcal{Y}}$	$\mathrm{J}_{\mathrm{R}}^{-1}(\mathcal{Y}\ominus\mathcal{X})$	$(\mathcal{Y}\oplus\Delta\boldsymbol{\tau})\ominus\mathcal{X}$	$\approx$	$(\mathcal{Y}\ominus\mathcal{X})+\left[\mathbf{J}^{+\,\mathcal{Y}\ominus\mathcal{X}}_{\oplus\,\mathcal{Y}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathcal{Y}\boxminus\mathcal{X}}_{\boxplus\,\mathcal{Y}}$	$\mathrm{J}_{\mathrm{L}}^{-1}(\mathcal{Y}\boxminus\mathcal{X})$	$(\Delta\boldsymbol{\tau}\boxplus\mathcal{Y})\boxminus\mathcal{X}$	$\approx$	$(\mathcal{Y}\boxminus\mathcal{X})+\left[\mathbf{J}^{+\,\mathcal{Y}\boxminus\mathcal{X}}_{\boxplus\,\mathcal{Y}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{\oplus\,\mathcal{X}}$	Tables	$\mathrm{g}_{\mathcal{X}\oplus\Delta\boldsymbol{\tau}}(\mathbf{u})$	$\approx$	$\mathrm{g}_{\mathcal{X}}(\mathbf{u})+\left[\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{\boxplus\,\mathcal{X}}$	8	$\mathrm{g}_{\Delta\boldsymbol{\tau}\boxplus\mathcal{X}}(\mathbf{u})$	$\approx$	$\mathrm{g}_{\mathcal{X}}(\mathbf{u})+\left[\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{+\,\mathbf{u}}$	&	$\mathrm{g}_{\mathcal{X}}(\mathbf{u}+\Delta\mathbf{u})$	$\approx$	$\mathrm{g}_{\mathcal{X}}(\mathbf{u})+\left[\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}(\mathbf{u})}_{+\,\mathbf{u}}\,\Delta\mathbf{u}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}^{-1}(\mathbf{u})}_{\oplus\,\mathcal{X}}$	16	$\mathrm{g}_{\mathcal{X}\oplus\Delta\boldsymbol{\tau}}^{-1}(\mathbf{u})$	$\approx$	$\mathrm{g}_{\mathcal{X}}^{-1}(\mathbf{u})+\left[\mathbf{J}^{+\,\mathrm{g}_{\mathcal{X}}^{-1}(\mathbf{u})}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$

[20] Obtained in sections 5.12 and 6.13 for rotational and rigid body motion, respectively.
[21] Obtained by replacing $\Delta\boldsymbol{\tau}$ by $\mathbf{J}_{\mathrm{R}}^{-1}\,\Delta\boldsymbol{\tau}$ in the expression above.
[22] Obtained in sections 5.12 and 6.13 for rotational and rigid body motion, respectively.
[23] Obtained by replacing $\Delta\boldsymbol{\tau}$ by $\mathbf{J}_{\mathrm{L}}^{-1}\,\Delta\boldsymbol{\tau}$ in the expression above.

Jacobian	Result	Taylor Expansion			
$\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u})}_{\boxplus\,\mathcal{X}}$		$\mathrm{g}^{-1}_{\Delta\boldsymbol{\tau}\boxplus\mathcal{X}}(\mathbf{u})$	$\approx$	$\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u}) + \left[\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u})}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u})}_{+\,\mathbf{u}}$		$\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u}+\Delta\mathbf{u})$	$\approx$	$\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u}) + \left[\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathcal{X}}(\mathbf{u})}_{+\,\mathbf{u}}\,\Delta\mathbf{u}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{\oplus\,\mathcal{X}}$		$\mathbf{Ad}_{\mathcal{X}\oplus\Delta\boldsymbol{\tau}}(\mathbf{v})$	$\approx$	$\mathbf{Ad}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{\boxplus\,\mathcal{X}}$	Tables	$\mathbf{Ad}_{\Delta\boldsymbol{\tau}\boxplus\mathcal{X}}(\mathbf{v})$	$\approx$	$\mathbf{Ad}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{+\,\mathbf{v}}$	8	$\mathbf{Ad}_{\mathcal{X}}(\mathbf{v}+\Delta\mathbf{v})$	$\approx$	$\mathbf{Ad}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{X}}(\mathbf{v})}_{+\,\mathbf{v}}\,\Delta\mathbf{v}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{\oplus\,\mathcal{X}}$	&	$\mathbf{Ad}^{-1}_{\mathcal{X}\oplus\Delta\boldsymbol{\tau}}(\mathbf{v})$	$\approx$	$\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{\oplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{\boxplus\,\mathcal{X}}$	16	$\mathbf{Ad}^{-1}_{\Delta\boldsymbol{\tau}\boxplus\mathcal{X}}(\mathbf{v})$	$\approx$	$\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{\boxplus\,\mathcal{X}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{+\,\mathbf{v}}$		$\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v}+\Delta\mathbf{v})$	$\approx$	$\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v}) + \left[\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{X}}(\mathbf{v})}_{+\,\mathbf{v}}\,\Delta\mathbf{v}\right]$	$\in\mathbb{R}^{\mathrm{m}}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u})}_{+\,\boldsymbol{\tau}}$		$\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})}(\mathbf{u})$	$\approx$	$\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u}) + \left[\mathbf{J}^{+\,\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u})}_{+\,\boldsymbol{\tau}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$
$\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u})}_{+\,\boldsymbol{\tau}}$		$\mathrm{g}^{-1}_{\mathrm{Exp}(\boldsymbol{\tau}+\Delta\boldsymbol{\tau})}(\mathbf{u})$	$\approx$	$\mathrm{g}^{-1}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u}) + \left[\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{u})}_{+\,\boldsymbol{\tau}}\,\Delta\boldsymbol{\tau}\right]$	$\in\mathbb{R}^{\mathrm{u}}$

Table 1: Lie jacobians

4.5 Discrete Integration in Lie Groups

Following on the discrete integration in Euclidean spaces described in section 3.1, let's now consider that the state system is composed by a vector $\mathbf{y}\in\mathbb{R}^{\mathrm{n}}$, an element of a Lie group $\mathcal{X}\in\mathcal{G}$, and a vector $\mathbf{v}^{\mathcal{X}}\in\mathbb{R}^{\mathrm{m}}$ representing the velocity of $\mathcal{X}$ as it moves along its manifold, contained in the local or body tangent space $\mathrm{T}_{\mathcal{X}}\mathcal{G}$. Neither $\mathcal{X}$ nor its components are Euclidean, and hence the section 3.1 integration schemes are not applicable. If treated so, the resulting element $\mathcal{X}_{\mathrm{k+1}}$ would not be located on the manifold as it would not comply with the Lie group constraints, and it would be necessary to reproject it back to it, incurring in errors that although small for a single integration step may become significant when accumulated.

Let's group these states into a composite state vector made up by n plus m components of an Euclidean space plus an element of a Lie group. As in the Euclidean case, the initial composite vector value is known:

$$\mathbf{x} = \left[\mathbf{y},\,\mathbf{v}^{\mathcal{X}},\,\mathcal{X}\right]^{\mathrm{T}} \tag{162}$$

$$\mathbf{x}_{\mathrm{k}} = \mathbf{x}\left(\mathrm{t_k}\right) = \left[\mathbf{y}_{\mathrm{k}},\,\mathbf{v}^{\mathcal{X}}_{\mathrm{k}},\,\mathcal{X}_{\mathrm{k}}\right]^{\mathrm{T}} \tag{163}$$

As in the Euclidean case, the objective is the determination of the composite state vector value at a time $\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}\left(\mathrm{t_{k+1}}\right) = \mathbf{x}\left(\mathrm{t_k}+\Delta\mathrm{t}\right)$ by relying on evaluations of the $\mathbf{y}$ and $\mathbf{v}^{\mathcal{X}}$ time derivatives:

$$\dot{\mathbf{y}}\left(\mathrm{t}\right) = \mathrm{f_y}\big(\mathbf{y}\left(\mathrm{t}\right),\,\mathbf{v}^{\mathcal{X}}\left(\mathrm{t}\right),\,\mathcal{X}\left(\mathrm{t}\right),\,\mathrm{t}\big) \tag{164}$$

$$\dot{\mathbf{v}}^{\mathcal{X}}\left(\mathrm{t}\right) = \mathrm{f_v}\big(\mathbf{y}\left(\mathrm{t}\right),\,\mathbf{v}^{\mathcal{X}}\left(\mathrm{t}\right),\,\mathcal{X}\left(\mathrm{t}\right),\,\mathrm{t}\big) \tag{165}$$

The solution consists on employing the Euclidean integration method of choice to obtain $\mathbf{y}_{\mathrm{k+1}}$ and $\mathbf{v}^{\mathcal{X}}_{\mathrm{k+1}}$, but rely on the right plus operator and the capitalized exponential map of the Lie group to determine $\mathcal{X}_{\mathrm{k+1}}$. In case of Euler's method, the solution is the following:

$$\mathbf{y}_{\mathrm{k+1}} \approx \mathbf{y}_{\mathrm{k}} + \Delta\mathrm{t}\;\dot{\mathbf{y}}(\mathbf{y}_{\mathrm{k}},\,\mathbf{v}^{\mathcal{X}}_{\mathrm{k}},\,\mathcal{X}_{\mathrm{k}},\,\mathrm{t_k}) \tag{166}$$

$$\mathbf{v}^{\mathcal{X}}_{\mathrm{k+1}} \approx \mathbf{v}^{\mathcal{X}}_{\mathrm{k}} + \Delta\mathrm{t}\;\dot{\mathbf{v}}^{\mathcal{X}}(\mathbf{y}_{\mathrm{k}},\,\mathbf{v}^{\mathcal{X}}_{\mathrm{k}},\,\mathcal{X}_{\mathrm{k}},\,\mathrm{t_k}) \tag{167}$$

$$\mathcal{X}_{\mathrm{k+1}} \approx \mathcal{X}_{\mathrm{k}} \oplus \left[\Delta\mathrm{t}\;\mathbf{v}^{\mathcal{X}}_{\mathrm{k}}\right] = \mathcal{X}_{\mathrm{k}} \circ \mathrm{Exp}\left(\Delta\mathrm{t}\;\mathbf{v}^{\mathcal{X}}_{\mathrm{k}}\right) \tag{168}$$

$$\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}\left(\mathrm{t_{k+1}}\right) \approx \left[\mathbf{y}_{\mathrm{k+1}},\,\mathbf{v}^{\mathcal{X}}_{\mathrm{k+1}},\,\mathcal{X}_{\mathrm{k+1}}\right]^{\mathrm{T}} \tag{169}$$

In case the state vector velocity $\mathbf{v}^{\mathcal{E}}$ is that of the global or space tangent space $\mathrm{T}_{\mathcal{E}}\mathcal{G}$, it is necessary to modify (168) to employ the left plus operator:

$$\mathcal{X}_{\mathrm{k+1}} \approx \left[\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{E}}\right] \boxplus \mathcal{X}_{\mathrm{k}} = \mathrm{Exp}\left(\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{E}}\right) \circ \mathcal{X}_{\mathrm{k}} \tag{170}$$

It is necessary to proceed in a similar manner if a different integration method is selected. In the case of Heun's method with local velocity $\mathbf{v}^{\mathcal{X}}$, the modified integration scheme is the following:

$$\begin{aligned}
\dot{\mathbf{y}}_1 &= \dot{\mathbf{y}}(\mathbf{y}_{\mathrm{k}}, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}}, \mathrm{t}_{\mathrm{k}}) && (171)\\
\dot{\mathbf{y}}_2 &= \dot{\mathbf{y}}(\mathbf{y}_{\mathrm{k}} + \Delta \mathrm{t} \ \dot{\mathbf{y}}_1, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_1^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathrm{t}_{\mathrm{k}} + \Delta \mathrm{t}) && (172)\\
\dot{\mathbf{v}}_1^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}(\mathbf{y}_{\mathrm{k}}, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}}, \mathrm{t}_{\mathrm{k}}) && (173)\\
\dot{\mathbf{v}}_2^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}(\mathbf{y}_{\mathrm{k}} + \Delta \mathrm{t} \ \dot{\mathbf{y}}_1, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_1^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathrm{t}_{\mathrm{k}} + \Delta \mathrm{t}) && (174)\\
\mathbf{y}_{\mathrm{k+1}} &\approx \mathbf{y}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2} \ [\dot{\mathbf{y}}_1 + \dot{\mathbf{y}}_2] && (175)\\
\mathbf{v}_{\mathrm{k+1}}^{\mathcal{X}} &\approx \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \left[\dot{\mathbf{v}}_1^{\mathcal{X}} + \dot{\mathbf{v}}_2^{\mathcal{X}}\right] && (176)\\
\mathcal{X}_{\mathrm{k+1}} &\approx \mathcal{X}_{\mathrm{k}} \oplus \left[\frac{\Delta \mathrm{t}}{2} \ \left[\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \left(\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right)\right]\right] = \mathcal{X}_{\mathrm{k}} \oplus \left[\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right] \\
&= \mathcal{X}_{\mathrm{k}} \circ \mathrm{Exp}\left(\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right) && (177)
\end{aligned}$$

In case of the 4^{th} order Runge-Kutta integration scheme, it results in the following:

$$\begin{aligned}
\dot{\mathbf{y}}_1 &= \dot{\mathbf{y}}\left(\mathbf{y}_{\mathrm{k}}, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}}, \mathrm{t}_{\mathrm{k}}\right) && (178)\\
\dot{\mathbf{v}}_1^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}\left(\mathbf{y}_{\mathrm{k}}, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}}, \mathrm{t}_{\mathrm{k}}\right) && (179)\\
\dot{\mathbf{y}}_2 &= \dot{\mathbf{y}}\left(\mathbf{y}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{y}}_1, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \frac{\Delta \mathrm{t}}{2} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathrm{t}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2}\right) && (180)\\
\dot{\mathbf{v}}_2^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}\left(\mathbf{y}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{y}}_1, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \frac{\Delta \mathrm{t}}{2} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}}, \mathrm{t}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2}\right) && (181)\\
\dot{\mathbf{y}}_3 &= \dot{\mathbf{y}}\left(\mathbf{y}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{y}}_2, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_2^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \frac{\Delta \mathrm{t}}{2} \ \left[\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right], \mathrm{t}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2}\right) && (182)\\
\dot{\mathbf{v}}_3^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}\left(\mathbf{y}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{y}}_2, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_2^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \frac{\Delta \mathrm{t}}{2} \ \left[\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right], \mathrm{t}_{\mathrm{k}} + \frac{\Delta \mathrm{t}}{2}\right) && (183)\\
\dot{\mathbf{y}}_4 &= \dot{\mathbf{y}}\left(\mathbf{y}_{\mathrm{k}} + \Delta \mathrm{t} \ \dot{\mathbf{y}}_3, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_3^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \Delta \mathrm{t} \ \left[\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_2^{\mathcal{X}}\right], \mathrm{t}_{\mathrm{k}} + \Delta \mathrm{t}\right) && (184)\\
\dot{\mathbf{v}}_4^{\mathcal{X}} &= \dot{\mathbf{v}}^{\mathcal{X}}\left(\mathbf{y}_{\mathrm{k}} + \Delta \mathrm{t} \ \dot{\mathbf{y}}_3, \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_3^{\mathcal{X}}, \mathcal{X}_{\mathrm{k}} \oplus \Delta \mathrm{t} \ \left[\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_2^{\mathcal{X}}\right], \mathrm{t}_{\mathrm{k}} + \Delta \mathrm{t}\right) && (185)\\
\mathbf{y}_{\mathrm{k+1}} &\approx \mathbf{y}_{\mathrm{k}} + \Delta \mathrm{t} \ \left[\frac{\dot{\mathbf{y}}_1}{6} + \frac{\dot{\mathbf{y}}_2}{3} + \frac{\dot{\mathbf{y}}_3}{3} + \frac{\dot{\mathbf{y}}_4}{6}\right] && (186)\\
\mathbf{v}_{\mathrm{k+1}}^{\mathcal{X}} &\approx \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \left[\frac{\dot{\mathbf{v}}_1^{\mathcal{X}}}{6} + \frac{\dot{\mathbf{v}}_2^{\mathcal{X}}}{3} + \frac{\dot{\mathbf{v}}_3^{\mathcal{X}}}{3} + \frac{\dot{\mathbf{v}}_4^{\mathcal{X}}}{6}\right] && (187)\\
\mathcal{X}_{\mathrm{k+1}} &\approx \mathcal{X}_{\mathrm{k}} \oplus \left[\frac{\Delta \mathrm{t}}{6} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{3} \ \left(\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_1^{\mathcal{X}}\right) + \frac{\Delta \mathrm{t}}{3} \ \left(\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}}{2} \ \dot{\mathbf{v}}_2^{\mathcal{X}}\right) + \frac{\Delta \mathrm{t}}{6} \ \left(\mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \Delta \mathrm{t} \ \dot{\mathbf{v}}_3^{\mathcal{X}}\right)\right] \\
&= \mathcal{X}_{\mathrm{k}} \oplus \left[\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_1^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_2^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_3^{\mathcal{X}}\right] \\
&= \mathcal{X}_{\mathrm{k}} \circ \mathrm{Exp}\left(\Delta \mathrm{t} \ \mathbf{v}_{\mathrm{k}}^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_1^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_2^{\mathcal{X}} + \frac{\Delta \mathrm{t}^2}{6} \ \dot{\mathbf{v}}_3^{\mathcal{X}}\right) && (188)
\end{aligned}$$

Similar modifications to that of (170) are required if the tangent space velocity $\mathbf{v}^{\mathcal{E}}$ is viewed in the global space.

4.6 Gradient Descent Optimization in Lie Groups

The Gauss-Newton implementation of the gradient descent optimization method is described in section 3.2 for the case of Euclidean spaces. Let's now consider a Lie group element $\mathcal{X} \in \mathcal{G}$ and a nonlinear map based on its tangent space $\{\mathbf{f} : \mathbb{R}^{\mathrm{m}} \rightarrow \mathbb{R}^{\mathrm{n}} \mid \mathbf{f}(\boldsymbol{\tau}) \in \mathbb{R}^{\mathrm{n}}, \boldsymbol{\tau} = \mathrm{Log}(\mathcal{X}) \in \mathbb{R}^{\mathrm{m}}, \boldsymbol{\tau}^{\wedge} = \log(\mathcal{X}) \in \mathfrak{m}, \forall\, \mathcal{X} \in \mathcal{G}\}$. As in the Euclidean case, it is possible to evaluate its jacobian $\{\mathbf{J} : \mathbb{R}^{\mathrm{m}} \rightarrow \mathbb{R}^{\mathrm{nxm}} \mid \mathbf{J}(\boldsymbol{\tau}) = \partial \mathbf{f}(\boldsymbol{\tau}) / \partial \boldsymbol{\tau} \in \mathbb{R}^{\mathrm{nxm}}, \forall\, \boldsymbol{\tau} \in \mathbb{R}^{\mathrm{m}}\}$, and the error function is defined as $\boldsymbol{\mathcal{E}}(\boldsymbol{\tau}) = \boldsymbol{\mathcal{E}}\big(\mathrm{Log}(\mathcal{X})\big) = \mathbf{f}(\boldsymbol{\tau}) - \mathbf{f}_{\mathrm{T}}$.

Given an initial state $\mathcal{X}_0 = \mathrm{Exp}(\boldsymbol{\tau}_0)$, the objective is to determine a Lie group element $\mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau}) = \Delta\boldsymbol{\tau}^{\mathcal{E}} \boxplus \mathcal{X}_0 = \Delta\boldsymbol{\tau}^{\mathcal{E}} \circ \mathrm{Exp}(\boldsymbol{\tau}_0)$ in the vicinity of $\mathcal{X}_0$ for which the cost function norm $\|\boldsymbol{\mathcal{E}}(\boldsymbol{\tau})\| \in \mathbb{R}$ holds a local minimum. As the $\boldsymbol{\tau}^{\mathcal{E}\wedge}$ perturbation belongs to the spatial tangent space $\mathrm{T}_{\mathcal{E}}\mathcal{G}$, all gradient descent methods advance the solution by means of (189):

$$\mathcal{X}_{\mathrm{k}+1} \longleftarrow \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} \boxplus \mathcal{X}_{\mathrm{k}} = \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} \circ \mathrm{Exp}(\boldsymbol{\tau}_{\mathrm{k}}) \tag{189}$$

A robust formulation is necessary to ensure that the transform vector τ, which is treated as Euclidean by the functions $\boldsymbol{\mathcal{E}}$ and $\mathbf{f}$, is advanced as a member of the tangent space that adheres to the Lie group constraints, as guaranteed by the use of the left jacobian inverse $\mathbf{J}_{\mathrm{L}}^{-1}(\boldsymbol{\tau})$ (table 1). The first step of the Gauss-Newton method hence consists of two first order Taylor expansions, one for the logarithmic map and a second for the function $\mathbf{f}$:

$$\begin{aligned}
\boldsymbol{\mathcal{E}}_{\mathrm{k}+1} &= \mathbf{f}_{\mathrm{k}+1} - \mathbf{f}_{\mathrm{T}} = \mathbf{f}\Big(\mathrm{Log}\big(\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} \circ \mathrm{Exp}(\boldsymbol{\tau}_{\mathrm{k}})\big)\Big) - \mathbf{f}_{\mathrm{T}} \approx \mathbf{f}\Big(\boldsymbol{\tau}_{\mathrm{k}} + \mathbf{J}_{\mathrm{L}}^{-1}\Big|_{\mathrm{Exp}(\boldsymbol{\tau}_{\mathrm{k}})} \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}}\Big) - \mathbf{f}_{\mathrm{T}} \\
&\approx \mathbf{f}_{\mathrm{k}} + \mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{L}}^{-1}\Big|_{\mathrm{Exp}(\boldsymbol{\tau}_{\mathrm{k}})} \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} - \mathbf{f}_{\mathrm{T}} = \boldsymbol{\mathcal{E}}_{\mathrm{k}} + \mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\,\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}}
\end{aligned} \tag{190}$$

The remaining Gauss-Newton steps are modified accordingly:

$$\|\boldsymbol{\mathcal{E}}_{\mathrm{k}+1}\| = \boldsymbol{\mathcal{E}}_{\mathrm{k}+1}^{\mathrm{T}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}+1} = \boldsymbol{\mathcal{E}}_{\mathrm{k}}^{\mathrm{T}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}} + \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}\mathrm{T}}\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]^{\mathrm{T}}\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} + 2\,\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}\mathrm{T}}\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{k}} \tag{191}$$

$$\begin{aligned}
\frac{\partial\|\boldsymbol{\mathcal{E}}_{\mathrm{k}+1}\|}{\partial\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}}} &= 0 \;\longrightarrow\; \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} = -\left[\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]^{\mathrm{T}}\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]\right]^{-1}\left[\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]^{\mathrm{T}}\boldsymbol{\mathcal{E}}_{\mathrm{k}} \;\longrightarrow \\
\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{E}} &= -\left[\mathbf{J}_{\mathrm{Lk}}^{-\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}^{\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Lk}}^{-1}\right]^{-1}\mathbf{J}_{\mathrm{Lk}}^{-\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}^{\mathrm{T}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}}
\end{aligned} \tag{192}$$

If the perturbation $\boldsymbol{\tau}^{\mathcal{X}\wedge}$ instead belongs to the local tangent space $\mathrm{T}_{\mathcal{X}}\mathcal{G}$, it is necessary to employ the right jacobian instead of the left one, resulting in:

$$\mathcal{X}_{\mathrm{k}+1} \longleftarrow \mathcal{X}_{\mathrm{k}} \oplus \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X}} = \mathrm{Exp}(\boldsymbol{\tau}_{\mathrm{k}}) \circ \Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X}} \tag{193}$$

$$\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X}} = -\left[\mathbf{J}_{\mathrm{Rk}}^{-\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}^{\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{Rk}}^{-1}\right]^{-1}\mathbf{J}_{\mathrm{Rk}}^{-\mathrm{T}}\,\mathbf{J}_{\mathrm{k}}^{\mathrm{T}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}} \tag{194}$$

4.7 State Estimation in Lie Groups

The state estimation `EKF` discussion of section 3.3 states that given a continuous time nonlinear Euclidean state system (63) with process noise provided by (64) and (65), together with a series of discrete time nonlinear observations (88) with measurement noise given by (89) and (90), and considering no correlation between both noises (91), it is possible to compute estimations of the Euclidean state and its covariance at the same time points at which the observations are provided, in such a way that the estimation errors (difference with respect to the true state) are zero mean. The estimations are obtained by means of (118) through (124).

Instead of the Euclidean space time varying state vector $\mathbf{x}(\mathrm{t}) \in \mathbb{R}^{\mathrm{m}}$ considered in section 3.3, let's now consider that the continuous time nonlinear state system is composed by a vector $\mathbf{z}(\mathrm{t}) \in \mathbb{R}^{\mathrm{n}}$, an element of a Lie group $\mathcal{X}(\mathrm{t}) \in \mathcal{G}$, and a vector $\mathbf{v}^{\mathcal{X}}(\mathrm{t}) \in \mathbb{R}^{\mathrm{m}}$ representing the velocity of $\mathcal{X}$ as it moves along its manifold, contained in the local or body tangent space $\mathrm{T}_{\mathcal{X}}\mathcal{G}$. As $\mathcal{X}$ is not Euclidean, the direct application of the `EKF` state estimation scheme of section 3.3 would result in the need to continuously reproject the estimated Lie group elements $\mathcal{X}_{\mathrm{k}}$ back to the manifold as otherwise the estimated states would not comply with the Lie group constraints. The repeated deviations and reprojections from and to the manifold may result in a significant degradation in the estimation accuracy.

The most rigorous and precise way to adapt the EKF scheme is to exclude the Lie group element $\mathcal{X} \in \mathcal{G}$ from the state system, replacing it by a local tangent space perturbation $\Delta\boldsymbol{\tau}^{\mathcal{X}} \in \mathrm{T}_{\mathcal{X}}\mathcal{G}$. Each filter step now consists on estimating the Lie group element $\hat{\mathcal{X}}_{\mathrm{k}}^{+} = \hat{\mathcal{X}}^{+}\left(\mathrm{t_k}\right) \in \mathcal{G}$, the state vector $\hat{\mathbf{x}}_{\mathrm{k}}^{+} = \hat{\mathbf{x}}^{+}\left(\mathrm{t_k}\right) = \left[\Delta\hat{\boldsymbol{\tau}}_{\mathrm{k}}^{\mathcal{X}+},\, \hat{\mathbf{v}}_{\mathrm{k}}^{\mathcal{X}+}, \hat{\mathbf{z}}_{\mathrm{k}}^{+}\right]^{\mathrm{T}} \in \mathbb{R}^{2\,\mathrm{m+n}}$, and its covariance $\mathbf{P}_{\mathrm{k}}^{+} = \mathbf{P}^{+}\left(\mathrm{t_k}\right) \in \mathbb{R}^{(2\,\mathrm{m+n})\mathrm{x}(2\,\mathrm{m+n})}$, based on their values at $\mathrm{t_{k-1}} = \left(\mathrm{k}-1\right)\Delta\mathrm{t}$. For clarity purposes the Lie velocity $\mathbf{v}^{\mathcal{X}}$ and the state vector $\mathbf{z}$ can be grouped into a bigger Euclidean state vector $\mathbf{p} = \left[\mathbf{v}^{\mathcal{X}},\, \mathbf{z}\right]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{m+n}}$, so $\hat{\mathbf{x}}_{\mathrm{k}}^{+} = \left[\Delta\hat{\boldsymbol{\tau}}_{\mathrm{k}}^{\mathcal{X}},\, \hat{\mathbf{p}}_{\mathrm{k}}\right]^{\mathrm{T}}$.

Considering with no loss of generality that the local tangent state perturbation $\Delta\boldsymbol{\tau}^{\mathcal{X}}$ is located on the first positions of the combined state vector $\mathbf{x}$, the definition of the covariance matrix is a combination of that of its Euclidean components (87) and its local Lie counterparts (153), with additional combined members:

$$\mathbf{P}_{\mathrm{k}} = \begin{bmatrix} \mathbf{C}_{\mathcal{X}\mathcal{X},\mathrm{k}}^{\mathcal{X}} & \mathbf{C}_{\mathcal{X}\mathrm{p},\mathrm{k}}^{\mathcal{X}} \\ \mathbf{C}_{\mathrm{p}\mathcal{X},\mathrm{k}}^{\mathcal{X}} & \mathbf{C}_{\mathrm{pp},\mathrm{k}} \end{bmatrix} \in \mathbb{R}^{(2\,\mathrm{m+n})\mathrm{x}(2\,\mathrm{m+n})} \tag{195}$$

$$\mathbf{C}_{\mathcal{X}\mathcal{X},\mathrm{k}}^{\mathcal{X}} = \mathrm{E}\left[\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X}}\,\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X},\mathrm{T}}\right] = \mathrm{E}\left[\left(\mathcal{X}_{\mathrm{k}} \ominus \boldsymbol{\mu}_{\mathcal{X},\mathrm{k}}\right)\,\left(\mathcal{X}_{\mathrm{k}} \ominus \boldsymbol{\mu}_{\mathcal{X},\mathrm{k}}\right)^{\mathrm{T}}\right] \in \mathbb{R}^{\mathrm{mxm}} \tag{196}$$

$$\mathbf{C}_{\mathcal{X}\mathrm{p},\mathrm{k}}^{\mathcal{X}} = \mathrm{E}\left[\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X}}\,\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)^{\mathrm{T}}\right] = \mathrm{E}\left[\left(\mathcal{X}_{\mathrm{k}} \ominus \boldsymbol{\mu}_{\mathcal{X},\mathrm{k}}\right)\,\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)^{\mathrm{T}}\right] \in \mathbb{R}^{\mathrm{mx(m+n)}} \tag{197}$$

$$\mathbf{C}_{\mathrm{p}\mathcal{X},\mathrm{k}}^{\mathcal{X}} = \mathrm{E}\left[\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)\,\Delta\boldsymbol{\tau}_{\mathrm{k}}^{\mathcal{X},\mathrm{T}}\right] = \mathrm{E}\left[\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)\,\left(\mathcal{X}_{\mathrm{k}} \ominus \boldsymbol{\mu}_{\mathcal{X},\mathrm{k}}\right)^{\mathrm{T}}\right] \in \mathbb{R}^{\mathrm{(m+n)xm}} \tag{198}$$

$$\mathbf{C}_{\mathrm{pp},\mathrm{k}} = \mathrm{E}\left[\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)\,\left(\mathbf{p}_{\mathrm{k}} - \boldsymbol{\mu}_{\mathrm{p,k}}\right)^{\mathrm{T}}\right] \in \mathbb{R}^{\mathrm{(m+n)x(m+n)}} \tag{199}$$

The following paragraphs do not describe the full state estimation process, but only the changes with respect to the Euclidean case described in section 3.3:

- **Initialization**. The Lie group element $\hat{\mathcal{X}}_{0}^{+}$, Lie velocity $\hat{\mathbf{v}}_{0}^{\mathcal{X}+}$, and Euclidean state vector $\hat{\mathbf{y}}_{0}^{+}$ are initialized with their expected values, while the local tangent space perturbation $\Delta\hat{\boldsymbol{\tau}}_{0}^{\mathcal{X}+}$ is initialized to zero. The covariance of the initial estimation error $\mathbf{P}_{0}^{+}$ represents the uncertainty in the initial estimation $\hat{\mathbf{x}}_{0}^{+}$.

$$\hat{\mathcal{X}}_{0}^{+} = \boldsymbol{\mu}_{\mathcal{X},0} = \mathrm{E}\left[\mathcal{X}_0\right] \tag{200}$$

$$\hat{\mathbf{x}}_{0}^{+} = \boldsymbol{\mu}_{\mathrm{x},0} = \mathrm{E}\left[\mathbf{x}_0\right] = \mathrm{E}\left[\,\left[\mathbf{0}_{\mathrm{m}}, \mathbf{v}_{0}^{\mathcal{X}}, \mathbf{z}_0\right]^{\mathrm{T}}\,\right] \tag{201}$$

$$\mathbf{P}_{0}^{+} = \mathrm{E}\left[\left(\mathbf{x}_0 - \boldsymbol{\mu}_{\mathrm{x},0}\right)\,\left(\mathbf{x}_0 - \boldsymbol{\mu}_{\mathrm{x},0}\right)^{\mathrm{T}}\right] \tag{202}$$

$$\tag{203}$$

- **Time Update Equations**. The first EKF step propagates the state estimation without the use of any observations, and is similar to that described in section 3.3. The (63) continuous time non linear state system is however replaced by (204):

$$\dot{\mathbf{x}}\left(\mathrm{t}\right) = \left[\Delta\dot{\boldsymbol{\tau}}^{\mathcal{X}},\, \dot{\mathbf{v}}^{\mathcal{X}},\, \dot{\mathbf{z}}\right]^{\mathrm{T}} = \mathrm{f}\big(\mathcal{X}\left(\mathrm{t}\right) \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}}\left(\mathrm{t}\right),\, \mathbf{v}^{\mathcal{X}}\left(\mathrm{t}\right),\, \mathbf{z}\left(\mathrm{t}\right),\, \mathbf{u}\left(\mathrm{t}\right),\, \mathbf{w}\left(\mathrm{t}\right),\, \mathrm{t}\big) \tag{204}$$

$$\Delta\dot{\boldsymbol{\tau}}^{\mathcal{X}} = \mathbf{v}^{\mathcal{X}} \tag{205}$$

$$\dot{\mathbf{p}} = \mathrm{f_p}\big(\mathcal{X}\left(\mathrm{t}\right) \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}}\left(\mathrm{t}\right),\, \mathbf{v}^{\mathcal{X}}\left(\mathrm{t}\right),\, \mathbf{z}\left(\mathrm{t}\right),\, \mathbf{u}\left(\mathrm{t}\right),\, \mathbf{w}\left(\mathrm{t}\right),\, \mathrm{t}\big) \tag{206}$$

Linearization of the continuous time system results in the (68) system matrix $\mathbf{A}\left(\mathrm{t}\right) \in \mathbb{R}^{(2\,\mathrm{m+n})\mathrm{x}(2\,\mathrm{m+n})}$, where various blocks are likely to be based on the $\oplus$ and $+$ jacobians of table 1. Its discretization by means of (81) leads to the system state transition matrix $\mathbf{F}_{\mathrm{k}} \in \mathbb{R}^{(2\,\mathrm{m+n})\mathrm{x}(2\,\mathrm{m+n})}$. With no other differences with respect to the section 3.3 Euclidean process, the manifold element is left unchanged, while the a priori state vector and error covariance are obtained by means of (208) and (209):

$$\hat{\mathcal{X}}_{\mathrm{k}}^{-} = \hat{\mathcal{X}}_{\mathrm{k-1}}^{+} \tag{207}$$

$$\hat{\mathbf{x}}_{\mathrm{k}}^{-} = \hat{\mathbf{x}}_{\mathrm{k-1}}^{+} + \Delta\mathrm{t}\cdot\mathrm{f}\left(\hat{\mathcal{X}}_{\mathrm{k-1}}^{+} \oplus \Delta\hat{\boldsymbol{\tau}}_{\mathrm{k-1}}^{\mathcal{X}+},\, \hat{\mathbf{v}}_{\mathrm{k-1}}^{\mathcal{X}+},\, \hat{\mathbf{z}}_{\mathrm{k-1}}^{+},\, \mathbf{u}_{\mathrm{k-1}},\, \mathbf{0},\, \mathrm{t_{k-1}}\right) \tag{208}$$

$$\mathbf{P}_{\mathrm{k}}^{-} = \mathbf{F}_{\mathrm{k-1}}\,\mathbf{P}_{\mathrm{k-1}}^{+}\,\mathbf{F}_{\mathrm{k-1}}^{\mathrm{T}} + \widetilde{\mathbf{Q}}_{\mathrm{d,k-1}} \tag{209}$$

- **Measurement Update Equations**. The second EKF step updates the estimations by means of the observations, and is very similar to that described in section 3.3. The (88) discrete time non linear observation system is however replaced by (210):

$$\mathbf{y}_k = h\left(\mathcal{X}_k \oplus \Delta\boldsymbol{\tau}_k^{\mathcal{X}},\, \mathbf{v}_k^{\mathcal{X}},\, \mathbf{z}_k,\ \mathbf{v}_k,\, t_k\right) \tag{210}$$

Its linearization results in the (94) output matrix $\mathbf{H}_k \in \mathbb{R}^{qx(2\,m+n)}$, where it is also likely for various blocks to be based on the $\oplus$ and $+$ jacobians of table 1. With no other differences with respect to the section 3.3 Euclidean process, the manifold element is again left unchanged, while the a posteriori state vector and error covariance are obtained by means of (213) and (214):

$$\mathbf{K}_k = \mathbf{P}_k^- \, \mathbf{H}_k^T \left(\mathbf{H}_k \, \mathbf{P}_k^- \, \mathbf{H}_k^T + \widetilde{\mathbf{R}}_k\right)^{-1} \tag{211}$$

$$\hat{\mathcal{X}}_k^+ = \hat{\mathcal{X}}_k^- \tag{212}$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left[\mathbf{y}_k - h\left(\hat{\mathcal{X}}_k^- \oplus \Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{X}-},\, \hat{\mathbf{v}}_k^{\mathcal{X}-},\, \hat{\mathbf{z}}_k^-,\, \mathbf{0},\, t_k\right)\right] \tag{213}$$

$$\mathbf{P}_k^+ = \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k\right) \mathbf{P}_k^- \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k\right)^T + \mathbf{K}_k \, \widetilde{\mathbf{R}}_k \, \mathbf{K}_k^T \tag{214}$$

- **Reset Equations**. The third EKF step, which is not necessary for purely Euclidean systems, resets the a posteriori estimation of the tangent space perturbation $\Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{X}+}$ to zero while modifying the a posteriori estimations for the Lie group element $\hat{\mathcal{X}}_k^+$ and the error covariance $\mathbf{P}_k^+$ accordingly[24]. Note that the accuracy of the linearizations present in the other two steps, which result in the $\mathbf{A}(t)$ and $\mathbf{H}_k$ system and output matrices, is based on the first order Taylor expansions present in table 1, which are directly related to the size of the tangent space perturbations. Although it is not strictly necessary to execute this step in every EKF cycle, the accuracy of the whole state estimation process is improved by maintaining the perturbations as small as possible, so it is recommended to never bypass the reset step.

  Taking into account that the Lie group element is going to be updated per (217), the error covariance is propagated to the new Lie group element per (157) as follows:

$$\mathbf{P}_k^+ \longleftarrow \mathbf{D} \, \mathbf{P}_k^+ \, \mathbf{D}^T \tag{215}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{J}_{\oplus\,\mathcal{X}}^{\oplus\,\hat{\mathcal{X}}_k^+ \oplus \Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{X}+}} & \mathbf{0}_{mx(m+n)} \\ \mathbf{0}_{(m+n)xm} & \mathbf{I}_{(m+n)x(m+n)} \end{bmatrix} \in \mathbb{R}^{(2\,m+n)x(2\,m+n)} \tag{216}$$

  Once propagated, the Lie group element is updated with the local tangent space perturbation, which is itself reset to zero:

$$\hat{\mathcal{X}}_k^+ \longleftarrow \hat{\mathcal{X}}_k^+ \oplus \Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{X}+} \tag{217}$$

$$\Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{X}+} \longleftarrow \mathbf{0}_m \tag{218}$$

Although the above process has been described making use of a local tangent space perturbation $\Delta\boldsymbol{\tau}^{\mathcal{X}} \in T_{\mathcal{X}}\mathcal{G}$ and a vector $\mathbf{v}^{\mathcal{X}} \in \mathbb{R}^m$ also viewed in the local tangent space that represents the velocity of $\mathcal{X} \in \mathcal{G}$ as it moves along its manifold, there exists an equivalent formulation that employs perturbations and velocities viewed in the manifold global tangent space, this is, $\Delta\boldsymbol{\tau}^{\mathcal{E}} \in T_{\mathcal{E}}\mathcal{G}$ and $\mathbf{v}^{\mathcal{E}} \in \mathbb{R}^m$. In this case, the filter estimates the Lie group element $\hat{\mathcal{X}}_k^+ \in \mathcal{G}$, the state vector $\hat{\mathbf{x}}_k^+ = \left[\Delta\hat{\boldsymbol{\tau}}_k^{\mathcal{E}+},\, \hat{\mathbf{v}}_k^{\mathcal{E}+},\, \hat{\mathbf{z}}_k^+\right]^T \in \mathbb{R}^{2\,m+n}$, and the error covariance $\mathbf{P}_k^+$:

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{C}_{\mathcal{X}\mathcal{X},k}^{\mathcal{E}} & \mathbf{C}_{\mathcal{X}p,k}^{\mathcal{E}} \\ \mathbf{C}_{p\mathcal{X},k}^{\mathcal{E}} & \mathbf{C}_{pp,k} \end{bmatrix} \in \mathbb{R}^{(2\,m+n)x(2\,m+n)} \tag{219}$$

$$\mathbf{C}_{\mathcal{X}\mathcal{X},k}^{\mathcal{E}} = E\left[\Delta\boldsymbol{\tau}_k^{\mathcal{E}} \, \Delta\boldsymbol{\tau}_k^{\mathcal{E},T}\right] = E\left[\left(\mathcal{X}_k \boxminus \boldsymbol{\mu}_{\mathcal{X},k}\right) \left(\mathcal{X}_k \boxminus \boldsymbol{\mu}_{\mathcal{X},k}\right)^T\right] \in \mathbb{R}^{mxm} \tag{220}$$

$$\mathbf{C}_{\mathcal{X}p,k}^{\mathcal{E}} = E\left[\Delta\boldsymbol{\tau}_k^{\mathcal{E}} \left(\mathbf{p}_k - \boldsymbol{\mu}_{p,k}\right)^T\right] = E\left[\left(\mathcal{X}_k \boxminus \boldsymbol{\mu}_{\mathcal{X},k}\right) \left(\mathbf{p}_k - \boldsymbol{\mu}_{p,k}\right)^T\right] \in \mathbb{R}^{mx(m+n)} \tag{221}$$

$$\mathbf{C}_{p\mathcal{X},k}^{\mathcal{E}} = E\left[\left(\mathbf{p}_k - \boldsymbol{\mu}_{p,k}\right) \Delta\boldsymbol{\tau}_k^{\mathcal{E},T}\right] = E\left[\left(\mathbf{p}_k - \boldsymbol{\mu}_{p,k}\right) \left(\mathcal{X}_k \boxminus \boldsymbol{\mu}_{\mathcal{X},k}\right)^T\right] \in \mathbb{R}^{(m+n)xm} \tag{222}$$

[24] The a posteriori estimations for the Lie velocity $\hat{\mathbf{v}}_k^{\mathcal{X}+}$ and the Euclidean components $\hat{\mathbf{z}}_k^+$ are not modified in this step.

The only additional changes are the use of $(\Delta\boldsymbol{\tau}^{\mathcal{E}} \boxplus \mathcal{X})$ instead of $(\mathcal{X} \oplus \Delta\boldsymbol{\tau}^{\mathcal{X}})$, the fact that the blocks of the $\mathbf{A}\,(\mathrm{t})$ and $\mathbf{H}_{\mathrm{k}}$ matrices are now based on the $\boxplus$ and $+$ jacobians of table 1, and the use of $\mathbf{J}^{\boldsymbol{\tau} \boxplus \mathcal{X}}_{\boxplus\,\mathcal{X}}\Big|_{\Delta\hat{\boldsymbol{\tau}}_{\mathrm{k}}^{\mathcal{E}+} \boxplus \hat{\mathcal{X}}_{\mathrm{k}}^{+}}$ to propagate the covariance.

5 Rotation of Rigid Bodies

A *rigid body* is an object in which the distance between any two of its points is constant, as is the orientation between any two of its vectors (refer to section 4.2 for the definition of points and vectors in $\mathbb{R}^3$). Rotational rigid body motion is that in which one of its points, named the *center of rotation* $\mathbf{O}_{\mathrm{CR}}$, does not move. Rotations of rigid bodies do not comply with the axioms of an Euclidean space (section 4.1) but with those of Lie groups (section 4.3), and hence this section heavily relies on the concepts of Lie theory discussed in sections 4.3 and 4.4. Table 2 provides a comparison between the generic nomenclature employed in section 4 and their rotation equivalents. The different representations discussed in this section are summarized in Table 3.

Concept	**Lie Theory**	**Rotation**	**Concept**	**Lie Theory**	**Rotation**
Lie group	$\mathcal{G}$	$\mathbb{SO}(3)$	Lie group elements	$\mathcal{X}$, $\mathcal{Y}$	$\mathcal{R}$, $\mathcal{S}$
Concatenation	$\circ$	$\circ$	Lie algebra	$\mathfrak{m}$	$\mathfrak{so}(3)$
Identity	$\mathcal{E}$	$\mathcal{I}_{\mathcal{R}}$	Inverse	$\mathcal{X}^{-1}$	$\mathcal{R}^{-1}$
Velocity	$\mathbf{v}$	$\boldsymbol{\omega}$	Tangent element	$\boldsymbol{\tau}$	$\mathbf{r}$
Local frame	$\mathcal{X}$	B	Global frame	$\mathcal{E}$	N
Point action	$\mathrm{g}_{\mathcal{X}}()$	$\mathbf{g}_{\mathcal{R}}(\mathbf{p})$	Vector action	$\mathrm{g}_{\mathcal{X}}()$	$\mathbf{g}_{\mathcal{R}*}(\mathbf{v})$
Adjoint	$\mathbf{Ad}_{\mathcal{X}}\,(\boldsymbol{\tau}^\wedge)$	$\mathbf{Ad}_{\mathcal{R}}\,(\mathbf{r}^\wedge)$	Adjoint matrix	$\mathbf{Ad}_{\mathcal{X}}\,\boldsymbol{\tau}$	$\mathbf{Ad}_{\mathcal{R}}\,\mathbf{r}$

Table 2: Comparison between generic Lie elements and those of rigid body rotations

This section begins with an introduction to rotational motion in section 5.1, followed by a description of the different rotation Lie group representations: the rotation matrix (section 5.2), the rotation vector (section 5.3), the unit quaternion (section 5.4), the half rotation vector (section 5.5), and the Euler angles (section 5.6). Algebraic operations on rigid body rotations, such as powers, linear interpolation, and the plus and minus operators, are introduced in section 5.7. Section 5.8 presents the rotation time derivative that leads to the definition of the angular velocity in the tangent space. The velocity of the rigid body points is discussed in section 5.9, followed by the adjoint map in section 5.10, which transforms elements of the tangent space while the rotation advances on its manifold, and by an analysis of uncertainty and covariances applied to rotation motion (section 5.11). An extensive analysis of the rotation jacobians is presented in section 5.12. Sections 5.13, 5.14, and 5.15 apply the discrete integration of Lie groups, the Gauss-Newton optimization of Lie group functions, and the state estimation of Lie groups contained in sections 4.5, 4.6, and 4.7 to the case of rotations. Finally, the advantages and disadvantages of each rotation representation are discussed in section 5.16.

5.1 Special Orthogonal (Lie) Group

A rigid body can be represented with a cartesian frame attached to any of its points (the origin), with the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ being simply unit vectors along the main axes. It can be assumed with no loss of generality that the frame origin coincides with the center of rotation. Rigid body rotations can be combined and reversed, complying with the algebraic concept of group, but are not endowed with a metric, so they are not part of a metric or Euclidean space (section 4.1). They do however comply with the axioms of a Lie group (section 4.3), and hence the set of rigid body rotations together with the operation of rotation concatenation comprises $\langle\mathbb{SO}(3), \circ\rangle$, known as the *rotation group* or *special orthogonal group* of $\mathbb{R}^3$ [26], where its elements are denoted by $\mathcal{R}$, the identify rotation by $\mathcal{I}_{\mathcal{R}}$, and the inverse by $\mathcal{R}^{-1}$. The rotation group has two main actions, which are the rotation of points $\left\{\mathbf{g}() : \mathbb{SO}(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \mid \mathbf{p} \rightarrow \mathbf{g}_{\mathcal{R}}\,(\mathbf{p})\right\}$ and that of vectors $\left\{\mathbf{g}_*() : \mathbb{SO}(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \mid \mathbf{v} \rightarrow \mathbf{g}_{\mathcal{R}*}\,(\mathbf{v})\right\}$.

Based on the rigid body definition above, its motion corresponds to an *orthogonal transformation*, this is, one that preserves the norm (maintaining distances between points as well as angles between vectors) and the cross product (maintaining orientation)[25] [22]. These are called *orthogonality* and *handedness* [26]:

- Norm: $\|\mathbf{g}_{\mathcal{R}*}(\mathbf{v})\| = \|\mathbf{v}\|, \forall\, \mathbf{v} \in \mathbb{R}^3$
- Cross product: $\mathbf{g}_{\mathcal{R}*}(\mathbf{u}) \times \mathbf{g}_{\mathcal{R}*}(\mathbf{v}) = \mathbf{g}_{\mathcal{R}*}(\mathbf{u} \times \mathbf{v}), \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^3$

Noting that a vector represents the difference between two points, $\mathbf{g}_{\mathcal{R}*}(\mathbf{v}) = \mathbf{g}_{\mathcal{R}*}(\mathbf{q} - \mathbf{p}) = \mathbf{g}_{\mathcal{R}}(\mathbf{q}) - \mathbf{g}_{\mathcal{R}}(\mathbf{p})$, and considering the possibility that one of the points may be the origin, $\mathbf{g}_{\mathcal{R}}(\mathbf{p}) = \mathbf{g}_{\mathcal{R}}(\mathbf{0}) = \mathbf{0}$, results in the equivalence between the point and vector rotation maps:

$$\mathbf{g}_{\mathcal{R}*}(\mathbf{q} - \mathbf{0}) = \mathbf{g}_{\mathcal{R}*}(\mathbf{q}) = \mathbf{g}_{\mathcal{R}}(\mathbf{q}) - \mathbf{g}_{\mathcal{R}}(\mathbf{0}) = \mathbf{g}_{\mathcal{R}}(\mathbf{q}) \rightarrow \mathbf{g}_{\mathcal{R}}() = \mathbf{g}_{\mathcal{R}*}() \tag{223}$$

Representation	**Symbol**	**Structure**	**Space**
Rotation matrix	$\mathbf{R}$	orthogonal 3x3 matrix	$\mathbb{SO}(3)$
Angular velocity	$\boldsymbol{\omega}^\wedge = \widehat{\boldsymbol{\omega}}$	skew-symmetric matrix	$\mathfrak{so}(3)$
	$\boldsymbol{\omega}$	free 3-vector	
Rotation vector	$\mathbf{r}^\wedge = \widehat{\mathbf{r}}$	skew-symmetric matrix	$\mathbb{SO}(3)$ & $\mathfrak{so}(3)$
	$\mathbf{r} = \boldsymbol{\omega}\,\mathrm{t} = \mathbf{n}\,\phi$	free 3-vector	
Unit quaternion	$\mathbf{q}$	unit quaternion	$\mathbb{SO}(3)$
Half angular velocity	$\boldsymbol{\Omega}^\wedge$	pure quaternion	$\mathfrak{so}(3)$
	$\boldsymbol{\Omega} = \boldsymbol{\omega}/2$	free 3-vector	
Half rotation vector	$\mathbf{h}^\wedge$	pure quaternion	$\mathbb{SO}(3)$ & $\mathfrak{so}(3)$
	$\mathbf{h} = \boldsymbol{\Omega}\,\mathrm{t} = \mathbf{n}\,\theta = \mathbf{r}/2$	free 3-vector	
Euler angles	$\boldsymbol{\phi}$	3 angles	$\mathbb{SO}(3)$

Table 3: Summary of rotational motion representations

The $\mathbb{SO}(3)$ analysis below adopts the convention introduced in section 4.3, in which all actions, including concatenation $\{\circ : \mathbb{SO}(3) \times \mathbb{SO}(3) \rightarrow \mathbb{SO}(3)\}$, transform elements viewed in the local or body frame $\mathrm{F_B} = \{\mathrm{O_B}, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ into elements viewed in the global or spatial frame $\mathrm{F_N} = \{\mathrm{O_N}, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\} = \{\mathbf{g}_{\mathcal{R}}(\mathrm{O_B}), \mathbf{g}_{\mathcal{R}*}(\mathbf{b}_1), \mathbf{g}_{\mathcal{R}*}(\mathbf{b}_2), \mathbf{g}_{\mathcal{R}*}(\mathbf{b}_3)\}$[26]:

$$\mathbf{p}^{\mathrm{N}} = \mathbf{g}_{\mathcal{R}_{\mathrm{NB}}}(\mathbf{p}^{\mathrm{B}}) \tag{224}$$

$$\mathbf{v}^{\mathrm{N}} = \mathbf{g}_{\mathcal{R}*_{\mathrm{NB}}}(\mathbf{v}^{\mathrm{B}}) \tag{225}$$

5.2 Rotation Matrix

The three basis vectors of the output frame can be stacked side by side into a matrix $\mathbf{R} = \mathbf{R}_{\mathrm{NB}} = [\mathbf{n}_1\ \mathbf{n}_2\ \mathbf{n}_3] \in \mathbb{R}^{3\mathrm{x}3}$, called the *rotation matrix*. Since its columns form a right handed orthonormal basis, it complies with the orthogonality and handedness conditions, and it can be proven that the rotation matrix $\mathbf{R}$ is an special orthogonal matrix[27]. Rotation matrices hence represent rigid body rotations, and their space $\mathbb{SO}(3) = \{\mathbf{R} \in \mathbb{R}^{3\mathrm{x}3} \mid \mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}_3,\ \det(\mathbf{R}) = +1\}$ has group structure under matrix multiplication $\{\mathbb{R}^{3\mathrm{x}3} \times \mathbb{R}^{3\mathrm{x}3} \rightarrow \mathbb{R}^{3\mathrm{x}3} \mid \mathbf{R}_{\mathrm{a}}\,\mathbf{R}_{\mathrm{b}} \in \mathbb{R}^{3\mathrm{x}3}, \forall\, \mathbf{R}_{\mathrm{a}},\ \mathbf{R}_{\mathrm{b}} \in \mathbb{R}^{3\mathrm{x}3}\}$ [15]. While having dimension nine, the special orthogonal group $\mathbb{SO}(3)$ defined by means of rotation matrices constitutes a three dimensional manifold to euclidean space $\mathbb{E}^3$. Note that in this group the identity element is given by the identity matrix ($\mathbf{I} = \mathbf{I}_3$), and the inverse coincides with the transpose $\left(\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}\right)$.

Concept	$\mathbb{SO}^3$	**Rotation Matrix**	**Concept**	$\mathbb{SO}^3$	**Rotation Matrix**
Lie group element	$\mathcal{R}$	$\mathbf{R}$	Concatenation	$\circ$	Matrix product
Identity	$\mathcal{I}_{\mathcal{R}}$	$\mathbf{I}_3$	Inverse	$\mathcal{R}^{-1}$	$\mathbf{R}^{\mathrm{T}}$
Point rotation	$\mathbf{g}_{\mathcal{R}}(\mathbf{p})$	$\mathbf{R}\,\mathbf{p}$	Vector rotation	$\mathbf{g}_{\mathcal{R}*}(\mathbf{v})$	$\mathbf{R}\,\mathbf{v}$

[25] An orthogonal transformation can also be defined as one that preserves both the inner and cross products.
[26] The local and spatial bases are denoted B and N as they usually correspond to the body and NED frames respectively.
[27] Orthogonal means that the transpose equals the inverse, while special or proper means that the determinant is positive one.

Table 4: Comparison between generic $\mathbb{SO}(3)$ and rotation matrix

The rotation matrix $\mathbf{R}$ represents the actual coordinate transformation from the local to the global frame:

$$\mathbf{g}_{\mathcal{R}*}(\mathbf{v}) = \mathbf{R}\,\mathbf{v} \tag{226}$$

The inverse rotation (from the global to the local frame), is simply the transpose:

$$\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}} \tag{227}$$

The concatenation of rotations is also straight forward as it coincides with matrix multiplication. Note that $\mathbb{SO}(3)$ as defined above is not an abelian group, so the order of the factors is important.

$$\mathbf{R}_{\mathrm{EB}} = \mathbf{R}_{\mathrm{EN}}\,\mathbf{R}_{\mathrm{NB}} \tag{228}$$

5.3 Rotation Vector as Tangent Space

As discussed in section 4.3.1, the structure of the Lie algebra associated to $\mathbb{SO}(3)$ can be obtained by time derivating the Lie group inverse constraint, $\mathbf{R}^{\mathrm{T}}(\mathrm{t})\ \mathbf{R}(\mathrm{t}) = \mathbf{R}(\mathrm{t})\ \mathbf{R}^{\mathrm{T}}(\mathrm{t}) = \mathbf{I}_3$, resulting in the following particularizations of (128) and (129):

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}\wedge} = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}} \;=\; \dot{\mathbf{R}}_{\mathrm{NB}}\,\mathbf{R}_{\mathrm{NB}}^{\mathrm{T}} = -\mathbf{R}_{\mathrm{NB}}\,\dot{\mathbf{R}}_{\mathrm{NB}}^{\mathrm{T}} \tag{229}$$

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}\wedge} = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}} \;=\; \mathbf{R}_{\mathrm{NB}}^{\mathrm{T}}\,\dot{\mathbf{R}}_{\mathrm{NB}} = -\dot{\mathbf{R}}_{\mathrm{NB}}^{\mathrm{T}}\,\mathbf{R}_{\mathrm{NB}} \tag{230}$$

The Lie algebra velocity $\mathbf{v}^{\wedge}$ of $\mathbb{SO}(3)$ is known as the *angular velocity* $\boldsymbol{\omega}^{\wedge}$, and as shown in (229) and (230), has the structure of a skew-symmetric matrix because its negative coincides with its transpose, so it is generally denoted as $\widehat{\boldsymbol{\omega}}$. An alternative definition of the angular velocity is presented in section 5.8. Inverting the previous equations results in the rotation matrix time derivative, which is linear:

$$\dot{\mathbf{R}}_{\mathrm{NB}} = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}\,\mathbf{R}_{\mathrm{NB}} = \mathbf{R}_{\mathrm{NB}}\,\widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}} \tag{231}$$

Notice that if $\mathbf{R}(\mathrm{t}_0) = \mathbf{I}_3$, then $\dot{\mathbf{R}}(\mathrm{t}_0) = \widehat{\boldsymbol{\omega}}(\mathrm{t}_0)$, and hence the skew-symmetric matrix $\widehat{\boldsymbol{\omega}}(\mathrm{t}_0)$ provides a first order approximation of the rotation matrix around the identity matrix $\mathbf{I}_3$:

$$\mathbf{R}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \mathbf{I}_3 + \widehat{\boldsymbol{\omega}}(\mathrm{t}_0)\ \Delta\mathrm{t} \tag{232}$$

The *space of skew-symmetric matrices* $\mathfrak{so}(3) = \{\widehat{\boldsymbol{\omega}} \in \mathbb{R}^{3\mathrm{x}3} \mid \boldsymbol{\omega} \in \mathbb{R}^3,\ -\widehat{\boldsymbol{\omega}} = \widehat{\boldsymbol{\omega}}^{\mathrm{T}}\}$ is hence the *tangent space* of $\mathbb{SO}(3)$ at the identity $\mathbf{I}_3$ [22], denoted as $\mathrm{T}_{\mathbf{I}_3}\mathcal{R}$. The *hat* $\left\{\cdot^{\wedge} : \mathbb{R}^3 \to \mathfrak{so}(3) \mid \boldsymbol{\omega} \to \boldsymbol{\omega}^{\wedge} = \widehat{\boldsymbol{\omega}}\right\}$ and *vee* $\left\{\cdot^{\vee} : \mathfrak{so}(3) \to \mathbb{R}^3 \mid \left(\widehat{\boldsymbol{\omega}}^{\vee} \to \boldsymbol{\omega}\right)\right\}$ operators convert the cartesian vector form of the angular velocity into its skew-symmetric form, and viceversa.

If $\mathbf{R}(\mathrm{t}_0) \neq \mathbf{I}_3$, the tangent space needs to be transported right multiplying by $\mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0)$ (in the case of space angular velocity), or left multiplying for the local based velocity:

$$\mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0 + \Delta\mathrm{t}) \;\approx\; \mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0) + \left[\widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}(\mathrm{t}_0)\ \Delta\mathrm{t}\right]\mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0) = \left[\mathbf{I}_3 + \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}(\mathrm{t}_0)\ \Delta\mathrm{t}\right]\mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0) \tag{233}$$

$$\mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0 + \Delta\mathrm{t}) \;\approx\; \mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0) + \mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0)\left[\widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}}(\mathrm{t}_0)\ \Delta\mathrm{t}\right] = \mathbf{R}_{\mathrm{NB}}(\mathrm{t}_0)\left[\mathbf{I}_3 + \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}}(\mathrm{t}_0)\ \Delta\mathrm{t}\right] \tag{234}$$

Note that the solution to the ordinary differential equation $\dot{\mathbf{x}}(\mathrm{t}) = \mathbf{x}(\mathrm{t})\ \widehat{\boldsymbol{\omega}},\ \mathbf{x}(\mathrm{t}) \in \mathbb{R}^3$, where $\widehat{\boldsymbol{\omega}}$ is constant, is $\mathbf{x}(\mathrm{t}) = \mathbf{x}(0)\ \mathrm{e}^{\widehat{\boldsymbol{\omega}}\mathrm{t}}$. Based on it, assuming $\mathbf{R}(0) = \mathbf{I}_3$ as initial condition, and considering for the time being that $\widehat{\boldsymbol{\omega}}$ is constant,

$$\mathbf{R}(\mathrm{t}) = \mathrm{e}^{\widehat{\boldsymbol{\omega}}\mathrm{t}} = \mathbf{I}_3 + \widehat{\boldsymbol{\omega}}\mathrm{t} + \frac{(\widehat{\boldsymbol{\omega}}\mathrm{t})^2}{2!} + \cdots + \frac{(\widehat{\boldsymbol{\omega}}\mathrm{t})^{\mathrm{n}}}{\mathrm{n}!} + \ldots \tag{235}$$

which is indeed a rotation matrix as it complies with the $\mathbb{SO}(3)$ conditions of orthogonality and handedness [22].

Concept	Lie Theory	$\mathbb{SO}^3$
Tangent space element	$\boldsymbol{\tau}^\wedge$	$\mathbf{r}^\wedge = \widehat{\mathbf{r}}$
Velocity element	$\mathbf{v}^\wedge$	$\boldsymbol{\omega}^\wedge = \widehat{\boldsymbol{\omega}}$
Structure	$\wedge$	skew symmetric matrix

Table 5: Comparison between generic $\mathbb{SO}(3)$ and rotation vector as tangent space

Remembering that so far $\widehat{\boldsymbol{\omega}}$ is constant, (235) means that any rotation $\mathbf{R}\,(\mathrm{t}) = \mathrm{e}^{\widehat{\boldsymbol{\omega}}\mathrm{t}}$ can be realized by maintaining a constant angular velocity $\boldsymbol{\omega}$ for a given time t. This is analogous to stating that any angular displacement $\mathbf{R}\,(\phi) = \mathrm{e}^{\widehat{\mathbf{n}}\phi}$ can be achieved by rotating an angle ϕ about a fixed unitary rotation axis $\mathbf{n}$, which enables the definition of the *rotation vector* $\mathbf{r}$, also known as the *exponential coordinates* of the $\mathcal{R}$ rotation, as

$$\mathbf{r} = \boldsymbol{\omega}\,\mathrm{t} = \mathbf{n}\,\phi \in \mathbb{R}^3 \tag{236}$$

Note that the rotation vector $\mathbf{r}$ belongs to the tangent space as it is a multiple of the angular velocity $\boldsymbol{\omega} \in \mathfrak{so}\,(3)$, and hence tends to coincide with it as time tends to zero. The *exponential map* $\{\exp\,() : \mathfrak{so}(3) \rightarrow \mathbb{SO}(3) \mid \mathcal{R} = \exp\,(\mathbf{r}^\wedge)\}$ and its capitalized form $\big\{\mathrm{Exp}\,() : \mathbb{R}^3 \rightarrow \mathbb{SO}(3) \mid \mathcal{R} = \mathrm{Exp}\,(\mathbf{r})\big\}$ wrap the rotation vector around the rotation group. In the case of the rotation matrix, the exponential map can be obtained from (235) based on the fact that all skew-symmetric matrices verify that $\widehat{\mathbf{r}}^2 = \mathbf{r}\,\mathbf{r}^{\mathrm{T}} - \mathbf{I}_3$ and $\widehat{\mathbf{r}}^3 = -\widehat{\mathbf{r}}$, converting skew symmetric matrices into orthogonal ones:

$$\mathbf{R}\,(\mathbf{r}) = \exp\,(\widehat{\mathbf{r}}) = \mathrm{Exp}\,(\mathbf{r}) = \mathrm{e}^{\widehat{\mathbf{r}}} = \mathbf{I}_3 + \frac{\widehat{\mathbf{r}}}{\|\mathbf{r}\|}\sin\|\mathbf{r}\| + \frac{\widehat{\mathbf{r}}^2}{\|\mathbf{r}\|^2}\,(1 - \cos\|\mathbf{r}\|) \tag{237}$$

Geometrically, the skew symmetric matrix corresponds to an axis of rotation (via the mapping $\mathbf{n} \rightarrow \widehat{\mathbf{n}}$) and the exponential map generates the rotation corresponding to rotating about that axis by an amount ϕ [27].

The angular velocity $\boldsymbol{\omega}$ however is in fact not required to be constant. Given a rotation matrix $\mathbf{R} \in \mathbb{SO}\,(3)$, it can be proven that there exists a not necessarily unique vector $\mathbf{r} \in \mathbb{R}^3$ such that $\mathbf{R} = \mathrm{e}^{\widehat{\mathbf{r}}}$. The *logarithmic map* $\{\log\,() : \mathbb{SO}(3) \rightarrow \mathfrak{so}(3) \mid \mathbf{r}^\wedge = \log\,(\mathcal{R})\}$ and its capitalized version $\big\{\mathrm{Log}\,() : \mathbb{SO}(3) \rightarrow \mathbb{R}^3 \mid \mathbf{r} = \mathrm{Log}\,(\mathcal{R})\big\}$ hence convert rigid body rotations into rotation vectors.

$$\mathbf{R} = \mathbf{R}_{\mathrm{NB}} = \begin{bmatrix} \mathrm{R}_{11} & \mathrm{R}_{12} & \mathrm{R}_{13} \\ \mathrm{R}_{21} & \mathrm{R}_{22} & \mathrm{R}_{23} \\ \mathrm{R}_{31} & \mathrm{R}_{32} & \mathrm{R}_{33} \end{bmatrix} \;\rightarrow\; \|\mathbf{r}\| = \arccos\left(\frac{\mathrm{trace}\,(\mathbf{R}) - 1}{2}\right),\; \mathbf{r} = \frac{\|\mathbf{r}\|}{2\sin\|\mathbf{r}\|}\begin{bmatrix} \mathrm{R}_{32} - \mathrm{R}_{23} \\ \mathrm{R}_{13} - \mathrm{R}_{31} \\ \mathrm{R}_{21} - \mathrm{R}_{12} \end{bmatrix} \tag{238}$$

Any rotation matrix can hence be realized by rotating a certain angle about a given axis, as indicated in (236). The vector $\mathbf{n} = \mathbf{r}/\|\mathbf{r}\|$ indicates the rotation direction while $\phi = \|\mathbf{r}\|$ represents the turn angle. The exponential map described by (237) is thus surjective (there is at least one rotation vector for every rotation matrix) but not injective, as a rotation of $(\|\mathbf{r}\| + 2\,\mathrm{k}\,\pi)\;\forall\;\mathrm{k} \in \mathbb{Z}$ about $\mathbf{r}/\|\mathbf{r}\|$ or a rotation of $(-\|\mathbf{r}\| + 2\,\mathrm{k}\,\pi)$ about $-\mathbf{r}/\|\mathbf{r}\|$ produce exactly the same rotation matrix.

Although inverting the rotation by means of the rotation vector is straightforward,

$$\mathbf{r}_{\mathrm{BN}} = {\mathbf{r}_{\mathrm{NB}}}^{-1} = -\mathbf{r}_{\mathrm{NB}} \tag{239}$$

the different $\mathbb{SO}(3)$ actions (concatenation, point rotation, vector rotation), as well as the relationship between the rotation vector derivative with time and the angular velocities, are complex and rarely used.

5.4 Unit Quaternion

The quaternions with unity norm, known as unit quaternions, comprise an additional representation of the rotation group $\mathbb{SO}\,(3)$, as shown below. Quaternions in turn are generalizations of complex numbers in the same way that these are generalizations of real ones [22]. It is hence necessary to first describe the complex numbers in section 5.4.1 and the quaternions in section 5.4.2 before focusing on the unit quaternions in section 5.4.3.

5.4.1 Complex Numbers

The set of *complex numbers* $\mathbb{C}$ is composed of two real numbers $\left\{\mathbb{C} = \mathbb{R} + \mathbb{R}\,\mathrm{i} \mid \mathrm{i}^2 = \mathrm{i} \cdot \mathrm{i} = -1\right\}$. Given two complex numbers $c_1 = x_1 + y_1\,\mathrm{i} \in \mathbb{C}, c_2 = x_2 + y_2\,\mathrm{i} \in \mathbb{C}, \forall\, x_1, y_1, x_2, y_2 \in \mathbb{R}$, it is possible to define the operations of addition $\{+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}\}$ and multiplication $\{\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}\}$.

$$c_1 + c_2 = (x_1 + y_1\,\mathrm{i}) + (x_2 + y_2\,\mathrm{i}) = (x_1 + x_2) + (y_1 + y_2)\,\mathrm{i} \tag{240}$$

$$c_1 \cdot c_2 = c_1\,c_2 = (x_1 + y_1\,\mathrm{i}) \cdot (x_2 + y_2\,\mathrm{i}) = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)\,\mathrm{i} \tag{241}$$

The conjugate is defined as $c^* = x - y\,\mathrm{i} \in \mathbb{C}$ and verifies that $(c_1 \cdot c_2)^* = c_1^* \cdot c_2^*$, while the norm $\|c\| = \sqrt{c \cdot c^*} = \sqrt{c^* \cdot c} = \sqrt{x^2 + y^2} \in \mathbb{R}$ satisfies that $\|c_1 \cdot c_2\| = \|c_1\| \cdot \|c_2\|$. The set of complex numbers $\mathbb{C}$ endowed with the operations of addition $+$ and multiplication $\cdot$ forms a field (not ordered), known as the field of complex numbers $\langle \mathbb{C}, +, \cdot \rangle$, nearly always abbreviated to simply $\mathbb{C}$. The additive identity is $0 = 0 + 0\,\mathrm{i}$ and the inverse $-c = -x - y\,\mathrm{i}$, while the multiplication identity is $1 = 1 + 0\,\mathrm{i}$ and the inverse $c^{-1} = c^*/\|c\|^2$.

Complex numbers can always be written in polar form $\left(c = r\,(\cos\phi + \sin\phi\,\mathrm{i}) = r\,e^{\mathrm{i}\,\phi}\right)$, and as such are valid representations of the circle group or plane rotations group $\mathbb{SO}(2)$, similarly to the case of rotation vectors in $\mathbb{SO}(3)$ described in section 5.3.

5.4.2 Quaternions

The set of *quaternions* $\mathbb{H}$ is defined as $\{\mathbb{H} = \mathbb{C} + \mathbb{C}\,\mathrm{j} \mid \mathrm{j}^2 = -1,\ \mathrm{i} \cdot \mathrm{j} = -\mathrm{j} \cdot \mathrm{i}\}$. A quaternion $\mathbf{q} \in \mathbb{H}$ has the form $\mathbf{q} = q_0 + q_1\,\mathrm{i} + q_2\,\mathrm{j} + q_3\,\mathrm{i}\,\mathrm{j}$, with $q_i \in \mathbb{R}$. *Pure quaternions* $\mathbf{q} = q_1\,\mathrm{i} + q_2\,\mathrm{j} + q_3\,\mathrm{i}\,\mathrm{j} \in \mathbb{H}_p$ are those defined in the tridimensional imaginary subspace of $\mathbb{H}$, and verify that $\mathbf{q} = -\,\mathbf{q}^*$.

There are many different conventions for the quaternion found in the literature [26]. This article adopts the *Hamilton convention*, characterized by locating the real part first (instead of last), being right handed (left), passive (rotates frames and not vectors as in active), and local to global rotations (global to local). Any variation to these choices would result in different expressions below, although the physical concepts do not vary.

The real plus imaginary notation $\{1, \mathrm{i}, \mathrm{j}, \mathrm{i}\,\mathrm{j}\}$ is not always the most convenient. A quaternion can also be expressed as the sum of a scalar plus a vector in the form $\mathbf{q} = q_0 + \mathbf{q}_v$, where q_0 is the real or scalar part and $\mathbf{q}_v = q_1\,\mathrm{i} + q_2\,\mathrm{j} + q_3\,\mathrm{i}\,\mathrm{j}$ is the imaginary or vector part. Quaternions are however mostly represented as 4-vectors $\mathbf{q} = [q_0, \mathbf{q}_v]^T = [q_0, q_1, q_2, q_3]^T$, which enables the usage of matrix algebra for quaternion operations. It is also convenient to abuse the equal operator by combining general, real, and pure quaternions as in $\mathbf{q} = q_0 + \mathbf{q}_v$, where $q_0 = [q_0, \mathbf{0}_v]^T$ and $\mathbf{q}_v = [0, \mathbf{q}_v]^T$.

The following expressions define the addition $\{+ : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}\}$ and inner product $\{\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}\}$ of two quaternions, which commute, as well as the scalar multiplication $\{\cdot : \mathbb{R} \times \mathbb{H} \rightarrow \mathbb{H}\}$:

$$\mathbf{q} + \mathbf{p} = [q_0, \mathbf{q}_v]^T + [p_0, \mathbf{p}_v]^T = [q_0 + p_0, q_1 + p_1, q_2 + p_2, q_3 + p_3]^T = [q_0 + p_0, \mathbf{q}_v + \mathbf{p}_v]^T \tag{242}$$

$$\langle \mathbf{q}, \mathbf{p} \rangle = \mathbf{q} \cdot \mathbf{p} = \mathbf{q}^T\,\mathbf{p} = q_0\,p_0 + q_1\,p_1 + q_2\,p_2 + q_3\,p_3 \tag{243}$$

$$a \cdot \mathbf{q} = a \cdot [q_0, \mathbf{q}_v]^T = [a\,q_0, a \cdot \mathbf{q}_v]^T \tag{244}$$

The multiplication of quaternions $\{\otimes : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}\}$ is not commutative as it includes the cross product:

$$\mathbf{q} \otimes \mathbf{p} = \begin{bmatrix} q_0 \cdot p_0 - q_1 \cdot p_1 - q_2 \cdot p_2 - q_3 \cdot p_3 \\ q_1 \cdot p_0 + q_0 \cdot p_1 - q_3 \cdot p_2 + q_2 \cdot p_3 \\ q_2 \cdot p_0 + q_3 \cdot p_1 + q_0 \cdot p_2 - q_1 \cdot p_3 \\ q_3 \cdot p_0 - q_2 \cdot p_1 + q_1 \cdot p_2 + q_0 \cdot p_3 \end{bmatrix} = \begin{bmatrix} q_0\,p_0 - \mathbf{q}_v^T\,\mathbf{p}_v \\ q_0\,\mathbf{p}_v + p_0\,\mathbf{q}_v + \widehat{\mathbf{q}}_v\,\mathbf{p}_v \end{bmatrix} \tag{245}$$

It is also bilinear [26]:

$$\mathbf{q} \otimes \mathbf{p} = [\mathbf{q}]_\mathrm{L}\, \mathbf{p} = \begin{bmatrix} +q_0 & -q_1 & -q_2 & -q_3 \\ +q_1 & +q_0 & -q_3 & +q_2 \\ +q_2 & +q_3 & +q_0 & -q_1 \\ +q_3 & -q_2 & +q_1 & +q_0 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} = [\mathbf{p}]_\mathrm{R}\, \mathbf{q} = \begin{bmatrix} +p_0 & -p_1 & -p_2 & -p_3 \\ +p_1 & +p_0 & +p_3 & -p_2 \\ +p_2 & -p_3 & +p_0 & +p_1 \\ +p_3 & +p_2 & -p_1 & +p_0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{246}$$

These expressions can be simplified for the common case of the multiplication of a quaternion $\mathbf{q}$ by a pure quaternion $\mathbf{p}_\mathrm{v}$, this is, $\{\otimes : \mathbb{H} \times \mathbb{H}_\mathrm{p} \to \mathbb{H}\}$:

$$\mathbf{q} \otimes \mathbf{p}_\mathrm{v} \quad = \quad \begin{bmatrix} -q_1 \cdot p_1 - q_2 \cdot p_2 - q_3 \cdot p_3 \\ +q_0 \cdot p_1 - q_3 \cdot p_2 + q_2 \cdot p_3 \\ +q_3 \cdot p_1 + q_0 \cdot p_2 - q_1 \cdot p_3 \\ -q_2 \cdot p_1 + q_1 \cdot p_2 + q_0 \cdot p_3 \end{bmatrix} = \begin{bmatrix} -{\mathbf{q}_\mathrm{v}}^T \mathbf{p}_\mathrm{v} \\ q_0\, \mathbf{p}_\mathrm{v} + \widehat{\mathbf{q}}_\mathrm{v}\, \mathbf{p}_\mathrm{v} \end{bmatrix} \tag{247}$$

$$\mathbf{q} \otimes \mathbf{p}_\mathrm{v} \quad = \quad [\mathbf{q}]_\mathrm{L3}\, \mathbf{p}_\mathrm{v} = \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ +q_0 & -q_3 & +q_2 \\ +q_3 & +q_0 & -q_1 \\ -q_2 & +q_1 & +q_0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = [\mathbf{p}_\mathrm{v}]_\mathrm{R3}\, \mathbf{q} = \begin{bmatrix} 0 & -p_1 & -p_2 & -p_3 \\ +p_1 & 0 & +p_3 & -p_2 \\ +p_2 & -p_3 & 0 & +p_1 \\ +p_3 & +p_2 & -p_1 & 0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{248}$$

The conjugate quaternion is defined as $\mathbf{q}^* = q_0 - \mathbf{q}_\mathrm{v} \in \mathbb{H}$ and verifies that $(\mathbf{q} \otimes \mathbf{p})^* = \mathbf{p}^* \otimes \mathbf{q}^*$, while the quaternion norm $\|\mathbf{q}\| = \sqrt{\langle \mathbf{q}\ ,\ \mathbf{q}^* \rangle} = \sqrt{\langle \mathbf{q}^*\ ,\ \mathbf{q} \rangle} = \sqrt{\mathbf{q} \otimes \mathbf{q}^*} = \sqrt{\mathbf{q}^* \otimes \mathbf{q}} \in \mathbb{R}$ satisfies that $\|\mathbf{q} \otimes \mathbf{p}\| = \|\mathbf{p} \otimes \mathbf{q}\| = \|\mathbf{q}\|\, \|\mathbf{p}\|$. Quaternions endowed with $\otimes$ form the non-commutative group $\langle \mathbb{H}, \otimes \rangle$, where $\mathbf{q_1} = 1 + \mathbf{0}_\mathrm{v}$ is the identity and $\mathbf{q}^{-1} = \mathbf{q}^*/\|\mathbf{q}\|^2$ the inverse [26]. Additionally, quaternions endowed with addition $+$ and multiplication $\otimes$ form the ring $\langle \mathbb{H}, +, \otimes \rangle$ where $\mathbf{q_0} = 0 + \mathbf{0}_\mathrm{v}$ is the addition identity and $-\mathbf{q}$ the addition inverse or negative.

The quaternion rotation operator or *double product* is defined as $\{\mathbb{H} \times \mathbb{R}^3 \to \mathbb{R}^3 \mid \mathbf{q} \in \mathbb{H},$ $\mathbf{v} \in \mathbb{R}^3 \to \mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^* \in \mathbb{R}^3\}$[28], while the natural power of a quaternion $\mathbf{q}^n, n \in \mathbb{N}$ is obtained by multiplying the quaternion by itself $n - 1$ times.

5.4.3 Unit Quaternion

Unit quaternions verify that $\|\mathbf{q}\| = 1$, which implies that $\mathbf{q}^{-1} = \mathbf{q}^*$. They can always be written as

$$\mathbf{q} = \cos\theta + \mathbf{n}\, \sin\theta \tag{249}$$

where $\mathbf{n}$ is a unit vector and θ is a scalar. The exponential of a quaternion $e^{\mathbf{q}}$ is defined analogously to that of real numbers [26]. For pure quaternions $\mathbf{q} = \mathbf{q}_\mathrm{v}$, if abusing notation with $\mathbf{q}_\mathrm{v} = \mathbf{v} = \mathbf{n}\, \theta$ where $\theta = \|\mathbf{v}\|$ and $\|\mathbf{n}\| = 1$, it can be proven that $e^{\mathbf{q}_\mathrm{v}} = \cos\theta + \mathbf{n}\, \sin\theta$, which can be considered an extension of the $e^{i\theta} = \cos\theta + i\, \sin\theta$ expression for complex numbers introduced in section 5.4.1 [26]. Notice that since $\|e^{\mathbf{q}_\mathrm{v}}\| = 1$, the exponential of a pure quaternion is a unit quaternion. If $\|\mathbf{q}\| = 1$, it is easy to verify that $\log(\mathbf{q}) = \log\left(e^{\mathbf{n}\, \theta}\right) = \mathbf{q}_\mathrm{v}$, so the logarithm of a unit quaternion is a pure quaternion [26].

Unit quaternions endowed with $\otimes$ constitute a subgroup that represents a 3-sphere, this is, the three dimensional surface of the unit sphere of $\mathbb{R}^4$, and is commonly noted as $\mathbb{S}^3$. They comply with the orthogonality and handedness conditions required in section 5.1 for rigid body rotations, and hence their space $\mathbb{SO}(3) = \{\mathbf{q} \in \mathbb{S}^3 \mid \mathbf{q}^* \otimes \mathbf{q} = \mathbf{q} \otimes \mathbf{q}^* = \mathbf{q_1}\}$ possesses group structure under quaternion multiplication $\{\otimes : \mathbb{S}^3 \times \mathbb{S}^3 \to \mathbb{S}^3 \mid \mathbf{q}_\mathrm{a} \otimes \mathbf{q}_\mathrm{b} \in \mathbb{S}^3, \forall\ \mathbf{q}_\mathrm{a},\ \mathbf{q}_\mathrm{b} \in \mathbb{S}^3\}$ [15]. While having dimension four, the special orthogonal group $\mathbb{SO}(3)$ defined by means of unit quaternions constitutes a three dimensional manifold to euclidean space $\mathbb{E}^3$. Note that in this group $\mathbf{q_1}$ constitutes the identity and $\mathbf{q}^*$ the inverse.

[28] It is easily proven that the double quaternion product results in $\mathbb{R}^3$ and not $\mathbb{R}^4$.

Concept	$\mathbb{SO}^3$	$\mathbb{S}^3$	**Concept**	$\mathbb{SO}^3$	$\mathbb{S}^3$
Lie group element	$\mathcal{R}$	$\mathbf{q}$	Concatenation	$\circ$	$\otimes$
Identity	$\mathcal{I}_{\mathcal{R}}$	$\mathbf{q_1}$	Inverse	$\mathcal{R}^{-1}$	$\mathbf{q}^*$
Point rotation	$\mathbf{g}_{\mathcal{R}}(\mathbf{p})$	$\mathbf{q}\otimes\mathbf{p}\otimes\mathbf{q}^*$	Vector rotation	$\mathbf{g}_{\mathcal{R}*}(\mathbf{v})$	$\mathbf{q}\otimes\mathbf{v}\otimes\mathbf{q}^*$

Table 6: Comparison between generic $\mathbb{SO}(3)$ and unit quaternion

Coordinate transformation (point or vector rotation) and rotation concatenation are both linear:

$$\mathbf{g}_{\mathcal{R}*}(\mathbf{v}) = \mathbf{q}\otimes\mathbf{v}\otimes\mathbf{q}^* \tag{250}$$

$$\mathbf{q}_{\text{EB}} = \mathbf{q}_{\text{EN}}\otimes\mathbf{q}_{\text{NB}} \tag{251}$$

5.5 Half Rotation Vector as Tangent Space

It is interesting to point out that in the case of rotation matrices (section 5.2), the orthogonality ($\mathbf{R}^{\text{T}}\,\mathbf{R} = \mathbf{I}_3$) and handedness ($\det(\mathbf{R}) = +1$) constraints constitute two different expressions, while in the case of quaternions both are contained within ($\mathbf{q}^*\otimes\mathbf{q} = \mathbf{q}\otimes\mathbf{q}^* = \mathbf{q_1}$). As in other Lie groups, the time derivation of this constraint results in the structure of the Lie algebra. Derivating leads to $\mathbf{q}^*\otimes\dot{\mathbf{q}} = -\left(\mathbf{q}^*\otimes\dot{\mathbf{q}}\right)^*$, which indicates that $\mathbf{q}^*\otimes\dot{\mathbf{q}}$ is in fact a pure quaternion, as is $\dot{\mathbf{q}}\otimes\mathbf{q}^*$. This results in the following particularizations of (128) and (129):

$$\mathbf{\Omega}_{\text{NB}}^{\text{N}\wedge} = \dot{\mathbf{q}}_{\text{NB}}\otimes\mathbf{q}_{\text{NB}}^* = -\mathbf{q}_{\text{NB}}\otimes\dot{\mathbf{q}}_{\text{NB}}^* \tag{252}$$

$$\mathbf{\Omega}_{\text{NB}}^{\text{B}\wedge} = \mathbf{q}_{\text{NB}}^*\otimes\dot{\mathbf{q}}_{\text{NB}} = -\dot{\mathbf{q}}_{\text{NB}}^*\otimes\mathbf{q}_{\text{NB}} \tag{253}$$

The Lie algebra velocity $\mathbf{v}^\wedge$ of $\mathbb{S}^3$ is known as the *half angular velocity* $\mathbf{\Omega}^\wedge$ [26], and as shown in (252) and (253), has the structure of a pure quaternion because its negative coincides with its conjugate:

$$\mathbf{\Omega}^\wedge(\text{t}) = \left[0, \mathbf{\Omega}(\text{t})\right]^{\text{T}} \in \mathbb{H}_{\text{p}} \tag{254}$$

Inverting the previous equations results in the unit quaternion time derivative, which is linear:

$$\dot{\mathbf{q}}_{\text{NB}} = \mathbf{\Omega}_{\text{NB}}^{\text{N}\wedge}\otimes\mathbf{q}_{\text{NB}} = \mathbf{q}_{\text{NB}}\otimes\mathbf{\Omega}_{\text{NB}}^{\text{B}\wedge} \tag{255}$$

Notice that if $\mathbf{q}(\text{t}_0) = \mathbf{q_1}$, then $\dot{\mathbf{q}}(\text{t}_0) = \mathbf{\Omega}(\text{t}_0)$, and hence the pure quaternion $\mathbf{\Omega}^\wedge(\text{t}_0)$ provides a first order approximation of the unit quaternion around the identity $\mathbf{q_1}$:

$$\mathbf{q}(\text{t}_0 + \Delta\text{t}) \approx \mathbf{q}_1 + \mathbf{\Omega}^\wedge(\text{t}_0)\,\Delta\text{t} \tag{256}$$

The *space of pure quaternions* $\mathfrak{so}(3) = \{\mathbf{\Omega}^\wedge \in \mathbb{H}_{\text{p}} \mid \mathbf{\Omega} \in \mathbb{R}^3\}$ is hence the *tangent space* of the unit sphere $\mathbb{S}^3$ of quaternions at the identity $\mathbf{q}_1$, denoted as $\text{T}_{\mathbf{q}_1}\mathcal{R}$. The *hat* $\left\{\cdot^\wedge : \mathbb{R}^3 \to \mathfrak{so}(3) \mid \mathbf{\Omega} \to \mathbf{\Omega}^\wedge\right\}$ and *vee* $\left\{\cdot^\vee : \mathfrak{so}(3) \to \mathbb{R}^3 \mid \left(\mathbf{\Omega}^\wedge\right)^\vee \to \mathbf{\Omega}\right\}$ operators convert the half angular velocity vector into its pure quaternion form, and viceversa.

If $\mathbf{q}(\text{t}_0) \neq \mathbf{q}_1$, the tangent space needs to be transported right multiplying by $\mathbf{q}_{\text{NB}}(\text{t}_0)$ (in the case of space tangent space), or left multiplying for the local space:

$$\mathbf{q}_{\text{NB}}(\text{t}_0 + \Delta\text{t}) \approx \mathbf{q}_{\text{NB}}(\text{t}_0) + \left[\mathbf{\Omega}_{\text{NB}}^{\text{N}\wedge}(\text{t}_0)\,\Delta\text{t}\right]\otimes\mathbf{q}_{\text{NB}}(\text{t}_0) = \left[\mathbf{q}_1 + \mathbf{\Omega}_{\text{NB}}^{\text{N}\wedge}(\text{t}_0)\,\Delta\text{t}\right]\otimes\mathbf{q}_{\text{NB}}(\text{t}_0) \tag{257}$$

$$\mathbf{q}_{\text{NB}}(\text{t}_0 + \Delta\text{t}) \approx \mathbf{q}_{\text{NB}}(\text{t}_0) + \mathbf{q}_{\text{NB}}(\text{t}_0)\otimes\left[\mathbf{\Omega}_{\text{NB}}^{\text{B}\wedge}(\text{t}_0)\,\Delta\text{t}\right] = \mathbf{q}_{\text{NB}}(\text{t}_0)\otimes\left[\mathbf{q}_1 + \mathbf{\Omega}_{\text{NB}}^{\text{B}\wedge}(\text{t}_0)\,\Delta\text{t}\right] \tag{258}$$

Note that the solution to the ordinary differential equation $\dot{\mathbf{x}}(\text{t}) = \mathbf{x}(\text{t})\otimes\mathbf{\Omega}^\wedge$, $\mathbf{x}(\text{t}) \in \mathbb{R}^4$, where $\mathbf{\Omega}^\wedge$ is constant, is $\mathbf{x}(\text{t}) = \mathbf{x}(0)\,\text{e}^{\mathbf{\Omega}^\wedge\text{t}}$. Based on it, assuming $\mathbf{q}(0) = \mathbf{q_1}$ as initial condition, and considering for the time being that $\mathbf{\Omega}$ is constant,

$$\mathbf{q}(\text{t}) = \text{e}^{\mathbf{\Omega}^\wedge\text{t}} = \mathbf{q_1} + \mathbf{\Omega}^\wedge\text{t} + \frac{\left(\mathbf{\Omega}^\wedge\text{t}\right)^2}{2!} + \cdots + \frac{\left(\mathbf{\Omega}^\wedge\text{t}\right)^{\text{n}}}{\text{n}!} + \ldots \tag{259}$$

which is indeed a unit quaternion [26].

Concept	Lie Theory	$\mathbb{SO}^3$
Tangent space element	$\boldsymbol{\tau}^\wedge$	$\mathbf{h}^\wedge = [0, \mathbf{h}]^T$
Velocity element	$\mathbf{v}^\wedge$	$\boldsymbol{\Omega}^\wedge = [0, \boldsymbol{\Omega}]^T$
Structure	$\wedge$	pure quaternion

Table 7: Comparison between generic $\mathbb{SO}(3)$ and half rotation vector as tangent space

Remembering that so far $\boldsymbol{\Omega}^\wedge$ is constant, (259) means that any rotation $\mathbf{q}(\mathrm{t}) = \mathrm{e}^{\boldsymbol{\Omega}^\wedge \mathrm{t}}$ can be realized by maintaining a constant half angular velocity $\boldsymbol{\Omega}^\wedge$ in $\mathbb{H}_\mathrm{p}$ for a given time t. This is analogous to stating that any angular displacement $\mathbf{q}(\theta) = \mathrm{e}^{\mathbf{n}^\wedge \theta}$ can be achieved by rotating an angle θ about a fixed unitary rotation axis $\mathbf{n}^\wedge \in \mathbb{H}_\mathrm{p}$. It is customary to absorb t into $\boldsymbol{\Omega}$ or θ into $\mathbf{n}$, resulting in *the half rotation vector* $\mathbf{h}$:

$$\mathbf{h} = \boldsymbol{\Omega}\,\mathrm{t} = \mathbf{n}\,\theta \quad \in \mathbb{R}^3 \tag{260}$$

Expression (259) represents the *exponential map* $\{\exp : \mathfrak{so}(3) \to \mathbb{SO}(3) \mid \mathbf{h}^\wedge \in \mathbb{H}_\mathrm{p} \to \exp\left(\mathbf{h}^\wedge\right) \in \mathbb{S}^3\}$ [26], which transforms pure quaternions into unit quaternions. Before continuing, let's compare the similarities and differences between the exponential maps when applied to rotation matrices versus quaternions, as both represent maps between the tangent space $\mathfrak{so}(3)$ and the special orthogonal group $\mathbb{SO}(3)$. In the case of rotation matrices, this translates to a map between skew-symmetric matrices and orthogonal ones, while for quaternions the exponential map converts pure quaternions into unitary ones. There is one additional difference, however. In the case of rotation matrices, the map encodes through the rotation vector $\mathbf{r} = \boldsymbol{\omega}\,\mathrm{t} = \mathbf{n}\,\phi$, this is, the axis of rotation n and the rotated angle ϕ. In the case of quaternions, the encoding is through the half rotation vector $\mathbf{h} = \boldsymbol{\Omega}\,\mathrm{t} = \mathbf{n}\,\theta$. Since the rotation is accomplished by the double product $\mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*$ as noted in (250), the vector $\mathbf{v}$ experiences a rotation that is twice that encoded in $\mathbf{q}$, which means that $\mathbf{q}$ encodes half the intended rotation on $\mathbf{v}$. This implies that the space of unit quaternions $\mathbb{S}^3$ is in fact a double covering of $\mathbb{SO}(3)$, not $\mathbb{SO}(3)$ itself [26].

$$\mathbf{r} \;=\; 2 \cdot \mathbf{h} \tag{261}$$

$$\omega \;=\; 2 \cdot \Omega \tag{262}$$

$$\phi \;=\; 2 \cdot \theta \tag{263}$$

Taking these relationships into consideration, it is possible to obtain a more practical expression for the exponential map $\left\{\exp() : \mathfrak{so}(3) \to \mathbb{SO}(3) \mid \mathcal{R} = \exp\left(\mathbf{h}^\wedge\right) = \exp\left(\mathbf{r}^\wedge/2\right)\right\}$ and its capitalized form $\left\{\mathrm{Exp}() : \mathbb{R}^3 \to \mathbb{SO}(3) \mid \mathcal{R} = \mathrm{Exp}(\mathbf{h}) = \mathrm{Exp}(\mathbf{r}/2)\right\}$ than (259) [26]:

$$\mathbf{q}(\mathbf{h}) = \mathbf{q}(\mathbf{r}/2) = \mathrm{Exp}(\mathbf{h}) = \mathrm{Exp}(\mathbf{r}/2) = \mathrm{e}^{\boldsymbol{\Omega}^\wedge \mathrm{t}} = \mathrm{e}^{\mathbf{n}^\wedge \theta} = \mathrm{e}^{\mathbf{n}^\wedge \phi/2} = \cos\frac{\phi}{2} + \mathbf{n}\,\sin\frac{\phi}{2} \tag{264}$$

When regarded as a pure quaternion in $\mathbb{H}_\mathrm{p}$, the angle θ between a unit quaternion $\mathbf{q}$ and the identity $\mathbf{q_1}$ is $\cos\theta = \mathbf{q_1}^T\,\mathbf{q} = \mathrm{q}_0$. At the same time, the angle ϕ rotated by the unit quaternion $\mathbf{q}$ on objects in $\mathbb{R}^3$ satisfies (264), so the angle between a unit quaternion vector and the identify in $\mathbb{H}_\mathrm{p}$ is half the angle rotated by the unit quaternion in $\mathbb{R}^3$ space, as depicted by (263).

So far the exponential map has been obtained based on the assumption of constant angular velocity, but this does not need to be the case. Given a unit quaternion $\mathbf{q} \in \mathbb{S}^3$, there exists a not necessarily unique rotation vector $\mathbf{r} = 2 \cdot \mathbf{h} \in \mathbb{R}^3$ such that $\mathbf{q} = \mathrm{Exp}(\mathbf{r}/2) = \mathrm{Exp}(\mathbf{n}\,\phi/2)$:

$$\phi \;=\; 2\,\arctan\left(\frac{\|\mathbf{q_v}\|}{\mathrm{q}_0}\right) \tag{265}$$

$$\mathbf{n} \;=\; \frac{\mathbf{q_v}}{\|\mathbf{q_v}\|} \tag{266}$$

This expression represents the capitalized *logarithmic map* $\{\log : \mathbb{SO}(3) \to \mathfrak{so}(3) \mid \mathbf{q} \in \mathbb{S}^3 \to \mathbf{h} = \mathbf{r}/2 \in \mathbb{R}^3\}$ [26]. As in the case of the rotation matrix described in section 5.3, the exponential map described by (264) is surjective but not injective, as a rotation of $(\|\mathbf{r}\| + 2\,\mathrm{k}\,\pi)\;\forall\;\mathrm{k} \in \mathbb{Z}$ about $\mathbf{r}/\|\mathbf{r}\|$ produces exactly the same unit quaternion.

In contrast with the case of rotation matrices, a rotation of $(-\|\mathbf{r}\| + 2\,\mathrm{k}\,\pi)$ about $-\mathbf{r}/\|\mathbf{r}\|$ produces the opposite (negative) unit quaternion, although both represent the same rotation.

This shows that the map from rotation matrix to unit quaternion $\{\mathbb{SO}\,(3) \rightarrow \mathbb{SO}\,(3)\,|\,\mathbf{R} \in \mathbb{R}^{3\mathrm{x}3} \rightarrow \mathbf{q} \in \mathbb{S}^3\}$ is also surjective but not injective, as there are two and only two quaternions corresponding to the same rotation matrix $(\mathbf{R}\,(\mathbf{q}) = \mathbf{R}\,(-\mathbf{q}))$. The reason is again the double covering of $\mathbb{SO}\,(3)$ by the unit quaternion [26]. One quaternion induces a rotation in $\mathbb{R}^3$ that follows the shortest direction to the final angle $(\phi < \pi/2)$, while the opposite quaternion rotates the opposite way reaching the same final angle after rotating $(\phi > \pi/2)$.

As the vector $\boldsymbol{\Omega}$ represents half the angular velocity $\boldsymbol{\omega}$, it is possible to adjust expressions (252), (253), and (255):

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}\wedge} = 2\,\dot{\mathbf{q}}_{\mathrm{NB}} \otimes \mathbf{q}_{\mathrm{NB}}^{*} \tag{267}$$

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}\wedge} = 2\,\mathbf{q}_{\mathrm{NB}}^{*} \otimes \dot{\mathbf{q}}_{\mathrm{NB}} \tag{268}$$

$$\dot{\mathbf{q}}_{\mathrm{NB}} = \frac{1}{2}\,\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}\wedge} \otimes \mathbf{q}_{\mathrm{NB}} = \frac{1}{2}\,\mathbf{q}_{\mathrm{NB}} \otimes \boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}\wedge} \tag{269}$$

5.6 Euler Angles

All the previous representations have some type of redundancy as their dimension is higher than three. It is always possible, however, to pick three unitary vectors $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$ forming a basis[29] and perform three consecutive rotations to define a map $\{\mathbb{R}^3 \rightarrow \mathbb{SO}\,(3) \;|\; (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3 \rightarrow \mathrm{R} = e^{\widehat{\mathbf{n}}_1 \beta_1}\, e^{\widehat{\mathbf{n}}_2 \beta_2}\, e^{\widehat{\mathbf{n}}_3 \beta_3} \in \mathbb{SO}\,(3)\}$ [22].

In case the selected basis is orthonormal there are only twelve possible combinations or Euler angles, of which $3-2-1$ is the one employed in this document[30]. The rotation from the spatial or global frame $\mathrm{F}_{\mathrm{N}} = \{\mathbf{O}_{\mathrm{CR}}, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ to the local or body frame $\mathrm{F}_{\mathrm{B}} = \{\mathbf{O}_{\mathrm{CR}}, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is performed by first rotating a *yaw* angle (y) about $\mathbf{n}_3$, followed by rotating a *pitch* angle (p) about $\mathbf{n}_2'$, and finally rotating a *roll* angle (r) about $\mathbf{n}_1''$, where $\mathbf{n}_2'$ is the result of applying the first rotation to $\mathbf{n}_2$ and $\mathbf{n}_1''$, which coincides with $\mathbf{b}_1$, that of applying the first two rotations to $\mathbf{n}_1$.

$$\mathbf{R}_1(\mathrm{r}) = e^{\widehat{\mathbf{n}}_1 \mathrm{r}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mathrm{cr} & -\mathrm{sr} \\ 0 & \mathrm{sr} & \mathrm{cr} \end{bmatrix} \quad \mathbf{R}_2(\mathrm{p}) = e^{\widehat{\mathbf{n}}_2 \mathrm{p}} = \begin{bmatrix} \mathrm{cp} & 0 & \mathrm{sp} \\ 0 & 1 & 0 \\ -\mathrm{sp} & 0 & \mathrm{cp} \end{bmatrix} \quad \mathbf{R}_3(\mathrm{y}) = e^{\widehat{\mathbf{n}}_3 \mathrm{y}} = \begin{bmatrix} \mathrm{cy} & -\mathrm{sy} & 0 \\ \mathrm{sy} & \mathrm{cy} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{270}$$

where s and c stand for sine and cosine respectively. The complete map for the three rotations is:

$$\mathbf{R}_{\mathrm{NB}} = \mathbf{R}_3(\mathrm{y})\;\mathbf{R}_2(\mathrm{p})\;\mathbf{R}_1(\mathrm{r}) = \begin{bmatrix} +\mathrm{cp}\cdot\mathrm{cy} & -\mathrm{cr}\cdot\mathrm{sy} + \mathrm{sr}\cdot\mathrm{sp}\cdot\mathrm{cy} & +\mathrm{sr}\cdot\mathrm{sy} + \mathrm{cr}\cdot\mathrm{sp}\cdot\mathrm{cy} \\ +\mathrm{cp}\cdot\mathrm{sy} & +\mathrm{cr}\cdot\mathrm{cy} + \mathrm{sr}\cdot\mathrm{sp}\cdot\mathrm{sy} & -\mathrm{sr}\cdot\mathrm{cy} + \mathrm{cr}\cdot\mathrm{sp}\cdot\mathrm{sy} \\ -\mathrm{sp} & +\mathrm{sr}\cdot\mathrm{cp} & +\mathrm{cr}\cdot\mathrm{cp} \end{bmatrix} \tag{271}$$

In this document the Euler angles are denoted by $\boldsymbol{\phi}^{\mathrm{NB}} = [\mathrm{y},\,\mathrm{p},\,\mathrm{r}]^{\mathrm{T}}$. They can also be obtained from the rotation matrix, but there are singular instances $(\mathrm{p} = \pm\pi/2)$ where the angles can not be uniquely determined.

$$\mathrm{y} = \arctan\frac{\mathrm{R}_{21}}{\mathrm{R}_{11}} \qquad \mathrm{p} = \arcsin\left(-\mathrm{R}_{31}\right) \qquad \mathrm{r} = \arctan\frac{\mathrm{R}_{32}}{\mathrm{R}_{33}} \tag{272}$$

5.7 Rotational Motion Algebraic Operations

The basic algebraic operations of addition, subtraction, multiplication, division, and exponentiation are not defined for objects of the special orthogonal group $\mathbb{SO}\,(3)$, no matter if they are represented by a rotation matrix $\mathbf{R} \in \mathbb{R}^{3\mathrm{x}3}$, a rotation vector $\mathbf{r} \in \mathbb{R}^3$, or a unit quaternion $\mathbf{q} \in \mathbb{S}^3$. However, as members of the special orthogonal group $\mathbb{SO}\,(3)$, all rotation representations are closed under a given operation that represents the concatenation of rotations, and define not only an identity rotation that represents the lack of rotation, but also an inverse operation representing the opposite rotation. The concatenation of rotations and the identity and inverse operations enable

[29] The base vectors do not need to be orthogonal, just linearly independent.

[30] This means to first rotate about the 3rd axis, then about the resulting 2nd axis, and finally about the ensuing 1st axis.

the definition of the power, exponential and logarithmic operators (section 5.7.1), the spherical linear interpolation (section 5.7.2), and the perturbations together with the plus and minus operators (section 5.7.3).

5.7.1 Powers, Exponentials and Logarithms

Any rotation can be executed by rotating a given angle ϕ about a fixed rotation axis $\mathbf{n}$, resulting in the rotation vector $\mathbf{r} = \mathbf{n}\,\phi$ (section 5.3) or its half $\mathbf{h} = \mathbf{n}\,\phi/2$ (section 5.5). Taking a multiple or a fraction of a rotation vector is hence straightforward, as $\mathrm{t}\,\mathbf{r} = \mathrm{t}\,\mathbf{n}\,\phi = \mathbf{n}\,(\mathrm{t}\,\phi)\ \forall\,\mathrm{t} \in \mathbb{R}, \mathbf{r} \in \mathfrak{so}\,(3)$. The exponential maps defined in (237) and (264) are named that way because they comply with the behavior of the real exponential function $\exp^{\mathrm{b}}(\mathrm{a}) = \exp(\mathrm{a} \cdot \mathrm{b})\ \forall\,\mathrm{a},\,\mathrm{b} \in \mathbb{R}$. As such, the exponential function $\{\exp() : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{SO}\,(3) \mid \mathbf{r} \in \mathbb{R}^3, \mathrm{t} \in \mathbb{R} \to \mathcal{R}^{\mathrm{t}} = \mathrm{Exp}\,(\mathrm{t}\,\mathbf{r}) \in \mathbb{SO}\,(3)\}$ is defined as:

$$\mathbf{R}^{\mathrm{t}}(\mathbf{r}) = \mathbf{R}\,(\mathrm{t}\,\mathbf{r}) = \mathrm{Exp}\,(\mathrm{t}\,\mathbf{r}) = \mathbf{I}_3 + \frac{\widehat{\mathbf{r}}}{\|\mathbf{r}\|}\sin\|\mathrm{t}\,\mathbf{r}\| + \frac{\widehat{\mathbf{r}}^2}{\|\mathbf{r}\|^2}\left(1 - \cos\|\mathrm{t}\,\mathbf{r}\|\right) \tag{273}$$

$$\mathbf{q}^{\mathrm{t}}(\mathbf{h} = \mathbf{r}/2) = \mathbf{q}\,(\mathrm{t}\,\mathbf{h} = \mathrm{t}\,\mathbf{r}/2) = \mathrm{Exp}\,(\mathrm{t}\,\mathbf{r}/2) = \cos\frac{\mathrm{t}\,\phi}{2} + \mathbf{n}\,\sin\frac{\mathrm{t}\,\phi}{2} \tag{274}$$

In a similar way, the logarithmic maps defined in (238), (265), and (266) also comply with the behavior of the real logarithmic function $\mathrm{b} \cdot \log(\mathrm{a}) = \log\left(\mathrm{a}^{\mathrm{b}}\right)\ \forall\,\mathrm{a},\,\mathrm{b} \in \mathbb{R}$. As such, the logarithmic function $\{\log() : \mathbb{SO}\,(3) \times \mathbb{R} \to \mathbb{R}^3 \mid \mathcal{R} \in \mathbb{SO}\,(3), \mathrm{t} \in \mathbb{R} \to \mathrm{t}\,\mathbf{r} = \mathrm{Log}\left(\mathcal{R}^{\mathrm{t}}\right) \in \mathbb{R}^3\}$ is defined as:

$$\log\left(\mathbf{R}^{\mathrm{t}}(\mathbf{r})\right) = \mathrm{Log}\,(\mathrm{Exp}\,(\mathrm{t}\,\mathbf{r})) = \mathrm{t}\,\mathrm{Log}\,(\mathbf{R}\,(\mathbf{r})) = \mathrm{t}\,\mathrm{Log}\,(\mathrm{Exp}\,(\mathbf{r})) = \mathrm{t}\,\mathbf{r} \tag{275}$$

$$\log\left(\mathbf{q}^{\mathrm{t}}(\mathbf{h} = \mathbf{r}/2)\right) = \mathrm{Log}\,(\mathrm{Exp}\,(\mathrm{t}\,\mathbf{r}/2)) = \mathrm{t}\,\mathrm{Log}\,(\mathbf{q}\,(\mathbf{r}/2)) = \mathrm{t}\,\mathrm{Log}\,(\mathrm{Exp}\,(\mathbf{r}/2)) = \mathrm{t}\,\mathbf{r}/2 \tag{276}$$

5.7.2 Spherical Linear Interpolation

Given two rotations $\mathcal{R}_0, \mathcal{R}_1 \in \mathbb{SO}\,(3)$, *spherical linear interpolation* (`SLERP`) seeks to obtain a rotation function $\mathcal{R}\,(\mathrm{t})\,,\ \mathrm{t} \in \mathbb{R}$ that linearly interpolates from $\mathcal{R}\,(0) = \mathcal{R}_0$ to $\mathcal{R}\,(1) = \mathcal{R}_1$ in such a way that the rotation occurs at constant angular velocity along a fixed axis [26].

If employing unit quaternions, $\Delta\mathbf{q}$ is according to (251) the full rotation required to go from $\mathbf{q}_0$ to $\mathbf{q}_1$, such that $\mathbf{q}_1 = \mathbf{q}_0 \otimes \Delta\mathbf{q}$, from where $\Delta\mathbf{q} = \mathbf{q}_0^* \otimes \mathbf{q}_1$. The corresponding rotation vector is then $\Delta\mathbf{r} = \mathbf{n}\,\Delta\phi = 2\,\mathrm{Log}\,(\Delta\mathbf{q})$. Let's take a fraction of the full rotation $\delta\phi = \mathrm{t}\,\Delta\phi$ and obtain the corresponding quaternion:

$$\begin{aligned} \delta\mathbf{q} &= \mathrm{Exp}\left(\frac{\mathbf{n}\,\delta\phi}{2}\right) = \mathrm{Exp}\left(\mathrm{t}\,\frac{\mathbf{n}\,\Delta\phi}{2}\right) = \mathrm{Exp}\left(\mathrm{t}\,\frac{\Delta\mathbf{r}}{2}\right) \\ &= \mathrm{Exp}\big(\mathrm{t}\,\mathrm{Log}\,(\Delta\mathbf{q})\big) = \mathrm{Exp}\big(\mathrm{t}\,\mathrm{Log}\,(\mathbf{q}_0^* \otimes \mathbf{q}_1)\big) = \left(\mathbf{q}_0^* \otimes \mathbf{q}_1\right)^{\mathrm{t}} \end{aligned} \tag{277}$$

The interpolated unit quaternion is hence the following:

$$\mathbf{q}\,(\mathrm{t}) = \mathbf{q}_0 \otimes \left(\mathbf{q}_0^* \otimes \mathbf{q}_1\right)^{\mathrm{t}} \tag{278}$$

Because of the double covering of $\mathbb{SO}\,(3)$ by the quaternion, only the interpolation between quaternions at acute angles ($\Delta\theta = \Delta\phi/2 \leq \pi/2$), is executed following the shortest path, which occurs if $\cos\Delta\theta = {\mathbf{q}_0}^{\mathrm{T}}\,\mathbf{q}_1 < 0$. If this is not the case, just replace $\mathbf{q}_1$ by $-\mathbf{q}_1$ and repeat the process.

A similar result is obtained when employing rotation matrices instead of unit quaternions:

$$\mathbf{R}\,(\mathrm{t}) = \mathbf{R}_0\,\left(\mathbf{R}_0^{\mathrm{T}}\,\mathbf{R}_1\right)^{\mathrm{t}} \tag{279}$$

5.7.3 Plus and Minus Operators

A perturbed rigid body rotation $\widetilde{\mathcal{R}} \in \mathbb{SO}\,(3)$ can always be expressed as the composition of the unperturbed rotation $\mathcal{R}$ with a (usually) small perturbation $\Delta\mathcal{R}$. Perturbations can be specified either at the local or body frame $\mathrm{F_B}$, this is, at the local vector space tangent to $\mathbb{SO}\,(3)$ at the actual orientation, in which case they are known as *local perturbations*. They can also be specified at the global frame $\mathrm{F_N}$, which coincides with the vector space tangent to $\mathbb{SO}(3)$ at the origin; in this case they are known as *global perturbations* [26]. Local perturbations appear on the right hand side of the rotation composition, resulting in $\widetilde{\mathcal{R}} = \mathcal{R} \circ \Delta\mathcal{R}^{\rm B}$, while global ones appear to the left, hence $\widetilde{\mathcal{R}} = \Delta\mathcal{R}^{\rm N} \circ \mathcal{R}$.

The *plus* and *minus operators* are introduced in section 4.3.2 and enable operating with increments of the nonlinear $\mathbb{SO}(3)$ manifold expressed in the linear tangent vector space $\mathfrak{so}(3)$. There exist right ($\oplus$, $\ominus$) or left ($\boxplus$, $\boxminus$) versions depending on whether the increments are viewed in the local frame (right) or the global one (left). It is important to remark that although perturbations and the plus and left operators are best suited to work with small rotation changes (perturbations), the expressions below are generic and work just the same no matter the size of the perturbation.

The right plus operator $\{\oplus : \mathbb{SO}(3) \times \mathfrak{so}(3) \rightarrow \mathbb{SO}(3) \mid \widetilde{\mathcal{R}} = \mathcal{R} \oplus \ \Delta\mathbf{r}^{\rm B} = \mathcal{R} \circ {\rm Exp}(\Delta\mathbf{r}^{\rm B})\}$ produces a rotation element $\widetilde{\mathcal{R}}$ resulting from the composition of a reference rotation $\mathcal{R}$ with an often small rotation $\Delta\mathbf{r}^{\rm B} = \mathbf{n}^{\rm B}\,\Delta\phi$, contained in the tangent space to the reference rotation $\mathcal{R}$, this is, in the local space [26]. The left plus operator $\{\boxplus : \mathfrak{so}(3) \times \mathbb{SO}(3) \rightarrow \mathbb{SO}(3) \mid \widetilde{\mathcal{R}} = \Delta\mathbf{r}^{\rm N} \boxplus \mathcal{R} = {\rm Exp}(\Delta\mathbf{r}^{\rm N}) \circ \mathcal{R}\}$ is similar but the often small rotation $\Delta\mathbf{r}^{\rm N} = \mathbf{n}^{\rm N}\,\Delta\phi$ is contained in the tangent space at the identify or global space. The expressions shown below are valid up to the first coverage of $\mathbb{SO}(3)$, this is, $\phi < \pi$. In the cases of rotation matrix and unit quaternion, the plus operator is defined as:

$$\widetilde{\mathbf{R}} = \mathbf{R} \oplus \Delta\mathbf{r}^{\rm B} = \mathbf{R}\ {\rm Exp}(\Delta\mathbf{r}^{\rm B}) = \mathbf{R}\ \Delta\mathbf{R}^{\rm B} \tag{280}$$

$$\widetilde{\mathbf{q}} = \mathbf{q} \oplus \Delta\mathbf{r}^{\rm B} = \mathbf{q} \otimes {\rm Exp}(\Delta\mathbf{r}^{\rm B}/2) = \mathbf{q} \otimes \Delta\mathbf{q}^{\rm B} \tag{281}$$

$$\widetilde{\mathbf{R}} = \Delta\mathbf{r}^{\rm N} \boxplus \mathbf{R} = {\rm Exp}(\Delta\mathbf{r}^{\rm N})\ \mathbf{R} = \Delta\mathbf{R}^{\rm N}\ \mathbf{R} \tag{282}$$

$$\widetilde{\mathbf{q}} = \Delta\mathbf{r}^{\rm N} \boxplus \mathbf{q} = {\rm Exp}(\Delta\mathbf{r}^{\rm N}/2) \otimes \mathbf{q} = \Delta\mathbf{q}^{\rm N} \otimes \mathbf{q} \tag{283}$$

The right minus operator $\{\ominus : \mathbb{SO}(3) \times \mathbb{SO}(3) \rightarrow \mathfrak{so}(3) \mid \Delta\mathbf{r}^{\rm B} = \widetilde{\mathcal{R}} \ominus \mathcal{R} = {\rm Log}\big(\mathcal{R}^{-1} \circ \widetilde{\mathcal{R}}\big)\}$, as well as the left $\{\boxminus : \mathbb{SO}(3) \times \mathbb{SO}(3) \rightarrow \mathfrak{so}(3) \mid \Delta\mathbf{r}^{\rm N} = \widetilde{\mathcal{R}} \boxminus \mathcal{R} = {\rm Log}\big(\widetilde{\mathcal{R}} \circ \mathcal{R}^{-1}\big)\}$, represent the inverse operations, returning the rotation vector difference $\Delta\mathbf{r}$ between two rotations $\mathcal{R}$ and $\widetilde{\mathcal{R}}$ expressed in either the local or global tangent spaces to $\mathcal{R}$.

$$\Delta\mathbf{r}^{\rm B} = \widetilde{\mathbf{R}} \ominus \mathbf{R} = {\rm Log}\left(\mathbf{R}^{\rm T}\ \widetilde{\mathbf{R}}\right) = {\rm Log}(\Delta\mathbf{R}^{\rm B}) \tag{284}$$

$$\Delta\mathbf{r}^{\rm B} = \widetilde{\mathbf{q}} \ominus \mathbf{q} = 2\ {\rm Log}(\mathbf{q}^* \otimes \widetilde{\mathbf{q}}) = 2\ {\rm Log}(\Delta\mathbf{q}^{\rm B}) \tag{285}$$

$$\Delta\mathbf{r}^{\rm N} = \widetilde{\mathbf{R}} \boxminus \mathbf{R} = {\rm Log}\left(\widetilde{\mathbf{R}}\ \mathbf{R}^{\rm T}\right) = {\rm Log}(\Delta\mathbf{R}^{\rm N}) \tag{286}$$

$$\Delta\mathbf{r}^{\rm N} = \widetilde{\mathbf{q}} \boxminus \mathbf{q} = 2\ {\rm Log}(\widetilde{\mathbf{q}} \otimes \mathbf{q}^*) = 2\ {\rm Log}(\Delta\mathbf{q}^{\rm N}) \tag{287}$$

If the $\Delta\mathbf{r}$ perturbation is small, the (235) and (259) Taylor expansions can be truncated, resulting in the following expressions, valid for both the body frame ($\Delta\mathbf{r}^{\rm B}$) or the global one ($\Delta\mathbf{r}^{\rm N}$):

$$\Delta\mathbf{R} = {\rm Exp}(\Delta\mathbf{r}) \approx \mathbf{I}_3 + \Delta\mathbf{r}^\wedge = \mathbf{I}_3 + \Delta\phi\,\widehat{\mathbf{n}} \tag{288}$$

$$\Delta\mathbf{q} = \exp(\Delta\mathbf{r}/2) \approx \mathbf{q_1} + \Delta\mathbf{r}^\wedge/2 = [1, \mathbf{n}\,\Delta\phi/2]^{\rm T} \tag{289}$$

5.8 Rotational Motion Time Derivative and Angular Velocity

Let's consider a rotating rigid body $\mathcal{R}(\mathrm{t}) \in \mathbb{SO}(3), \mathrm{t} \in \mathbb{R}$ and compute its derivative with time, which belongs to neither $\mathbb{SO}(3)$ nor $\mathfrak{so}(3)$ but to the Euclidean space of the chosen rotation representation, $\mathbb{R}^{3x3}$ for the rotation matrix and $\mathbb{H}$ for the unit quaternion:

$$\dot{\mathcal{R}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathcal{R}(\mathrm{t} + \Delta\mathrm{t}) - \mathcal{R}(\mathrm{t})}{\Delta\mathrm{t}} \tag{290}$$

Considering the time modified rotation $\mathcal{R}(\mathrm{t} + \Delta\mathrm{t})$ as the perturbed state (section 5.7.3), the resulting time derivatives for the rotation matrix and unit quaternion representations are the following:

$$\dot{\mathbf{R}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{R}\,\Delta\mathbf{R}^{\rm B} - \mathbf{R}}{\Delta\mathrm{t}} \approx \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{R}\ \left[\left(\mathbf{I}_3 + \Delta\phi\ \widehat{\mathbf{n}}^{\rm B}\right) - \mathbf{I}_3\right]}{\Delta\mathrm{t}} = \mathbf{R}\ \lim_{\Delta \mathrm{t} \to 0} \frac{\Delta\phi\,\widehat{\mathbf{n}}^{\rm B}}{\Delta\mathrm{t}} \tag{291}$$

$$\dot{\mathbf{q}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{q} \otimes \Delta \mathbf{q}^{\mathrm{B}} - \mathbf{q}}{\Delta \mathrm{t}} \approx \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{q} \otimes \left[[1, \mathbf{n}^{\mathrm{B}} \, \Delta \phi / 2]^{\mathrm{T}} - \mathbf{q_1}\right]}{\Delta \mathrm{t}} = \mathbf{q} \otimes \lim_{\Delta \mathrm{t} \to 0} \frac{[0, \mathbf{n}^{\mathrm{B}} \, \Delta \phi / 2]^{\mathrm{T}}}{\Delta \mathrm{t}} \tag{292}$$

Similar expressions based on $\mathbf{r}^{\mathrm{N}} = \Delta\phi \, \mathbf{n}^{\mathrm{N}}$ can be obtained if left multiplying by the perturbation instead of right multiplying. The $\dot{\mathbf{R}}(\mathrm{t})$ and $\dot{\mathbf{q}}(\mathrm{t})$ expressions (231) and (269) are then directly obtained when defining the *body angular velocity* $\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}}$ as the time derivative of the rotation vector $\mathbf{r}^{\mathrm{B}} = \mathbf{n}^{\mathrm{B}} \, \phi$ when viewed in local or body frame $\mathrm{F_B}$, and the *spatial angular velocity* $\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}}$ as the time derivative of the rotation vector $\mathbf{r}^{\mathrm{N}} = \mathbf{n}^{\mathrm{N}} \, \phi$ when viewed in global or spatial frame $\mathrm{F_N}$:

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}}(\mathrm{t}) = \Delta\dot{\mathbf{r}}^{\mathrm{B}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\Delta \mathbf{r}^{\mathrm{B}}}{\Delta \mathrm{t}} = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{n}^{\mathrm{B}} \, \Delta\phi}{\Delta \mathrm{t}} \tag{293}$$

$$\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}}(\mathrm{t}) = \Delta\dot{\mathbf{r}}^{\mathrm{N}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\Delta \mathbf{r}^{\mathrm{N}}}{\Delta \mathrm{t}} = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{n}^{\mathrm{N}} \, \Delta\phi}{\Delta \mathrm{t}} \tag{294}$$

Note that the rotation of the angular velocity (relationship between $\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}}$ and $\boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}}$) is not given by the rotation action $\mathbf{g}_{\mathcal{R}*}$ (225) but by the adjoint map $\mathbf{Ad}_{\mathcal{R}}$ described in section 5.10, although in the case of the $\mathbb{SO}(3)$ rotation group, both maps coincide.

5.9 Rotational Motion Point Velocity

There exists a direct relationship between the velocity of a point belonging to a rigid body and the elements of its tangent space, this is, the angular velocity in $\mathfrak{so}(3)$ in the case of rotational motion. This relationship is independent of the $\mathbb{SO}(3)$ representation, although the rotation matrix is employed in the expressions below. As discussed in section 5.1, the rotation actions have the same form for points as for vectors (223). Hence, if $\mathbf{p}^{\mathrm{B}}$ are the fixed coordinates of a point belonging to the $\mathrm{F_B}$ rigid body, the point spatial coordinates $\mathbf{p}^{\mathrm{N}}$ can be obtained by means of (226):

$$\mathbf{p}^{\mathrm{N}}(\mathrm{t}) = \mathbf{g}_{\mathcal{R}_{\mathrm{NB}(\mathrm{t})}}(\mathbf{p}^{\mathrm{B}}) = \mathbf{R}_{\mathrm{NB}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{B}} \tag{295}$$

The velocity of a point is the time derivative of its spatial or global coordinates. As $\mathbf{p}$ is fixed to $\mathrm{F_B}$, its time derivative is zero ($\dot{\mathbf{p}}^{\mathrm{B}} = \mathbf{0}$), so its velocity viewed in the spatial frame responds to:

$$\mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t}) = \dot{\mathbf{p}}^{\mathrm{N}}(\mathrm{t}) = \dot{\mathbf{R}}_{\mathrm{NB}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{B}} \tag{296}$$

Although $\dot{\mathbf{R}}_{\mathrm{NB}}$ maps the point body coordinates to its spatial velocity per (296), its high dimension makes it inefficient [27]. By making use of the spatial and body instantaneous angular velocities ($\widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}$, $\widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}}$) introduced in (231), the velocity of a point $\mathbf{p}^{\mathrm{B}}$ viewed in $\mathrm{F_N}$ can be obtained as follows:

$$\mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t}) = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}(\mathrm{t}) \; \mathbf{R}_{\mathrm{NB}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{B}} = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{N}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{N}}(\mathrm{t}) \tag{297}$$

$$\mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t}) = \mathbf{R}_{\mathrm{NB}}(\mathrm{t}) \; \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{B}} \tag{298}$$

The velocity of $\mathbf{p}^{\mathrm{B}}$ viewed in $\mathrm{F_B}$ can then be obtained by means of the vector action map:

$$\mathbf{v}_{\mathrm{p}}^{\mathrm{B}}(\mathrm{t}) = \mathbf{g}_{\mathcal{R}_{\mathrm{NB}(\mathrm{t})}*}^{-1}\big(\mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t})\big) = \mathbf{R}_{\mathrm{NB}}^{\mathrm{T}}(\mathrm{t}) \; \mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t}) = \widehat{\boldsymbol{\omega}}_{\mathrm{NB}}^{\mathrm{B}}(\mathrm{t}) \; \mathbf{p}^{\mathrm{B}} \tag{299}$$

The point velocity is hence the result of the cross product between the angular velocity and the point coordinates (298, 299). Similar expressions are obtained if employing the unit quaternion ($\mathbf{v}_{\mathrm{p}}^{\mathrm{N}}(\mathrm{t}) = \boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{N}\wedge} \otimes \mathbf{p}^{\mathrm{N}}(\mathrm{t})$, $\mathbf{v}_{\mathrm{p}}^{\mathrm{B}}(\mathrm{t}) = \boldsymbol{\omega}_{\mathrm{NB}}^{\mathrm{B}\wedge} \otimes \mathbf{p}^{\mathrm{B}}$).

5.10 Rotational Motion Adjoint

The *adjoint map* of a Lie group is defined in section 4.3.3 as an action of the Lie group on its own Lie algebra that converts between the local tangent space and that at the identity. In the case of rotational motion, both the rotation vector and the angular velocity belong to the tangent space, so

$\left\{\mathbf{Ad}\left(\right) : \mathbb{SO}(3) \times \mathfrak{so}(3) \rightarrow \mathfrak{so}(3) \mid \mathbf{Ad}_{\mathcal{R}}\left(\mathbf{r}^{\wedge}\right) = \mathcal{R} \circ \mathbf{r}^{\wedge} \circ \mathcal{R}^{-1},\ \mathbf{Ad}_{\mathcal{R}}\left(\boldsymbol{\omega}^{\wedge}\right) = \mathcal{R} \circ \boldsymbol{\omega}^{\wedge} \circ \mathcal{R}^{-1}\right\}$. This is equivalent to $\mathbf{q} \otimes \boldsymbol{\omega}^{\wedge} \otimes \mathbf{q}^{*}$ for unit quaternions or $\mathbf{R}\,\widehat{\boldsymbol{\omega}}\,\mathbf{R}^{T}$ for rotation matrices, which represents the congruency transformation[31] between the spatial and body angular velocities $\widehat{\boldsymbol{\omega}}_{NB}^{N}$ and $\widehat{\boldsymbol{\omega}}_{NB}^{B}$:

$$\widehat{\boldsymbol{\omega}}_{NB}^{N} = \mathbf{R}_{NB}\,\widehat{\boldsymbol{\omega}}_{NB}^{B}\,\mathbf{R}_{NB}^{T} \tag{300}$$

The application of the vee operator results in the adjoint matrix coinciding with the rotation matrix itself $\{\mathbf{Ad}_{\mathcal{R}}\,\boldsymbol{\omega} = \mathbf{R}\,\boldsymbol{\omega}\}$[32], implying that elements of the $\mathbb{SO}(3)$ tangent space (both rotation vectors and angular velocities) can be transformed by means of the rotation action as any other free vector. Note that this result only applies to rotational motion, as for example the vector action and adjoint matrix of $\mathbb{SE}(3)$ discussed in section 6 do not coincide:

$$\boldsymbol{\omega}_{NB}^{N} = \mathbf{Ad}_{\mathcal{R}_{NB}}\,\boldsymbol{\omega}_{NB}^{B} = \mathbf{R}_{NB}\,\boldsymbol{\omega}_{NB}^{B} \tag{301}$$

A similar process leads to the inverse adjoint matrix ($\mathbf{Ad}_{\mathcal{R}}^{-1}\,\boldsymbol{\omega} = \mathbf{Ad}_{\mathcal{R}^{-1}}\,\boldsymbol{\omega} = \mathbf{R}^{T}\,\boldsymbol{\omega}$):

$$\boldsymbol{\omega}_{NB}^{B} = \mathbf{Ad}_{\mathcal{R}_{NB}}^{-1}\,\boldsymbol{\omega}_{NB}^{N} = \mathbf{R}_{NB}^{T}\,\boldsymbol{\omega}_{NB}^{N} \tag{302}$$

5.11 Rotational Motion Uncertainty and Covariance

Following the analysis of uncertainty on Lie groups presented in section 4.3.5, the definitions of local and global autocovariances for $\mathbb{SO}(3)$ elements around a nominal or expected rotation $\mathrm{E}\left[\mathcal{R}\right] = \boldsymbol{\mu}_{\mathcal{R}} \in \mathbb{SO}(3)$ are the following:

$$\mathbf{C}_{\mathcal{R}\mathcal{R}}^{B} = \mathrm{E}\left[\Delta\mathbf{r}^{B}\,\Delta\mathbf{r}^{BT}\right] = \mathrm{E}\left[\left(\mathcal{R} \ominus \boldsymbol{\mu}_{\mathcal{R}}\right)\left(\mathcal{R} \ominus \boldsymbol{\mu}_{\mathcal{R}}\right)^{T}\right] \quad \in \mathbb{R}^{3x3} \tag{303}$$

$$\mathbf{C}_{\mathcal{R}\mathcal{R}}^{N} = \mathrm{E}\left[\Delta\mathbf{r}^{N}\,\Delta\mathbf{r}^{NT}\right] = \mathrm{E}\left[\left(\mathcal{R} \boxminus \boldsymbol{\mu}_{\mathcal{R}}\right)\left(\mathcal{R} \boxminus \boldsymbol{\mu}_{\mathcal{R}}\right)^{T}\right] \quad \in \mathbb{R}^{3x3} \tag{304}$$

Note that although the notation refers to the covariance of the rotation manifold $\mathcal{R} \in \mathbb{SO}(3)$, the definition in fact refers to the covariance of the rotation vectors $\Delta\mathbf{r}^{B}$ or $\Delta\mathbf{r}^{N}$ located in the tangent space, with its dimension (3) matching the number of degrees of freedom of the $\mathbb{SO}(3)$ manifold. The relationship between the local and global autocovariances responds to:

$$\mathbf{C}_{\mathcal{R}\mathcal{R}}^{N} = \mathbf{Ad}_{\mathcal{R}_{NB}}\,\mathbf{C}_{\mathcal{R}\mathcal{R}}^{B}\,\mathbf{Ad}_{\mathcal{R}_{NB}}^{T} = \mathbf{R}_{NB}\,\mathbf{C}_{\mathcal{R}\mathcal{R}}^{B}\,\mathbf{R}_{NB}^{T} \tag{305}$$

Given a function $\{\mathrm{f} : \mathcal{R} \rightarrow \mathcal{S} \mid \mathcal{S} = \mathrm{f}\left(\mathcal{R}\right) \in \mathbb{SO}(3),\ \forall\mathcal{R} \in \mathbb{SO}(3)\}$ between two rotations, the covariances are propagated as follows:

$$\mathbf{C}_{\mathcal{S}\mathcal{S}}^{B} = \mathbf{J}_{\oplus\,\mathcal{R}}^{\oplus\,\mathrm{f}\left(\mathcal{R}\right)}\,\mathbf{C}_{\mathcal{R}\mathcal{R}}^{B}\,\mathbf{J}_{\oplus\,\mathcal{R}}^{\oplus\,\mathrm{f}\left(\mathcal{R}\right),T} \quad \in \mathbb{R}^{3x3} \tag{306}$$

$$\mathbf{C}_{\mathcal{S}\mathcal{S}}^{N} = \mathbf{J}_{\boxplus\,\mathcal{R}}^{\boxplus\,\mathrm{f}\left(\mathcal{R}\right)}\,\mathbf{C}_{\mathcal{R}\mathcal{R}}^{N}\,\mathbf{J}_{\boxplus\,\mathcal{R}}^{\boxplus\,\mathrm{f}\left(\mathcal{R}\right),T} \quad \in \mathbb{R}^{3x3} \tag{307}$$

5.12 Rotational Motion Jacobians

Lie group jacobians are introduced in section 4.4 based on the right and left Lie group derivatives of section 4.3.4, and in this section are customized for the $\mathbb{SO}(3)$ case, with table 8 representing the particularization of table 1 to the case of rigid body rotations. The various jacobians listed in table 8 have been obtained by means of the chain rule, the expressions already introduced in this article, and those of section 4.3. Note that although in many cases the results include the rotation matrix, all jacobians are generic and do not depend on the specific $\mathbb{SO}(3)$ parameterization.

In addition to the adjoint matrix, two other jacobians are of particular importance as they appear repeatedly in table 8. These are the right and left jacobians of the capitalized exponential function, also known as simply

[31] Two square matrices $\mathbf{A}$ and $\mathbf{B}$ are called congruent if $\mathbf{B} = \mathbf{P}^{T}\,\mathbf{A}\,\mathbf{P}$ for some invertible square matrix $\mathbf{P}$.

[32] The adjoint matrix is generic and applicable to all $\mathbb{SO}(3)$ representations.

the *right jacobian* $\mathrm{J_R}(\mathbf{r})$ and the *left jacobian* $\mathrm{J_L}(\mathbf{r})$, and they evaluate the variation of the $\mathfrak{so}(3)$ tangent space provided by the output of the $\mathrm{Exp}(\mathbf{r})$ map (locally for $\mathrm{J_R}$ and globally for $\mathrm{J_L}$) while moving along the $\mathbb{SO}(3)$ manifold with respect to the (Euclidean) variations within the original tangent space provided by $\mathbf{r}$. Their closed forms as well as those of their inverses are included in table 8, and have been obtained from [28]; they verify that $\mathbf{J}_{\mathrm{L}}(\mathbf{r}) = \mathbf{J}_{\mathrm{R}}^{\mathrm{T}}(\mathbf{r})$, and $\mathbf{J}_{\mathrm{L}}^{-1}(\mathbf{r}) = \mathbf{J}_{\mathrm{R}}^{-\mathrm{T}}(\mathbf{r})$.

It is also worth noting the special importance of the $\mathbf{J}^{+\;g_{\mathrm{Exp}(\mathbf{r})*}(\mathbf{v})}_{+\;\mathbf{r}}$ jacobian present at the bottom of table 8, which represents the derivative of a rotated vector with respect to perturbations in the Euclidean tangent space (not on the curved manifold) that generates the rotation, as it enables tangent space optimization by calculus methods designed exclusively for Euclidean spaces.

Jacobian	Table 1	Expression	Size
$\mathbf{J}^{\oplus\;\mathcal{R}^{-1}}_{\oplus\;\mathcal{R}}$	$= -\mathbf{Ad}_{\mathcal{R}}$	$= -\mathbf{R}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathcal{R}^{-1}}_{\boxplus\;\mathcal{R}}$	$= -\mathbf{Ad}_{\mathcal{R}}^{-1}$	$= -\mathbf{R}^{\mathrm{T}}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\oplus\;\mathcal{R}\circ\mathcal{S}}_{\oplus\;\mathcal{R}}$	$= \mathbf{Ad}_{\mathcal{S}}^{-1}$	$= \mathbf{R}_{\mathcal{S}}^{\mathrm{T}}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathcal{R}\circ\mathcal{S}}_{\boxplus\;\mathcal{R}}$	$= \mathbf{I}$	$= \mathbf{I}_{3x3}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\oplus\;\mathcal{R}\circ\mathcal{S}}_{\oplus\;\mathcal{S}}$	$= \mathbf{I}$	$= \mathbf{I}_{3x3}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathcal{R}\circ\mathcal{S}}_{\boxplus\;\mathcal{S}}$	$= \mathbf{Ad}_{\mathcal{R}}$	$= \mathbf{R}_{\mathcal{R}}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\oplus\;\mathrm{Exp}(\mathbf{r})}_{+\;\mathbf{r}}$	$= \mathrm{J_R}(\mathbf{r})$	$= \mathbf{I}_3 - \dfrac{1-\cos\lVert\mathbf{r}\rVert}{\lVert\mathbf{r}\rVert^2}\,\widehat{\mathbf{r}} + \dfrac{\lVert\mathbf{r}\rVert - \sin\lVert\mathbf{r}\rVert}{\lVert\mathbf{r}\rVert^3}\,\widehat{\mathbf{r}}^2$	$\in \mathbb{R}^{3x3}$
$\mathrm{J_R^{-1}}(\mathbf{r})$		$= \mathbf{I}_3 + \dfrac{\widehat{\mathbf{r}}}{2} + \left(\dfrac{1}{\lVert\mathbf{r}\rVert^2} - \dfrac{1+\cos\lVert\mathbf{r}\rVert}{2\lVert\mathbf{r}\rVert\sin\lVert\mathbf{r}\rVert}\right)\widehat{\mathbf{r}}^2$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathrm{Exp}(\mathbf{r})}_{+\;\mathbf{r}}$	$= \mathrm{J_L}(\mathbf{r})$	$= \mathbf{I}_3 + \dfrac{1-\cos\lVert\mathbf{r}\rVert}{\lVert\mathbf{r}\rVert^2}\,\widehat{\mathbf{r}} + \dfrac{\lVert\mathbf{r}\rVert - \sin\lVert\mathbf{r}\rVert}{\lVert\mathbf{r}\rVert^3}\,\widehat{\mathbf{r}}^2$	$\in \mathbb{R}^{3x3}$
$\mathrm{J_L^{-1}}(\mathbf{r})$		$= \mathbf{I}_3 - \dfrac{\widehat{\mathbf{r}}}{2} + \left(\dfrac{1}{\lVert\mathbf{r}\rVert^2} - \dfrac{1+\cos\lVert\mathbf{r}\rVert}{2\lVert\mathbf{r}\rVert\sin\lVert\mathbf{r}\rVert}\right)\widehat{\mathbf{r}}^2$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathrm{Log}(\mathcal{R})}_{\oplus\;\mathcal{R}}$	$= \mathrm{J_R^{-1}}\big(\mathrm{Log}(\mathcal{R})\big)$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathrm{Log}(\mathcal{R})}_{\boxplus\;\mathcal{R}}$	$= \mathrm{J_L^{-1}}\big(\mathrm{Log}(\mathcal{R})\big)$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\oplus\;\mathcal{R}\oplus\mathbf{r}}_{\oplus\;\mathcal{R}}$	$= \mathbf{Ad}_{\mathrm{Exp}(\mathbf{r})}^{-1}$	$= \mathbf{R}^{\mathrm{T}}(\mathbf{r})$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathbf{r}\boxplus\mathcal{R}}_{\boxplus\;\mathcal{R}}$	$= \mathbf{Ad}_{\mathrm{Exp}(\mathbf{r})}$	$= \mathbf{R}(\mathbf{r})$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\oplus\;\mathcal{R}\oplus\mathbf{r}}_{+\;\mathbf{r}}$	$= \mathrm{J_R}(\mathbf{r})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{\boxplus\;\mathbf{r}\boxplus\mathcal{R}}_{+\;\mathbf{r}}$	$= \mathrm{J_L}(\mathbf{r})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathcal{S}\ominus\mathcal{R}}_{\oplus\;\mathcal{R}}$	$= -\mathrm{J_L^{-1}}(\mathcal{S}\ominus\mathcal{R})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathcal{S}\boxminus\mathcal{R}}_{\boxplus\;\mathcal{R}}$	$= -\mathrm{J_R^{-1}}(\mathcal{S}\boxminus\mathcal{R})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathcal{S}\ominus\mathcal{R}}_{\oplus\;\mathcal{S}}$	$= \mathrm{J_R^{-1}}(\mathcal{S}\ominus\mathcal{R})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;\mathcal{S}\boxminus\mathcal{R}}_{\boxplus\;\mathcal{S}}$	$= \mathrm{J_L^{-1}}(\mathcal{S}\boxminus\mathcal{R})$		$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;g_{\mathcal{R}*}(\mathbf{v})}_{\oplus\;\mathcal{R}}$		$= -\mathbf{R}\,\widehat{\mathbf{v}}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;g_{\mathcal{R}*}(\mathbf{v})}_{\boxplus\;\mathcal{R}}$		$= -(\mathbf{R}\,\mathbf{v})^{\wedge}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\;g_{\mathcal{R}*}(\mathbf{v})}_{+\;\mathbf{v}}$		$= \mathbf{R}$	$\in \mathbb{R}^{3x3}$

Jacobian	Table 1	Expression	Size
$\mathbf{J}^{+\,g_{\mathcal{R}*}^{-1}(\mathbf{v})}_{\oplus\,\mathcal{R}}$		$= \left(\mathbf{R}^T\,\mathbf{v}\right)^\wedge$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,g_{\mathcal{R}*}^{-1}(\mathbf{v})}_{\boxplus\,\mathcal{R}}$		$= \mathbf{R}^T\,\mathbf{v}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,g_{\mathcal{R}*}^{-1}(\mathbf{v})}_{+\,\mathbf{v}}$		$= \mathbf{R}^T$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}(\boldsymbol{\omega})}_{\oplus\,\mathcal{R}}$		$= -\mathbf{R}\,\widehat{\boldsymbol{\omega}}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}(\boldsymbol{\omega})}_{\boxplus\,\mathcal{R}}$		$= -\left(\mathbf{R}\,\boldsymbol{\omega}\right)^\wedge$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}(\boldsymbol{\omega})}_{+\,\boldsymbol{\omega}}$	$= \mathbf{Ad}_{\mathcal{R}}$	$= \mathbf{R}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}^{-1}(\boldsymbol{\omega})}_{\oplus\,\mathcal{R}}$		$= \left(\mathbf{R}^T\,\boldsymbol{\omega}\right)^\wedge$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}^{-1}(\boldsymbol{\omega})}_{\boxplus\,\mathcal{R}}$		$= \mathbf{R}^T\,\boldsymbol{\omega}$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{R}}^{-1}(\boldsymbol{\omega})}_{+\,\boldsymbol{\omega}}$	$= \mathbf{Ad}_{\mathcal{R}}^{-1}$	$= \mathbf{R}^T$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,g_{\mathrm{Exp}(\mathbf{r})*}(\mathbf{v})}_{+\,\mathbf{r}}$		$= -\mathbf{R}\,(\mathbf{r})\;\widehat{\mathbf{v}}\,J_R\,(\mathbf{r}) = -\big(\mathbf{R}\,(\mathbf{r})\;\mathbf{v}\big)^\wedge\,J_L\,(\mathbf{r})$	$\in \mathbb{R}^{3x3}$
$\mathbf{J}^{+\,g_{\mathrm{Exp}(\mathbf{r})*}^{-1}(\mathbf{v})}_{+\,\mathbf{r}}$		$= -\mathbf{R}^T\,(\mathbf{r})\;\widehat{\mathbf{v}}\,J_L\,(\mathbf{r}) = \big(\mathbf{R}^T\,(\mathbf{r})\;\mathbf{v}\big)^\wedge\,J_R\,(\mathbf{r})$	$\in \mathbb{R}^{3x3}$

Table 8: Rotational motion jacobians

5.13 Rotational Motion Discrete Integration

The discrete integration with time of an element of a Lie group based on its Lie algebra is discussed in detail in section 4.5, which includes expressions for the Euler, Heun and Runge-Kutta methods. In the case of rotational motion, the state vector includes the rotation element $\mathcal{R} \in \mathbb{SO}(3)$ and its angular velocity $\boldsymbol{\omega} \in \mathbb{R}^3$ contained in the tangent space, viewed either in the local ($\boldsymbol{\omega}_{NB}^{B}$) or global ($\boldsymbol{\omega}_{NB}^{N}$) frames. The Euler method expressions equivalent to (168) and (170) are shown below. Expressions for other integration schemes can easily be derived from those in section 4.5:

$$\mathcal{R}_{k+1} \approx \mathcal{R}_k \oplus [\Delta t\;\boldsymbol{\omega}_{NBk}^{B}] = \mathcal{R}_k \circ \mathrm{Exp}\,(\Delta t\;\boldsymbol{\omega}_{NBk}^{B}) \tag{308}$$

$$\mathcal{R}_{k+1} \approx [\Delta t\;\boldsymbol{\omega}_{NBk}^{N}] \boxplus \mathcal{R}_k = \mathrm{Exp}\,(\Delta t\;\boldsymbol{\omega}_{NBk}^{N}) \circ \mathcal{R}_k \tag{309}$$

5.14 Rotational Motion Gauss-Newton Optimization

The minimization by means of the Gauss-Newton iterative method of the Euclidean norm of a non linear function whose input is a Lie group element is presented in section 4.6. In the case of rotational motion, the resulting expressions for perturbations $\Delta\mathbf{r}_{NB}^{N} \in \mathfrak{so}(3)$ to an input rotation $\mathcal{R} \in \mathbb{SO}(3)$ viewed in the global frame F_N are shown in (310) and (311), which are equivalent to the generic (189) and (192). Refer to section 4.6 for the meaning of the function jacobian $\mathbf{J}$ and to section 5.12 for that of the left jacobian $\mathbf{J}_L$.

$$\mathcal{R}_{k+1} \longleftarrow \Delta\mathbf{r}_{NBk}^{N} \boxplus \mathcal{R}_k = \Delta\mathbf{r}_{NBk}^{N} \circ \mathrm{Exp}\,(\mathbf{r}_{NBk}) \tag{310}$$

$$\Delta\mathbf{r}_{NBk}^{N} = -\left[\mathbf{J}_{Lk}^{-T}\,\mathbf{J}_k^T\,\mathbf{J}_k\,\mathbf{J}_{Lk}^{-1}\right]^{-1}\,\mathbf{J}_{Lk}^{-T}\,\mathbf{J}_k^T\,\boldsymbol{\mathcal{E}}_k \tag{311}$$

If the perturbation is viewed in the local frame F_B, (193) and (194) are customized as follows, making use of the right jacobian $\mathbf{J}_R$ defined in section 5.12:

$$\mathcal{R}_{k+1} \longleftarrow \mathcal{R}_k \oplus \Delta\mathbf{r}_{NBk}^{B} = \mathrm{Exp}\,(\mathbf{r}_{NBk}) \circ \Delta\mathbf{r}_{NBk}^{B} \tag{312}$$

$$\Delta\mathbf{r}_{NBk}^{B} = -\left[\mathbf{J}_{Rk}^{-T}\,\mathbf{J}_k^T\,\mathbf{J}_k\,\mathbf{J}_{Rk}^{-1}\right]^{-1}\,\mathbf{J}_{Rk}^{-T}\,\mathbf{J}_k^T\,\boldsymbol{\mathcal{E}}_k \tag{313}$$

5.15 Rotational Motion State Estimation

The adaptation of the EKF state estimation introduced in section 3.3 to the case in which Lie group elements and their velocities are present is discussed in detail in section 4.7. For rotational motion with local perturbations, it is necessary to replace $\mathcal{X} \in \mathcal{G}$ by $\mathcal{R} \in \mathbb{SO}(3)$, $\Delta\boldsymbol{\tau}^{\mathcal{X}} \in T_{\mathcal{X}}\mathcal{G}$ by $\Delta\mathbf{r}^{B} \in \mathfrak{so}(3)$, $\mathbf{v}^{\mathcal{X}} \in \mathbb{R}^{m}$ by $\boldsymbol{\omega}^{B} \in \mathbb{R}^{3}$, $\mathbf{C}^{\mathcal{X}}_{\mathcal{X}\mathcal{X}} \in \mathbb{R}^{mxm}$ by $\mathbf{C}^{B}_{\mathcal{R}\mathcal{R}} \in \mathbb{R}^{3x3}$, and $\mathbf{J}^{\oplus\,\mathcal{X}\,\oplus\,\boldsymbol{\tau}}_{\oplus\,\mathcal{X}}$ by $\mathbf{J}^{\oplus\,\mathcal{R}\,\oplus\,\mathbf{r}}_{\oplus\,\mathcal{R}}$. The particularizations for global perturbations are similar.

5.16 Applications of the Various Rotation Representations

This section discusses five different representations of the rotation or special orthogonal group $\mathbb{SO}(3)$: the rotation matrix, the rotation vector, the unit quaternion, the half rotation vector, and the Euler angles. Although in theory all of them can be employed for each of the purposes described in this article, and the required expressions derived, each representation has its own advantages and disadvantages, being suited for certain purposes but not recommended for others.

- The rotation matrix $\mathbf{R}$ is the most natural representation, possesses an easy to obtain inverse, and linear expressions for composition and rotation. It provides a clear connection with the tangent space, together with the exponential and logarithmic maps, SLERP, and plus and minus operators, which are not complex. Its main drawbacks are the storage costs associated with its high dimension (9) and the expense involved in maintaining orthogonality if allowed to deviate from the manifold [17, 29]. Its high cost precludes its use to track the rotation over its manifold, although most implementations continuously compute it if the adjoint matrix or the jacobian blocks are required.

- The unit quaternion $\mathbf{q}$ is the preferred representation to track the rotation over its manifold, even if it is necessary to obtain the rotation matrix for the adjoint and jacobian blocks. Its advantages with respect to the rotation matrix are its small dimension (4) and ease to maintain unitary if allowed to deviate from the manifold. Unit quaternions are the least natural of the rotation representations, being necessary to convert to a different $\mathbb{SO}(3)$ representation for visualization [17]. While the inverse and concatenation are linear, the rotation action is bilinear, which presents a disadvantage with the rotation matrix. Unit quaternion expressions are slightly more complex than those of the rotation matrix, and present a slightly less obvious connection with the tangent space because in fact they represent a double covering of $\mathbb{SO}(3)$ instead of $\mathbb{SO}(3)$ itself.

- The main advantage of the rotation vector $\mathbf{r}$ is that it belongs to the $\mathfrak{so}(3)$ tangent space while simultaneously being an $\mathbb{SO}(3)$ representation. It is hence indicated for those uses related with incremental rotation changes by means of the exponential map together with the plus and minus operators (periodically adding the perturbations to the unit quaternion tracking the rotation), such as discrete integration, optimization, and state estimation. The rotation vector norm is the most adequate metric for evaluating the rotated distance (or estimation error) between two rigid bodies. Although it benefits from its straightforward inverse, its geometric appeal, and its small dimension (3+1)[33], its usage for other applications is discouraged by its complex non linear kinematics, rotation action, and composition [17], which are not shown in this article.

- The half rotation vector $\mathbf{h}$ is so similar (half) to the rotation vector that its usage is not recommended in order to avoid confusion. Its only real application as the tangent space of the unit quaternion is in practice solved by dividing the rotation vector by two when necessary.

- The Euler angles $\boldsymbol{\phi}$ have a long history and a clear physical meaning, which makes them the best choice for attitude visualization, and constitute the only representation in which its dimension (3) coincides with that of the manifold. However, they are not recommended for any other usage because of the presence of discontinuities, together with complex and non linear expressions for inversion, composition, and rotation action [17].

[33] Strictly speaking the dimension is 3, although any usage requires computing the norm, which is often stored to accelerate the transformations.

6 Motion of Rigid Bodies

This section can be considered as a continuation of the analysis of the rotational motion of rigid bodies contained in section 5, in which its center of rotation $\mathbf{O}_{\text{CR}}$ is not stationary but moves in the Euclidean space $\mathbb{E}^3$. It follows a similar scheme, relying on Lie theory concepts discussed in sections 4.3 and 4.4. Table 9 provides a comparison between the generic nomenclature employed in section 4 and their rigid body motion equivalents. The different representations discussed in this section are summarized in Table 10.

Concept	**Lie Theory**	**Motion**	**Concept**	**Lie Theory**	**Motion**
Lie group	$\mathcal{G}$	$\mathbb{SE}(3)$	Lie group element	$\mathcal{X}$, $\mathcal{Y}$	$\mathcal{M}$, $\mathcal{N}$
Concatenation	$\circ$	$\circ$	Lie algebra	$\mathfrak{m}$	$\mathfrak{se}(3)$
Identity	$\mathcal{E}$	$\mathcal{I}_{\mathcal{M}}$	Inverse	$\mathcal{X}^{-1}$	$\mathcal{M}^{-1}$
Velocity	$\mathbf{v}$	$\boldsymbol{\xi}$	Tangent element	$\boldsymbol{\tau}$	$\boldsymbol{\tau}$
Local frame	$\mathcal{X}$	B	Global frame	$\mathcal{E}$	E
Point action	$\mathrm{g}_{\mathcal{X}}()$	$\mathbf{g}_{\mathcal{M}}(\mathbf{p})$	Vector action	$\mathrm{g}_{\mathcal{X}}()$	$\mathbf{g}_{\mathcal{M}*}(\mathbf{v})$
Adjoint	$\mathbf{Ad}_{\mathcal{X}}\left(\boldsymbol{\tau}^\wedge\right)$	$\mathbf{Ad}_{\mathcal{M}}\left(\boldsymbol{\tau}^\wedge\right)$	Adjoint matrix	$\mathbf{Ad}_{\mathcal{X}}\,\boldsymbol{\tau}$	$\mathbf{Ad}_{\mathcal{M}}\,\boldsymbol{\tau}$

Table 9: Comparison between generic Lie elements and those of rigid body motions

This section begins with an introduction to rigid body motion in section 6.1, followed by a description of the different rigid body motion Lie group representations: the affine representation (section 6.2), the homogeneous matrix (section 6.3), the transform vector (section 6.4), the unit dual quaternion (section 6.5), the half transform vector (section 6.6), and the screw (section 6.7). Algebraic operations on rigid body motions are introduced in section 6.8, such as powers, linear interpolation, and the plus and minus operators. Section 6.9 presents the motion time derivative that leads to the definition of the twist or motion velocity in the tangent space. The velocity of the rigid body points is discussed in section 6.10, followed by the adjoint map in section 6.11, which transforms elements of the tangent space while the motion progresses on its manifold, and by an analysis of uncertainty and covariances applied to rigid body motion (section 6.12). An extensive analysis of the rigid body motion jacobians is presented in section 6.13. Sections 6.14, 6.15, and 6.16 apply the discrete integration of Lie groups, the Gauss-Newton optimization of Lie group functions, and the state estimation of Lie groups contained in sections 4.5, 4.6, and 4.7 to the case of rigid body motions. Finally, the advantages and disadvantages of each motion representation are discussed in section 6.17.

6.1 Special Euclidean (Lie) Group

A rigid body can be represented with a cartesian frame attached to any of its points (the origin), with the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ being simply unit vectors along the main axes. Rigid body motions can be combined and reversed, complying with the algebraic concept of group, but are not endowed with a metric, so they are not part of a metric or Euclidean space (section 4.1). They do however comply with the axioms of a Lie group (section 4.3), and hence the set of rigid body motions together with the operation of motion concatenation comprises $\langle\mathbb{SE}(3),\circ\rangle$, known as the *special Euclidean group* of $\mathbb{R}^3$ [22], where its elements are denoted by $\mathcal{M}$, the identify motion by $\mathcal{I}_{\mathcal{M}}$, and the inverse by $\mathcal{M}^{-1}$. The group has two main actions, which are the motion of points $\left\{\mathbf{g}() : \mathbb{SE}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \mid \mathbf{p} \rightarrow \mathbf{g}_{\mathcal{M}}\left(\mathbf{p}\right)\right\}$ and that of vectors $\left\{\mathbf{g}_*() : \mathbb{SE}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \mid \mathbf{v} \rightarrow \mathbf{g}_{\mathcal{M}*}\left(\mathbf{v}\right)\right\}$.

The movements of rigid bodies are introduced in section 5.1 as orthogonal transformations, this is, those that preserve orthogonality and handedness.

- Norm: $\|\mathbf{g}_{\mathcal{M}*}\left(\mathbf{v}\right)\| = \|\mathbf{v}\|, \forall\, \mathbf{v} \in \mathbb{R}^3$
- Cross product: $\mathbf{g}_{\mathcal{M}*}\left(\mathbf{u}\right) \times \mathbf{g}_{\mathcal{M}*}\left(\mathbf{v}\right) = \mathbf{g}_{\mathcal{M}*}\left(\mathbf{u} \times \mathbf{v}\right), \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^3$

It is also worth noting the relationship between the motions of vectors and points:

$$\mathbf{g}_{\mathcal{M}*}\left(\mathbf{v}\right) = \mathbf{g}_{\mathcal{M}*}\left(\mathbf{q} - \mathbf{p}\right) = \mathbf{g}_{\mathcal{M}}\left(\mathbf{q}\right) - \mathbf{g}_{\mathcal{M}}\left(\mathbf{p}\right) \tag{314}$$

The $\mathbb{SE}(3)$ analysis below adopts the convention introduced in section 4.3, in which all actions, including concatenation $\{\circ : \mathbb{SE}(3) \times \mathbb{SE}(3) \rightarrow \mathbb{SE}(3)\}$, transform elements viewed in the local or body frame $\mathrm{F_B} = \{\mathrm{O_B}, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ into elements viewed in the global or spatial frame $\mathrm{F_E} = \{\mathrm{O_E}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{\mathbf{g}_{\mathcal{M}}(\mathrm{O_B}), \mathbf{g}_{\mathcal{M}*}(\mathbf{b}_1), \mathbf{g}_{\mathcal{M}*}(\mathbf{b}_2), \mathbf{g}_{\mathcal{M}*}(\mathbf{b}_3)\}$[34], which overlap each other before the motion takes place.

Representation	Symbol	Structure	Space
Affine representation	$(\mathcal{R},\ \mathbf{T})$	$\mathbb{SO}(3)$ & free 3-vector	$\mathbb{SE}(3)$
Homogeneous matrix	$\mathbf{M}$	4x4 matrix (322)	$\mathbb{SE}(3)$
Twist	$\boldsymbol{\xi}^\wedge$	4x4 matrix (328)	$\mathfrak{se}(3)$
	$\boldsymbol{\xi} = [\boldsymbol{\nu},\ \boldsymbol{\omega}]^{\mathrm{T}}$	free 6-vector	
Transform vector	$\boldsymbol{\tau}^\wedge$	4x4 matrix (328)	$\mathbb{SE}(3)$ & $\mathfrak{se}(3)$
	$\boldsymbol{\tau} = \boldsymbol{\xi}\,\mathrm{t} = [\mathbf{s},\ \mathbf{r}]^{\mathrm{T}} = [\mathbf{k}\,\rho,\ \mathbf{n}\,\phi]^{\mathrm{T}}$	free 6-vector	
Unit dual quaternion	$\boldsymbol{\zeta}$	unit dual quaternion	$\mathbb{SE}(3)$
Half twist	$\boldsymbol{\Upsilon}^\wedge$	pure dual quaternion	$\mathfrak{se}(3)$
	$\boldsymbol{\Upsilon} = \boldsymbol{\xi}/2$	free 6-vector	
Half transform vector	$\boldsymbol{\Psi}^\wedge$	pure dual quaternion	$\mathbb{SE}(3)$ & $\mathfrak{se}(3)$
	$\boldsymbol{\Psi} = \boldsymbol{\Upsilon}\,\mathrm{t} = \boldsymbol{\xi}\,\mathrm{t}/2 = \boldsymbol{\tau}/2$	free 6-vector	
Screw	$\mathbf{S}^\wedge = (\phi^\diamond/2,\ \mathbf{n}\mathbf{m}^\diamond)$	dual number & dual vector	$\mathbb{SE}(3)$ & $\mathfrak{se}(3)$
	$\mathbf{S} = [\mathbf{n},\ \mathbf{m},\ \mathrm{h},\ \phi]^{\mathrm{T}}$	8-vector	

Table 10: Summary of rigid body motion representations

6.2 Affine Representation

The motion of a rigid body can always be divided into a rotation plus a translation, in which the point motion action responds to:

$$\mathbf{g}_{\mathcal{M}}(\mathbf{p}) = \mathbf{g}_{\mathcal{R}}(\mathbf{p}) + \mathbf{T} \tag{315}$$

where $\mathbf{g}_{\mathcal{R}}$ is the point rotation action discussed in section 5, the point $\mathbf{p}$ is viewed in the local or body frame, and $\mathbf{T}$ represents the vector going from the origin of the global or spatial frame to that of the body frame, viewed in the global frame. Any $\mathbb{SO}(3)$ representation can be employed for the above expression, but the rotation matrix (section 5.2) and the unit quaternion (section 5.4) are the most common:

$$\mathbf{p}^{\mathrm{E}} = \mathbf{R}_{\mathrm{EB}}\,\mathbf{p}^{\mathrm{B}} + \mathbf{T}^{\mathrm{E}}_{\mathrm{EB}} \tag{316}$$

$$\mathbf{p}^{\mathrm{E}} = \mathbf{q}_{\mathrm{EB}} \otimes \mathbf{p}^{\mathrm{B}} \otimes \mathbf{q}^*_{\mathrm{EB}} + \mathbf{T}^{\mathrm{E}}_{\mathrm{EB}} \tag{317}$$

The set of all possible rigid body motions $\mathbb{SE}(3) = \left\{\mathcal{M} = (\mathcal{R},\ \mathbf{T}) \,|\, \mathcal{R} \in \mathbb{SO}(3),\ \mathbf{T} \in \mathbb{R}^3\right\}$, coupled with the motion concatenation defined below, is a valid representation of the special Euclidean group. The inverse motion, as well as the concatenation operation, get slightly more complex because of the affine nature of the point action:

$$(\mathcal{R},\ \mathbf{T})^{-1} = \left(\mathcal{R}^{-1},\ -\mathrm{g}^{-1}_{\mathcal{R}*}(\mathbf{T})\right) \tag{318}$$

$$(\mathcal{R}_{\mathrm{EB}},\ \mathbf{T}^{\mathrm{E}}_{\mathrm{EB}}) = (\mathcal{R}_{\mathrm{EN}},\ \mathbf{T}^{\mathrm{E}}_{\mathrm{EN}}) \circ (\mathcal{R}_{\mathrm{NB}},\ \mathbf{T}^{\mathrm{N}}_{\mathrm{NB}}) = \left(\mathcal{R}_{\mathrm{EN}} \circ \mathcal{R}_{\mathrm{NB}},\ \mathbf{g}_{\mathcal{R}_{\mathrm{EN}}*}(\mathbf{T}^{\mathrm{N}}_{\mathrm{NB}}) + \mathbf{T}^{\mathrm{E}}_{\mathrm{EN}}\right) \tag{319}$$

Concept	$\mathbb{SE}^3$	Affine	Concept	$\mathbb{SE}^3$	Affine
Lie group element	$\mathcal{M}$	$(\mathcal{R},\ \mathbf{T})$	Concatenation	$\circ$	$\circ$
Identity	$\mathcal{I}_{\mathcal{M}}$	$(\mathcal{I}_{\mathcal{R}},\ \mathbf{0})$	Inverse	$\mathcal{M}^{-1}$	$(\mathcal{R},\ \mathbf{T})^{-1}$
Point motion	$\mathbf{g}_{\mathcal{M}}(\mathbf{p})$	$\mathbf{g}_{\mathcal{R}}(\mathbf{p}) + \mathbf{T}$	Vector motion	$\mathbf{g}_{\mathcal{M}*}(\mathbf{v})$	$\mathbf{g}_{\mathcal{R}*}(\mathbf{v})$

Table 11: Comparison between generic $\mathbb{SE}(3)$ and motion affine representation

[34] In contrast with the case of rotational motion described in section 5, the spatial frame is now named E as it usually corresponds to the `ECEF` frame. The `NED` case does not apply to this case as it shares origin with the body frame.

As indicated in (314), the effect of a rigid body motion on a vector requires a different map than that of points because the translational part has the same influence on both the vector initial and final points:

$$\mathbf{g}_{\mathcal{M}*}(\mathbf{v}) = \mathbf{g}_{\mathcal{M}}(\mathbf{q}) - \mathbf{g}_{\mathcal{M}}(\mathbf{p}) = (\mathbf{g}_{\mathcal{R}}(\mathbf{q}) + \mathbf{T}) - (\mathbf{g}_{\mathcal{R}}(\mathbf{p}) + \mathbf{T}) = \mathbf{g}_{\mathcal{R}}(\mathbf{q} - \mathbf{p}) = \mathbf{g}_{\mathcal{R}*}(\mathbf{v}) \neq \mathbf{g}_{\mathcal{M}}(\mathbf{v}) \tag{320}$$

6.3 Homogeneous Matrix

Homogeneous coordinates are introduced with the objective of replacing the affine transformation (316) representing the rigid body motion with a linear transformation. Given a point $\mathbf{p} = [\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3]^{\mathrm{T}} \in \mathbb{R}^3$, its homogeneous representation is obtained adding a "1" as a fourth coordinate, so that $\bar{\mathbf{p}} = [\mathbf{p}, 1]^{\mathrm{T}} = [\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, 1]^{\mathrm{T}} \in \mathbb{R}^4$. In the case of a vector $\mathbf{v} = \mathbf{q} - \mathbf{p} \in \mathbb{R}^3$, its homogeneous coordinates are $\bar{\mathbf{v}} = [\mathbf{v}, 0]^{\mathrm{T}} = \bar{\mathbf{q}} - \bar{\mathbf{p}} = [\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3, 0]^{\mathrm{T}} \in \mathbb{R}^4$. The affine coordinate transformation (316) can then be converted into a linear transformation:

$$\mathbf{g}_{\mathcal{M}}(\bar{\mathbf{p}}) = \begin{bmatrix} \mathbf{g}_{\mathcal{M}}(\mathbf{p}) \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} = \mathbf{M}\,\bar{\mathbf{p}} \tag{321}$$

where $\mathbf{M} \in \mathbb{R}^{4\mathrm{x}4}$ is the homogeneous representation of $(\mathbf{R},\, \mathbf{T}) \in \mathbb{SE}(3)$. This enables a natural matrix representation of the special Euclidean group [22]:

$$\mathbb{SE}(3) = \left\{ \mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \;\middle|\; \mathbf{R} \in \mathbb{R}^{3\mathrm{x}3},\ \mathbf{T} \in \mathbb{R}^3 \right\} \subset \mathbb{R}^{4\mathrm{x}4} \tag{322}$$

which has group structure under matrix multiplication $\{\mathbb{R}^{4\mathrm{x}4} \times \mathbb{R}^{4\mathrm{x}4} \rightarrow \mathbb{R}^{4\mathrm{x}4} \mid \mathbf{M}_{\mathrm{a}}\,\mathbf{M}_{\mathrm{b}} \in \mathbb{R}^{4\mathrm{x}4}, \forall\, \mathbf{M}_{\mathrm{a}},\ \mathbf{M}_{\mathrm{b}} \in \mathbb{R}^{4\mathrm{x}4}\}$ [15]. While having dimension sixteen, the special euclidean group $\mathbb{SE}(3)$ defined by means of homogeneous matrices constitutes a six dimensional manifold to euclidean space $\mathbb{E}^6$. Note that in this group the identity element is given by the identity matrix ($\mathbf{I} = \mathbf{I}_4$).

Concept	$\mathbb{SE}^3$	**Homogeneous**	**Concept**	$\mathbb{SE}^3$	**Homogeneous**
Lie group element	$\mathcal{M}$	$\mathbf{M}$	Concatenation	$\circ$	Matrix product
Identity	$\mathcal{I}_{\mathcal{M}}$	$\mathbf{I}_4$	Inverse	$\mathcal{M}^{-1}$	$\mathbf{M}^{-1}$
Point motion	$\mathbf{g}_{\mathcal{M}}(\mathbf{p})$	$\mathbf{M}\,\bar{\mathbf{p}}$	Vector motion	$\mathbf{g}_{\mathcal{M}*}(\mathbf{v})$	$\mathbf{M}\,\bar{\mathbf{v}}$

Table 12: Comparison between generic $\mathbb{SE}(3)$ and homogeneous matrix

The inversion and concatenation of transformations are linear when using homogeneous coordinates [22]:

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{R}^{\mathrm{T}} & -\mathbf{R}^{\mathrm{T}}\,\mathbf{T} \\ 0 & 1 \end{bmatrix} \tag{323}$$

$$\mathbf{M}_{\mathrm{EB}} = \begin{bmatrix} \mathbf{R}_{\mathrm{EB}} & \mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 1 \end{bmatrix} = \mathbf{M}_{\mathrm{EN}}\,\mathbf{M}_{\mathrm{NB}} = \begin{bmatrix} \mathbf{R}_{\mathrm{EN}} & \mathbf{T}_{\mathrm{EN}}^{\mathrm{E}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\mathrm{NB}} & \mathbf{T}_{\mathrm{NB}}^{\mathrm{N}} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathrm{EN}}\mathbf{R}_{\mathrm{NB}} & \mathbf{R}_{\mathrm{EN}}\mathbf{T}_{\mathrm{NB}}^{\mathrm{N}} + \mathbf{T}_{\mathrm{EN}}^{\mathrm{E}} \\ 0 & 1 \end{bmatrix} \tag{324}$$

An advantage of the homogeneous representation is that the motion actions on points and vectors share the same expression:

$$\mathbf{g}_{\mathcal{M}*}(\bar{\mathbf{v}}) = \mathbf{g}_{\mathcal{M}}(\bar{\mathbf{q}}) - \mathbf{g}_{\mathcal{M}}(\bar{\mathbf{p}}) = \mathbf{M}\,\bar{\mathbf{q}} - \mathbf{M}\,\bar{\mathbf{p}} = \mathbf{M}\ (\bar{\mathbf{q}} - \bar{\mathbf{p}}) = \mathbf{M}\,\bar{\mathbf{v}} = \mathbf{g}_{\mathcal{M}}(\bar{\mathbf{v}}) \tag{325}$$

6.4 Transform Vector as Tangent Space

As discussed in section 4.3.1, the structure of the Lie algebra associated to $\mathbb{SE}(3)$ can be obtained by time derivating the Lie group inverse constraint, $\mathbf{M}^{-1}(\mathrm{t})\,\mathbf{M}(\mathrm{t}) = \mathbf{M}(\mathrm{t})\,\mathbf{M}^{-1}(\mathrm{t}) = \mathbf{I}_4$, resulting in the following particularizations of (128) and (129):

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge} = \dot{\mathbf{M}}_{\mathrm{EB}}\,\mathbf{M}_{\mathrm{EB}}^{-1} = -\mathbf{M}_{\mathrm{EB}}\,\dot{\mathbf{M}}_{\mathrm{EB}}^{-1} \tag{326}$$

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge} = \mathbf{M}_{\mathrm{EB}}^{-1}\,\dot{\mathbf{M}}_{\mathrm{EB}} = -\dot{\mathbf{M}}_{\mathrm{EB}}^{-1}\,\mathbf{M}_{\mathrm{EB}} \tag{327}$$

The Lie algebra velocity $\mathbf{v}^\wedge$ of $\mathbb{SE}(3)$ is known as the *twist* $\boldsymbol{\xi}^\wedge$, and has the following structure, derived from (326) and (327):

$$\boldsymbol{\xi}^\wedge(\mathrm{t}) = \begin{bmatrix} \widehat{\boldsymbol{\omega}}(\mathrm{t}) & \boldsymbol{\nu}(\mathrm{t}) \\ 0 & 0 \end{bmatrix} \subset \mathbb{R}^{4x4} \tag{328}$$

The twist $\boldsymbol{\xi}$ represents the motion velocity and is composed by the angular velocity $\boldsymbol{\omega}$ defined in section 5.8 and the *linear velocity* $\boldsymbol{\nu}$, defined in section 6.9. Inverting (326) and (327) results in the homogeneous matrix time derivative, which is linear:

$$\dot{\mathbf{M}}_{\mathrm{EB}} = \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}\,\mathbf{M}_{\mathrm{EB}} = \mathbf{M}_{\mathrm{EB}}\,\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge} \tag{329}$$

Notice that if $\mathbf{M}(\mathrm{t}_0) = \mathbf{I}_4$, then $\dot{\mathbf{M}}(\mathrm{t}_0) = \boldsymbol{\xi}^\wedge(\mathrm{t}_0)$, and hence the twist matrix $\boldsymbol{\xi}^\wedge(\mathrm{t}_0)$ provides a first order approximation of the homogeneous matrix around the identity matrix $\mathbf{I}_4$:

$$\mathbf{M}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \mathbf{I}_4 + \boldsymbol{\xi}^\wedge(\mathrm{t}_0)\,\Delta\mathrm{t} \tag{330}$$

The space of matrices with the (328) structure, $\mathfrak{se}(3) = \{\boldsymbol{\xi}^\wedge = \begin{bmatrix}\widehat{\boldsymbol{\omega}} & \boldsymbol{\nu}; & \mathbf{0} & 0\end{bmatrix} \subset \mathbb{R}^{4x4} \mid \widehat{\boldsymbol{\omega}} \in \mathfrak{so}(3),\ \boldsymbol{\nu} \in \mathbb{R}^3\}$ is hence the *tangent space* of $\mathbb{SE}(3)$ at the identity $\mathbf{I}_4$ [22], denoted as $\mathrm{T}_{\mathbf{I}_4}\mathcal{M}$. With the twist cartesian coordinates defined as $\boldsymbol{\xi} = [\boldsymbol{\nu},\, \boldsymbol{\omega}]^{\mathrm{T}} \in \mathbb{R}^6$, the *hat* $\left\{\cdot^\wedge : \mathbb{R}^6 \to \mathfrak{se}(3) \mid \boldsymbol{\xi} \to \boldsymbol{\xi}^\wedge\right\}$ and *vee* $\left\{\cdot^\vee : \mathfrak{se}(3) \to \mathbb{R}^6 \mid \left(\boldsymbol{\xi}^\wedge\right)^\vee \to \boldsymbol{\xi}\right\}$ operators convert the cartesian vector form of the twist into its matrix form, and viceversa.

If $\mathbf{M}(\mathrm{t}_0) \neq \mathbf{I}_4$, the tangent space needs to be transported right multiplying by $\mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0)$ (in the case of space twist), or left multiplying for the local based twist:

$$\mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0) + [\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}(\mathrm{t}_0)\,\Delta\mathrm{t}]\,\mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0) = [\mathbf{I}_4 + \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}(\mathrm{t}_0)\,\Delta\mathrm{t}]\,\mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0) \tag{331}$$

$$\mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0) + \mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0)\,[\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}(\mathrm{t}_0)\,\Delta\mathrm{t}] = \mathbf{M}_{\mathrm{EB}}(\mathrm{t}_0)\,[\mathbf{I}_4 + \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}(\mathrm{t}_0)\,\Delta\mathrm{t}] \tag{332}$$

Note that the solution to the ordinary differential equation $\dot{\mathbf{x}}(\mathrm{t}) = \boldsymbol{\xi}^\wedge\,\mathbf{x}(\mathrm{t})\,,\ \mathbf{x}(\mathrm{t}) \in \mathbb{R}^6$, where $\boldsymbol{\xi}^\wedge$ is constant, is $\mathbf{x}(\mathrm{t}) = \mathrm{e}^{\boldsymbol{\xi}^\wedge\mathrm{t}}\,\mathbf{x}(0)$. Based on it, assuming $\mathbf{M}(0) = \mathbf{I}_4$ as initial condition, and considering for the time being that $\boldsymbol{\xi}^\wedge$ is constant,

$$\mathbf{M}(\mathrm{t}) = \mathrm{e}^{\boldsymbol{\xi}^\wedge\mathrm{t}} = \mathbf{I}_4 + \boldsymbol{\xi}^\wedge\mathrm{t} + \frac{\left(\boldsymbol{\xi}^\wedge\mathrm{t}\right)^2}{2!} + \dots + \frac{\left(\boldsymbol{\xi}^\wedge\mathrm{t}\right)^{\mathrm{n}}}{\mathrm{n}!} + \dots \tag{333}$$

It can be proved by means of (237) and the matrix exponential properties that (333) is indeed an homogeneous matrix that hence represents the special Euclidean $\mathbb{SE}(3)$ transformations.

Concept	**Lie Theory**	$\mathbb{SE}^3$
Tangent space element	$\boldsymbol{\tau}^\wedge$	$\boldsymbol{\tau}^\wedge = \begin{bmatrix}\widehat{\mathbf{r}}\ \mathbf{s};\ \mathbf{0}\ 0\end{bmatrix}^{\mathrm{T}}$
Velocity element	$\mathbf{v}^\wedge$	$\boldsymbol{\xi}^\wedge = \begin{bmatrix}\widehat{\boldsymbol{\omega}}\ \boldsymbol{\nu};\ \mathbf{0}\ 0\end{bmatrix}^{\mathrm{T}}$
Structure	$\wedge$	(328)

Table 13: Comparison between generic $\mathbb{SE}(3)$ and transform vector as tangent space

Remembering that so far $\boldsymbol{\xi}^\wedge$ is constant, which is equivalent to both $\widehat{\boldsymbol{\omega}}$ and $\boldsymbol{\nu}$ within (328) also being constant, (333) means that any rigid body motion $\mathbf{M}(\mathrm{t}) = \mathrm{e}^{\boldsymbol{\xi}^\wedge\mathrm{t}}$ can be realized by maintaining a constant twist $\boldsymbol{\xi}^\wedge$ for a given time t [22]. The vectors $\mathbf{n} = \boldsymbol{\omega}\,\mathrm{t}/\|\boldsymbol{\omega}\,\mathrm{t}\| = \mathbf{r}/\|\mathbf{r}\|$ and $\mathbf{k} = \boldsymbol{\nu}\,\mathrm{t}/\|\boldsymbol{\nu}\,\mathrm{t}\| = \mathbf{s}/\|\mathbf{s}\|$ indicate the twist directions, while $\phi = \|\mathbf{r}\|$ and $\rho = \|\mathbf{s}\|$ represent the twist magnitudes, respectively. This enables the definition of the *transform vector* $\boldsymbol{\tau}$, also known as the *exponential coordinates* of the $\mathcal{M}$ motion, as

$$\boldsymbol{\tau} = \boldsymbol{\xi}\,\mathrm{t} = [\boldsymbol{\nu}\,\mathrm{t},\, \boldsymbol{\omega}\,\mathrm{t}]^{\mathrm{T}} = [\mathbf{s},\, \mathbf{r}]^{\mathrm{T}} = [\mathbf{k}\,\rho,\, \mathbf{n}\,\phi]^{\mathrm{T}} \in \mathbb{R}^6 \tag{334}$$

Note that the transform vector $\boldsymbol{\tau}$ belongs to the tangent space as it is a multiple of the twist $\boldsymbol{\xi} \in \mathfrak{se}(3)$, and hence tends to coincide with it as time tends to zero. The *exponential map* $\{\exp() : \mathfrak{se}(3) \to \mathbb{SE}(3) \mid \mathcal{M} = \exp(\boldsymbol{\tau}^\wedge)\}$

and its capitalized form $\{\text{Exp}() : \mathbb{R}^6 \rightarrow \mathbb{SE}(3) \mid \mathcal{M} = \text{Exp}(\boldsymbol{\tau})\}$ wrap the transform vector around the special Euclidean group. However, the twist $\boldsymbol{\xi}^{\wedge}(\text{t})$ in fact is not required to be constant. Given a rigid body motion represented by its homogeneous matrix $\mathbf{M} \in \mathbb{SE}(3)$, it can be proved that there exists a not necessarily unique transform vector $\boldsymbol{\tau} = [\mathbf{s},\, \mathbf{r}]^{\text{T}} = [\mathbf{k}\,\rho,\, \mathbf{n}\,\phi]^{\text{T}}$ such that $\mathbf{M} = \text{e}^{\boldsymbol{\tau}^{\wedge}}$ [22, 27]. The exponential map has the following form [22]:

$$\mathbf{M}(\boldsymbol{\tau}) = \exp(\boldsymbol{\tau}^{\wedge}) = \begin{bmatrix} \exp(\widehat{\mathbf{r}}) & \dfrac{[\mathbf{I}_3 - \exp(\widehat{\mathbf{r}})]\,\widehat{\mathbf{r}}\,\mathbf{s} + \mathbf{r}\,\mathbf{r}^{\text{T}}\,\mathbf{s}}{\|\mathbf{r}\|^2} \\ 0 & 1 \end{bmatrix} \qquad \mathbf{r} \neq \mathbf{0} \tag{335}$$

$$\mathbf{M}(\boldsymbol{\tau}) = \exp(\boldsymbol{\tau}^{\wedge}) = \begin{bmatrix} \mathbf{I}_3 & \mathbf{s} \\ 0 & 1 \end{bmatrix} \qquad \mathbf{r} = \mathbf{0} \tag{336}$$

The exponential map described above is thus surjective but not injective, as in general there are infinitely many solutions to the map. The *logarithmic map* $\{\log() : \mathbb{SE}(3) \rightarrow \mathfrak{se}(3) \mid \boldsymbol{\tau}^{\wedge} = \log(\mathcal{M})\}$ and its capitalized version $\{\text{Log}() : \mathbb{SE}(3) \rightarrow \mathbb{R}^6 \mid \boldsymbol{\tau} = \text{Log}(\mathcal{M})\}$ hence convert rigid body motions into transform vectors [22]. The rotation vector $\mathbf{r}$ is provided by (238); if $\mathbf{r} = \mathbf{0}$, $\mathbf{s}$ coincides with $\mathbf{T}$, while otherwise it is obtained by solving for $\mathbf{s}$ from the following linear system [22], taken from (335):

$$\left[(\mathbf{I}_3 - \text{e}^{\widehat{\mathbf{r}}})\,\widehat{\mathbf{r}} + \mathbf{r}\,\mathbf{r}^{\text{T}}\right]\,\mathbf{s} = \|\mathbf{r}\|^2\,\mathbf{T} \tag{337}$$

Unlike the case of the rotation vector (239), the transform vector inverse coincides with its negative only if the motion is very small. The different $\mathbb{SE}(3)$ actions (concatenation, point motion, vector motion), the inverse, and the relationship between the transform vector derivative with time and the twist, are complex and rarely used.

6.5 Unit Dual Quaternion

The dual quaternions with unity norm, known as unit dual quaternions, comprise an additional representation of the special euclidean group $\mathbb{SE}(3)$, as shown below. Dual quaternions in turn are generalizations of quaternions in the same way as dual numbers are generalization or real ones [24, 30, 31]. For this reason, it is necessary to first describe the dual numbers and dual vectors in sections 6.5.1 and 6.5.2 before discussing the dual quaternions in section 6.5.3 and finally the unit dual quaternions in section 6.5.4.

6.5.1 Dual Numbers

The set of *dual numbers* $\mathbb{D}$ is defined as $\{\mathbb{D} = \mathbb{R} + \mathbb{R}\,\epsilon \mid \epsilon^2 = \epsilon \cdot \epsilon = 0\}$. Given two dual numbers $\text{d}_1^{\diamond} = \text{x}_1 + \text{y}_1\,\epsilon \in \mathbb{D}, \text{d}_2^{\diamond} = \text{x}_2 + \text{y}_2\,\epsilon \in \mathbb{D}, \forall\, \text{x}_1, \text{y}_1, \text{x}_2, \text{y}_2 \in \mathbb{R}$, it is possible to define the operations of addition $\{+ : \mathbb{D} \times \mathbb{D} \rightarrow \mathbb{D}\}$ and multiplication $\{\cdot : \mathbb{D} \times \mathbb{D} \rightarrow \mathbb{D}\}$ [24]:

$$\text{d}_1^{\diamond} + \text{d}_2^{\diamond} = (\text{x}_1 + \text{y}_1\,\epsilon) + (\text{x}_2 + \text{y}_2\,\epsilon) = (\text{x}_1 + \text{x}_2) + (\text{y}_1 + \text{y}_2)\,\epsilon \tag{338}$$

$$\text{d}_1^{\diamond} \cdot \text{d}_2^{\diamond} = (\text{x}_1 + \text{y}_1\,\epsilon) \cdot (\text{x}_2 + \text{y}_2\,\epsilon) = (\text{x}_1\text{x}_2) + (\text{x}_1\text{y}_2 + \text{y}_1\text{x}_2)\,\epsilon \tag{339}$$

The set of dual numbers $\mathbb{D}$ endowed with the operations of addition $+$ and multiplication $\cdot$ forms a ring, known as the ring of dual numbers $\langle\mathbb{D}, +, \cdot\rangle$, nearly always abbreviated to simply $\mathbb{D}$. The additive identity is $0^{\diamond} = 0 + 0\,\epsilon$ and the inverse $-\text{d}^{\diamond} = -\text{x} - \text{y}\,\epsilon$, while the multiplication identity is $1^{\diamond} = 1 + 0\,\epsilon$ and the inverse $\text{d}^{\diamond -1} = 1/\text{x} - \text{y}\,\epsilon/\text{x}^2$. Note that $\mathbb{D}$ is a ring instead of a field as the multiplicative inverse $\text{d}^{\diamond -1}$ is not defined when $\text{x} = 0$. The conjugate of a dual number is obtained by switching the sign of its dual part ($\text{d}^* = \text{x} - \text{y}\,\epsilon \in \mathbb{D}$).

The most useful property of dual numbers is the explicit relationship that exists between the value of any function evaluated at a dual number $\text{f}(\text{d}^{\diamond}) = \text{f}(\text{x} + \text{y}\,\epsilon)$ and its value when evaluated exclusively at its real part $\text{f}(\text{x})$ [30]. The Taylor expansion of $\text{f}(\text{x} + \text{y}\,\epsilon)$ around x reads:

$$\text{f}(\text{d}^{\diamond}) = \text{f}(\text{x} + \text{y}\,\epsilon) = \text{f}(\text{x}) + \frac{\partial \text{f}}{\partial \text{d}^{\diamond}}(\text{x})\,(\text{d}^{\diamond} - \text{x}) + \frac{1}{2!}\frac{\partial^2 \text{f}}{\partial \text{d}^{\diamond 2}}(\text{x})\,(\text{d}^{\diamond} - \text{x})^2 + \frac{1}{3!}\frac{\partial^3 \text{f}}{\partial \text{d}^{\diamond 3}}(\text{x})\,(\text{d}^{\diamond} - \text{x})^3 + \cdots \tag{340}$$

As $(\mathrm{d}^\diamond - \mathrm{x})^{\mathrm{n}} = \mathrm{y}^{\mathrm{n}}\,\epsilon^{\mathrm{n}}$ is zero when $\mathrm{n} > 1$, this translates into:

$$\mathrm{f}\,(\mathrm{d}^\diamond) = \mathrm{f}\,(\mathrm{x} + \mathrm{y}\,\epsilon) = \mathrm{f}\,(\mathrm{x}) + \frac{\partial \mathrm{f}}{\partial \mathrm{d}^\diamond}\,(\mathrm{x})\ \mathrm{y}\,\epsilon \tag{341}$$

6.5.2 Dual Vectors

Dual vectors in three dimensions are formed by grouping three dual numbers $\{\mathbf{d}^\diamond = [\mathrm{d}_1^\diamond,\, \mathrm{d}_2^\diamond,\, \mathrm{d}_3^\diamond]^{\mathrm{T}} \in \mathbb{D}^3,\ \forall\ \mathrm{d}_1^\diamond,\, \mathrm{d}_2^\diamond,\, \mathrm{d}_3^\diamond \in \mathbb{D}\}$. It is then possible to define, $\forall\ \mathrm{d}^\diamond \in \mathbb{D}, \mathbf{d}^\diamond, \mathbf{e}^\diamond \in \mathbb{D}^3$, the scalar multiplication of a double number by a double vector $\{\cdot : \mathbb{D} \times \mathbb{D}^3 \rightarrow \mathbb{D}^3\}$, the inner product between two double vectors $\{\langle\cdot\,,\cdot\rangle : \mathbb{D}^3 \times \mathbb{D}^3 \rightarrow \mathbb{D}\}$, and the cross product between two double vectors $\{\times : \mathbb{D}^3 \times \mathbb{D}^3 \rightarrow \mathbb{D}^3\}$. The results are similar to those of real numbers shown in section 4.1 [24]:

$$\mathrm{d}^\diamond \cdot \mathbf{d}^\diamond = [\mathrm{d}^\diamond\,\mathrm{d}_1^\diamond,\, \mathrm{d}^\diamond\,\mathrm{d}_2^\diamond,\, \mathrm{d}^\diamond\,\mathrm{d}_3^\diamond]^{\mathrm{T}} \tag{342}$$

$$\langle \mathbf{d}^\diamond, \mathbf{e}^\diamond \rangle = \mathbf{d}^\diamond \cdot \mathbf{e}^\diamond = \mathbf{d}^{\diamond\,\mathrm{T}}\,\mathbf{e}^\diamond = [\mathrm{d}_1^\diamond\,\mathrm{e}_1^\diamond,\, \mathrm{d}_2^\diamond\,\mathrm{e}_2^\diamond,\, \mathrm{d}_3^\diamond\,\mathrm{e}_3^\diamond]^{\mathrm{T}} \tag{343}$$

$$\mathbf{d}^\diamond \times \mathbf{e}^\diamond = \widehat{\mathbf{d}^\diamond}\,\mathbf{e}^\diamond = \begin{bmatrix} 0^\diamond & -\mathrm{d}_3^\diamond & +\mathrm{d}_2^\diamond \\ +\mathrm{d}_3^\diamond & 0^\diamond & -\mathrm{d}_1^\diamond \\ -\mathrm{d}_2^\diamond & +\mathrm{d}_1^\diamond & 0^\diamond \end{bmatrix} \begin{bmatrix} \mathrm{e}_1^\diamond \\ \mathrm{e}_2^\diamond \\ \mathrm{e}_3^\diamond \end{bmatrix} = \begin{bmatrix} \mathrm{d}_2^\diamond\,\mathrm{e}_3^\diamond - \mathrm{d}_3^\diamond\,\mathrm{e}_2^\diamond \\ \mathrm{d}_3^\diamond\,\mathrm{e}_1^\diamond - \mathrm{d}_1^\diamond\,\mathrm{e}_3^\diamond \\ \mathrm{d}_1^\diamond\,\mathrm{e}_2^\diamond - \mathrm{d}_2^\diamond\,\mathrm{e}_1^\diamond \end{bmatrix} = -\mathbf{e}^\diamond \times \mathbf{d}^\diamond = -\widehat{\mathbf{e}^\diamond}\,\mathbf{d}^\diamond \tag{344}$$

6.5.3 Dual Quaternions

The set of dual quaternions $\mathbb{H}_{\mathrm{d}}$ is defined as $\{\mathbb{H} = \mathbb{H} + \mathbb{H}\,\epsilon \mid \epsilon^2 = -1\}$. A dual quaternion $\boldsymbol{\zeta} \in \mathbb{H}_{\mathrm{d}}$ has the form $\boldsymbol{\zeta} = \mathbf{q}_{\mathrm{r}} + \mathbf{q}_{\mathrm{d}}\,\epsilon$, with $\mathbf{q}_{\mathrm{r}}, \mathbf{q}_{\mathrm{d}} \in \mathbb{H}$. The real plus dual notation $\{1, \epsilon\}$ is not always the most convenient. A dual quaternion can also be expressed as the sum of a dual number plus a dual vector in the form $\boldsymbol{\zeta} = \mathrm{d}_0^\diamond + \mathbf{d}_{\mathrm{v}}^\diamond$, where $\mathrm{d}_0^\diamond$ is the scalar part and $\mathbf{d}_{\mathrm{v}}^\diamond = \mathrm{d}_1^\diamond\,\mathrm{i} + \mathrm{d}_2^\diamond\,\mathrm{j} + \mathrm{d}_3^\diamond\,\mathrm{i}\,\mathrm{j}$ is the vector part. Dual quaternions are however mostly represented as 8-vectors $\boldsymbol{\zeta} = [\mathbf{q}_{\mathrm{r}}, \mathbf{q}_{\mathrm{d}}]^{\mathrm{T}} = [\mathrm{q}_{0\mathrm{r}}, \mathbf{q}_{\mathrm{vr}}, \mathrm{q}_{0\mathrm{d}}, \mathbf{q}_{\mathrm{vd}}]^{\mathrm{T}} = [\mathrm{q}_{0\mathrm{r}}, \mathrm{q}_{1\mathrm{r}}, \mathrm{q}_{2\mathrm{r}}, \mathrm{q}_{3\mathrm{r}}, \mathrm{q}_{0\mathrm{d}}, \mathrm{q}_{1\mathrm{d}}, \mathrm{q}_{2\mathrm{d}}, \mathrm{q}_{3\mathrm{d}}]^{\mathrm{T}}$, which enables the use of matrix algebra for quaternion operations [31]. It is also convenient to abuse the equal operator as required to combine general, real, and pure dual quaternions.

The dual quaternion addition $\{+ : \mathbb{H}_{\mathrm{d}} \times \mathbb{H}_{\mathrm{d}} \rightarrow \mathbb{H}_{\mathrm{d}}\}$ and the scalar product $\{\cdot : \mathbb{D} \times \mathbb{H}_{\mathrm{d}} \rightarrow \mathbb{H}_{\mathrm{d}}\}$ are both straightforward and commutative:

$$\begin{aligned} \boldsymbol{\zeta}_{\mathrm{a}} + \boldsymbol{\zeta}_{\mathrm{b}} &= (\mathbf{q}_{\mathrm{ra}} + \mathbf{q}_{\mathrm{da}}\,\epsilon) + (\mathbf{q}_{\mathrm{rb}} + \mathbf{q}_{\mathrm{db}}\,\epsilon) = (\mathbf{q}_{\mathrm{ra}} + \mathbf{q}_{\mathrm{rb}}) + (\mathbf{q}_{\mathrm{da}} + \mathbf{q}_{\mathrm{db}})\,\epsilon = \begin{bmatrix} \mathbf{q}_{\mathrm{ra}} + \mathbf{q}_{\mathrm{rb}} \\ \mathbf{q}_{\mathrm{da}} + \mathbf{q}_{\mathrm{db}} \end{bmatrix} \\ &= (\mathrm{d}_{0\mathrm{a}}^\diamond + \mathbf{d}_{\mathrm{va}}^\diamond) + (\mathrm{d}_{0\mathrm{b}}^\diamond + \mathbf{d}_{\mathrm{vb}}^\diamond) = (\mathrm{d}_{0\mathrm{a}}^\diamond + \mathrm{d}_{0\mathrm{b}}^\diamond) + (\mathbf{d}_{\mathrm{va}}^\diamond + \mathbf{d}_{\mathrm{vb}}^\diamond) \end{aligned} \tag{345}$$

$$\mathrm{d}^\diamond \cdot \boldsymbol{\zeta} = (\mathrm{x} + \mathrm{y}\,\epsilon) \cdot (\mathbf{q}_{\mathrm{r}} + \mathbf{q}_{\mathrm{d}}\,\epsilon) = \mathrm{x} \cdot \mathbf{q}_{\mathrm{r}} + (\mathrm{x} \cdot \mathbf{q}_{\mathrm{d}} + \mathrm{y} \cdot \mathbf{q}_{\mathrm{r}})\,\epsilon \tag{346}$$

The multiplication of dual quaternions $\{\otimes : \mathbb{H}_{\mathrm{d}} \times \mathbb{H}_{\mathrm{d}} \rightarrow \mathbb{H}_{\mathrm{d}}\}$ is not commutative as it includes the dual vector cross product (344). Depending on how it is expressed, the similarities with the multiplication of dual numbers (339) or that of quaternions (245) are obvious [24, 31, 32]:

$$\begin{aligned} \boldsymbol{\zeta}_{\mathrm{a}} \otimes \boldsymbol{\zeta}_{\mathrm{b}} &= (\mathbf{q}_{\mathrm{ra}} + \mathbf{q}_{\mathrm{da}}\,\epsilon) \otimes (\mathbf{q}_{\mathrm{rb}} + \mathbf{q}_{\mathrm{db}}\,\epsilon) = (\mathbf{q}_{\mathrm{ra}} \otimes \mathbf{q}_{\mathrm{rb}}) + (\mathbf{q}_{\mathrm{ra}} \otimes \mathbf{q}_{\mathrm{db}} + \mathbf{q}_{\mathrm{da}} \otimes \mathbf{q}_{\mathrm{rb}})\,\epsilon \\ &= \begin{bmatrix} \mathbf{q}_{\mathrm{ra}} \otimes \mathbf{q}_{\mathrm{rb}} \\ \mathbf{q}_{\mathrm{ra}} \otimes \mathbf{q}_{\mathrm{db}} + \mathbf{q}_{\mathrm{da}} \otimes \mathbf{q}_{\mathrm{rb}} \end{bmatrix} = (\mathrm{d}_{0\mathrm{a}}^\diamond + \mathbf{d}_{\mathrm{va}}^\diamond) \otimes (\mathrm{d}_{0\mathrm{b}}^\diamond + \mathbf{d}_{\mathrm{vb}}^\diamond) \\ &= \left(\mathrm{d}_{0\mathrm{a}}^\diamond\,\mathrm{d}_{0\mathrm{b}}^\diamond - \mathbf{d}_{\mathrm{va}}^{\diamond\,\mathrm{T}}\,\mathbf{d}_{\mathrm{vb}}^\diamond\right) + \left(\mathrm{d}_{0\mathrm{a}}^\diamond\,\mathbf{d}_{\mathrm{vb}}^\diamond + \mathrm{d}_{0\mathrm{b}}^\diamond\,\mathbf{d}_{\mathrm{va}}^\diamond + \widehat{\mathbf{d}}_{\mathrm{va}}^\diamond\,\mathbf{d}_{\mathrm{vb}}^\diamond\right) \end{aligned} \tag{347}$$

Dual quaternion multiplication is also bilinear [31], based on the operators defined in (246):

$$\boldsymbol{\zeta}_{\mathrm{a}} \otimes \boldsymbol{\zeta}_{\mathrm{b}} = [\boldsymbol{\zeta}_{\mathrm{a}}]_{\mathrm{L}}\,\boldsymbol{\zeta}_{\mathrm{b}} = \begin{bmatrix} [\mathbf{a}_{\mathrm{r}}]_{\mathrm{L}} & \mathbf{O}_{4\mathrm{x}4} \\ [\mathbf{a}_{\mathrm{d}}]_{\mathrm{L}} & [\mathbf{a}_{\mathrm{r}}]_{\mathrm{L}} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{\mathrm{r}} \\ \mathbf{b}_{\mathrm{d}} \end{bmatrix} = [\boldsymbol{\zeta}_{\mathrm{b}}]_{\mathrm{R}}\,\boldsymbol{\zeta}_{\mathrm{a}} = \begin{bmatrix} [\mathbf{b}_{\mathrm{r}}]_{\mathrm{R}} & \mathbf{O}_{4\mathrm{x}4} \\ [\mathbf{b}_{\mathrm{d}}]_{\mathrm{R}} & [\mathbf{b}_{\mathrm{r}}]_{\mathrm{R}} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{\mathrm{r}} \\ \mathbf{a}_{\mathrm{d}} \end{bmatrix} \tag{348}$$

It is possible to define three different conjugates for a dual quaternion [24], based on whether it only switches the sign of the dual part as in the case of dual numbers (349), it employs the conjugates of the real and dual quaternion components (350), or a combination of both (351):

$$\boldsymbol{\zeta}^\circ = \mathbf{q}_{\mathrm{r}} - \mathbf{q}_{\mathrm{d}}\,\epsilon \quad \rightarrow \quad (\boldsymbol{\zeta}_{\mathrm{a}} \otimes \boldsymbol{\zeta}_{\mathrm{b}})^\circ = \boldsymbol{\zeta}_{\mathrm{a}}^\circ \otimes \boldsymbol{\zeta}_{\mathrm{b}}^\circ \tag{349}$$

$$\boldsymbol{\zeta}^{*} = \mathbf{q}_{\mathrm{r}}^{*} + \mathbf{q}_{\mathrm{d}}^{*}\,\epsilon \quad \rightarrow \quad (\boldsymbol{\zeta}_{\mathrm{a}} \otimes \boldsymbol{\zeta}_{\mathrm{b}})^{*} = \boldsymbol{\zeta}_{\mathrm{b}}^{*} \otimes \boldsymbol{\zeta}_{\mathrm{a}}^{*} \tag{350}$$

$$\boldsymbol{\zeta}^{\bullet} = \mathbf{q}_{\mathrm{r}}^{*} - \mathbf{q}_{\mathrm{d}}^{*}\,\epsilon \quad \rightarrow \quad (\boldsymbol{\zeta}_{\mathrm{a}} \otimes \boldsymbol{\zeta}_{\mathrm{b}})^{\bullet} = \boldsymbol{\zeta}_{\mathrm{b}}^{\bullet} \otimes \boldsymbol{\zeta}_{\mathrm{a}}^{\bullet} \tag{351}$$

Pure dual quaternions $\boldsymbol{\zeta} = 0^{\diamond} + \mathbf{d}_{\mathrm{v}}^{\diamond} \in \mathbb{H}_{\mathrm{dp}}$ are those in which its dual number is zero ($0^{\diamond}$), or in which both its real and dual parts are pure quaternions ($\mathbf{q}_{\mathrm{r}},\, \mathbf{q}_{\mathrm{d}} \in \mathbb{H}_{\mathrm{p}}$), and verify that $\boldsymbol{\zeta} = -\,\boldsymbol{\zeta}^{*}$. The dual quaternion norm is defined as $\|\boldsymbol{\zeta}\| = \sqrt{\boldsymbol{\zeta} \otimes \boldsymbol{\zeta}^{*}} = \sqrt{\mathbf{q}_{\mathrm{r}} \otimes \mathbf{q}_{\mathrm{r}}^{*} + (\mathbf{q}_{\mathrm{r}} \otimes \mathbf{q}_{\mathrm{d}}^{*} + \mathbf{q}_{\mathrm{d}} \otimes \mathbf{q}_{\mathrm{r}}^{*})\,\epsilon} \in \mathbb{D}$ [32]. Dual quaternions endowed with $\otimes$ do not form a group, because although $\boldsymbol{\zeta}_{\mathbf{1}} = \mathbf{q}_{\mathbf{1}} + \mathbf{0}\,\epsilon$ is the identity, the inverse $\boldsymbol{\zeta}^{-1} = \mathbf{q}_{\mathrm{r}}^{-1} - \mathbf{q}_{\mathrm{r}}^{-1} \otimes \mathbf{q}_{\mathrm{d}} \otimes \mathbf{q}_{\mathrm{r}}^{-1}\,\epsilon = \mathbf{q}_{\mathrm{r}}^{-1} \otimes \left(\mathbf{q}_{1} - \mathbf{q}_{\mathrm{d}} \otimes \mathbf{q}_{\mathrm{r}}^{-1}\,\epsilon\right)$ is not defined when $\mathbf{q}_{\mathrm{r}} = \mathbf{0}$. Dual quaternions endowed with addition $+$ and multiplication $\otimes$ however do form the non abelian ring $\langle \mathbb{H}_{\mathrm{d}}, +, \otimes \rangle$.

As in the case of quaternions described in section 5.4.2, the natural power of a dual quaternion $\boldsymbol{\zeta}^{\mathrm{n}}, \mathrm{n} \in \mathbb{N}$ is obtained by multiplying the dual quaternion by itself $\mathrm{n}-1$ times. The double product of a dual quaternion by a vector $\{\mathbb{H}_{\mathrm{d}} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3\}$ is defined as the product $\otimes$ of the dual quaternion by the vector by the dual quaternion conjugate, resulting in three different versions based on the conjugate definition (349, 350, 351).

6.5.4 Unit Dual Quaternion

Unit dual quaternions are those dual quaternions in which $\boldsymbol{\zeta} \otimes \boldsymbol{\zeta}^{*} = \boldsymbol{\zeta}^{*} \otimes \boldsymbol{\zeta} = \boldsymbol{\zeta}_{\mathbf{1}}$, which implies that the inverse and the conjugate coincide as $\boldsymbol{\zeta}^{-1} = \mathbf{q}_{\mathrm{r}}^{*} - \mathbf{q}_{\mathrm{r}}^{*} \otimes \mathbf{q}_{\mathrm{d}} \otimes \mathbf{q}_{\mathrm{r}}^{*}\epsilon = \mathbf{q}_{\mathrm{r}}^{*} + \mathbf{q}_{\mathrm{d}}^{*}\,\epsilon = \boldsymbol{\zeta}^{*}$. Note that the norm $\|\boldsymbol{\zeta}\|$ has a unity real part and a zero dual part [32]. Based on (347), this translates into the following two conditions:

$$\mathbf{q}_{\mathrm{r}} \otimes \mathbf{q}_{\mathrm{r}}^{*} \;=\; 1 \rightarrow \|\mathbf{q}_{\mathrm{r}}\| = 1 \tag{352}$$

$$\mathbf{q}_{\mathrm{r}} \otimes \mathbf{q}_{\mathrm{d}}^{*} + \mathbf{q}_{\mathrm{d}} \otimes \mathbf{q}_{\mathrm{r}}^{*} \;=\; 0 \rightarrow \langle \mathbf{q}_{\mathrm{r}},\, \mathbf{q}_{\mathrm{d}} \rangle = 0 \tag{353}$$

In other words, unit dual quaternions are those in which the real part $\mathbf{q}_{\mathrm{r}}$ is a unit quaternion that is also orthogonal to the dual part $\mathbf{q}_{\mathrm{d}}$.

The rigid body motion between a body frame $\mathrm{F_B}$ and a spatial frame $\mathrm{F_E}$ represented by the unit quaternion $\mathbf{q}_{\mathrm{EB}}$ and the translation $\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}}$ (section 6.2) can always be represented by the following unit dual quaternion $\boldsymbol{\zeta}_{\mathrm{EB}}$ [24, 31], where the notation is abused to consider the quaternion $\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} = [0,\, \mathbf{T}_{\mathrm{EB}}^{\mathrm{E}}]^{\mathrm{T}}$:

$$\boldsymbol{\zeta}_{\mathrm{EB}} = \mathbf{q}_{\mathrm{EB}} + \frac{\epsilon}{2}\,\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \otimes \mathbf{q}_{\mathrm{EB}} \tag{354}$$

(354) is indeed a unit dual quaternion as $\boldsymbol{\zeta}_{\mathrm{EB}} \otimes \boldsymbol{\zeta}_{\mathrm{EB}}^{*} = \boldsymbol{\zeta}_{1} = \mathbf{q}_{1} = 1$ based on $\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}*} = -\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}}$ as it is a pure quaternion. The opposite map providing the affine representation based on the unit dual quaternion is the following:

$$\mathbf{q}_{\mathrm{EB}} \;=\; \boldsymbol{\zeta}_{\mathrm{EBr}} \tag{355}$$

$$\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \;=\; 2\,\boldsymbol{\zeta}_{\mathrm{EBd}} \otimes \boldsymbol{\zeta}_{\mathrm{EBr}}^{*} \tag{356}$$

Concept	$\mathbb{SE}^3$	$\mathbb{H}_{\mathrm{d}}$	**Concept**	$\mathbb{SE}^3$	$\mathbb{H}_{\mathrm{d}}$
Lie group element	$\mathcal{M}$	$\boldsymbol{\zeta}$	Concatenation	$\circ$	$\otimes$
Identity	$\mathcal{I}_{\mathcal{M}}$	$\boldsymbol{\zeta}_{\mathbf{1}}$	Inverse	$\mathcal{M}^{-1}$	$\boldsymbol{\zeta}^{*}$
Point motion	$\mathbf{g}_{\mathcal{M}}(\mathbf{p})$	$\boldsymbol{\zeta} \otimes \boldsymbol{\zeta}_{\mathbf{p}} \otimes \boldsymbol{\zeta}^{\bullet}$	Vector motion	$\mathbf{g}_{\mathcal{M}*}(\mathbf{v})$	$\boldsymbol{\zeta} \otimes \boldsymbol{\zeta}_{\mathbf{v}} \otimes \boldsymbol{\zeta}^{\bullet}$

Table 14: Comparison between generic $\mathbb{SE}(3)$ and unit dual quaternion

The unit dual quaternion endowed with the double product can be employed to transform both points and vectors, verifying that it complies with the rigid body motion orthogonality and handedness conditions described in section 5.1. Given a point $\mathbf{p} = [\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3]^{\mathrm{T}} \in \mathbb{R}^3$, its dual quaternion representation $\boldsymbol{\zeta}_{\mathbf{p}}$ is obtained by combining the unit quaternion $\mathbf{q}_1$ as the real part and the point coordinates as the dual part, resulting in $\boldsymbol{\zeta}_{\mathbf{p}} = \mathbf{q}_1 + \epsilon\,\mathbf{p} \in \mathbb{R}^8$ [24]. In the case of a vector $\mathbf{v} = \mathbf{q} - \mathbf{p} \in \mathbb{R}^3$, its dual quaternion representation is $\boldsymbol{\zeta}_{\mathbf{v}} = \boldsymbol{\zeta}_{\mathbf{q}} - \boldsymbol{\zeta}_{\mathbf{p}} = (\mathbf{q}_1 + \epsilon\,\mathbf{q}) - (\mathbf{q}_1 + \epsilon\,\mathbf{p}) = \epsilon\,(\mathbf{q} - \mathbf{p}) = \epsilon\,\mathbf{v} \in \mathbb{R}^8$. It is then possible, based on (317) and (320), to employ the double product to transform both points and vectors between different frames:

$$\boldsymbol{\zeta}_{\mathbf{p}^{\mathrm{E}}} = \boldsymbol{\zeta}_{\mathrm{EB}} \otimes \boldsymbol{\zeta}_{\mathbf{p}^{\mathrm{B}}} \otimes \boldsymbol{\zeta}_{\mathrm{EB}}^{\bullet} \;=\; \left(\mathbf{q}_{\mathrm{EB}} + \frac{\epsilon}{2}\,\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \otimes \mathbf{q}_{\mathrm{EB}}\right) \otimes (\mathbf{q}_1 + \epsilon\,\mathbf{p}^{\mathrm{B}}) \otimes \left(\mathbf{q}_{\mathrm{EB}}^{*} + \frac{\epsilon}{2}\,\mathbf{q}_{\mathrm{EB}}^{*} \otimes \mathbf{T}_{\mathrm{EB}}^{\mathrm{E}}\right)$$

$$\begin{aligned}
&= \mathbf{q_1} + \epsilon\,(\mathbf{q}_{\text{EB}} \otimes \mathbf{p}^{\text{B}} \otimes \mathbf{q}_{\text{EB}}^* + \mathbf{T}_{\text{EB}}^{\text{E}}) = \mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{E}} && (357)\\
\boldsymbol{\zeta}_{\mathbf{p}^{\text{E}}} = \boldsymbol{\zeta}_{\text{EB}} \otimes \boldsymbol{\zeta}_{\mathbf{p}^{\text{B}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\bullet} &= (\mathbf{q}_{\text{r}} + \mathbf{q}_{\text{d}}\,\epsilon) \otimes (\mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{B}}) \otimes (\mathbf{q}_{\text{r}}^* - \mathbf{q}_{\text{d}}^*\,\epsilon)\\
&= \mathbf{q_1} + \epsilon\,(\mathbf{q}_{\text{r}} \otimes \mathbf{p}^{\text{B}} \otimes \mathbf{q}_{\text{r}}^* + \mathbf{q}_{\text{d}} \otimes \mathbf{q}_{\text{r}}^* - \mathbf{q}_{\text{r}} \otimes \mathbf{q}_{\text{d}}^*) = \mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{E}} && (358)\\
\boldsymbol{\zeta}_{\mathbf{v}^{\text{E}}} = \boldsymbol{\zeta}_{\text{EB}} \otimes \boldsymbol{\zeta}_{\mathbf{v}^{\text{B}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\bullet} &= \Big(\mathbf{q}_{\text{EB}} + \frac{\epsilon}{2}\,\mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}\Big) \otimes \epsilon\,\mathbf{v}^{\text{B}} \otimes \Big(\mathbf{q}_{\text{EB}}^* + \frac{\epsilon}{2}\,\mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}}\Big)\\
&= \epsilon\,(\mathbf{q}_{\text{EB}} \otimes \mathbf{v}^{\text{B}} \otimes \mathbf{q}_{\text{EB}}^*) = \epsilon\,\mathbf{v}^{\text{E}} && (359)\\
\boldsymbol{\zeta}_{\mathbf{v}^{\text{E}}} = \boldsymbol{\zeta}_{\text{EB}} \otimes \boldsymbol{\zeta}_{\mathbf{v}^{\text{B}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\bullet} &= (\mathbf{q}_{\text{r}} + \mathbf{q}_{\text{d}}\,\epsilon) \otimes \epsilon\,\mathbf{v}^{\text{B}} \otimes (\mathbf{q}_{\text{r}}^* - \mathbf{q}_{\text{d}}^*\,\epsilon) = \epsilon\,(\mathbf{q}_{\text{r}} \otimes \mathbf{v}^{\text{B}} \otimes \mathbf{q}_{\text{r}}^*) = \epsilon\,\mathbf{v}^{\text{E}} && (360)
\end{aligned}$$

A disadvantage of the unit dual quaternion as an $\mathbb{SE}(3)$ representation is that a different expression is required for the inverse transformation:

$$\begin{aligned}
\boldsymbol{\zeta}_{\mathbf{p}^{\text{B}}} = \boldsymbol{\zeta}_{\text{EB}}^* \otimes \boldsymbol{\zeta}_{\mathbf{p}^{\text{E}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\circ} &= \Big(\mathbf{q}_{\text{EB}}^* - \frac{\epsilon}{2}\,\mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}}\Big) \otimes (\mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{E}}) \otimes \Big(\mathbf{q}_{\text{EB}} - \frac{\epsilon}{2}\,\mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}\Big)\\
&= \mathbf{q_1} + \epsilon\,(\mathbf{q}_{\text{EB}}^* \otimes \mathbf{p}^{\text{E}} \otimes \mathbf{q}_{\text{EB}} - \mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}) = \mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{B}} && (361)\\
\boldsymbol{\zeta}_{\mathbf{p}^{\text{B}}} = \boldsymbol{\zeta}_{\text{EB}}^* \otimes \boldsymbol{\zeta}_{\mathbf{p}^{\text{E}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\circ} &= (\mathbf{q}_{\text{r}}^* + \mathbf{q}_{\text{d}}^*\,\epsilon) \otimes (\mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{E}}) \otimes (\mathbf{q}_{\text{r}} - \mathbf{q}_{\text{d}}\,\epsilon)\\
&= \mathbf{q_1} + \epsilon\,(\mathbf{q}_{\text{r}}^* \otimes \mathbf{p}^{\text{E}} \otimes \mathbf{q}_{\text{r}} + \mathbf{q}_{\text{d}}^* \otimes \mathbf{q}_{\text{r}} - \mathbf{q}_{\text{r}}^* \otimes \mathbf{q}_{\text{d}}) = \mathbf{q_1} + \epsilon\,\mathbf{p}^{\text{B}} && (362)\\
\boldsymbol{\zeta}_{\mathbf{v}^{\text{B}}} = \boldsymbol{\zeta}_{\text{EB}}^* \otimes \boldsymbol{\zeta}_{\mathbf{v}^{\text{E}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\circ} &= \Big(\mathbf{q}_{\text{EB}}^* - \frac{\epsilon}{2}\,\mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}}\Big) \otimes \epsilon\,\mathbf{v}^{\text{E}} \otimes \Big(\mathbf{q}_{\text{EB}} - \frac{\epsilon}{2}\,\mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}\Big)\\
&= \epsilon\,(\mathbf{q}_{\text{EB}}^* \otimes \mathbf{v}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}) = \epsilon\,\mathbf{v}^{\text{B}} && (363)\\
\boldsymbol{\zeta}_{\mathbf{v}^{\text{B}}} = \boldsymbol{\zeta}_{\text{EB}}^* \otimes \boldsymbol{\zeta}_{\mathbf{v}^{\text{E}}} \otimes \boldsymbol{\zeta}_{\text{EB}}^{\circ} &= (\mathbf{q}_{\text{r}}^* + \mathbf{q}_{\text{d}}^*\,\epsilon) \otimes \epsilon\,\mathbf{v}^{\text{E}} \otimes (\mathbf{q}_{\text{r}} - \mathbf{q}_{\text{d}}\,\epsilon) = \epsilon\,(\mathbf{q}_{\text{r}}^* \otimes \mathbf{v}^{\text{E}} \otimes \mathbf{q}_{\text{r}}) = \epsilon\,\mathbf{v}^{\text{B}} && (364)
\end{aligned}$$

The inverse transformation coincides with the dual quaternion conjugate provided by (350):

$$\begin{aligned}
\boldsymbol{\zeta}_{\text{BE}} = \boldsymbol{\zeta}_{\text{EB}}^{-1} = \boldsymbol{\zeta}_{\text{EB}}^* &= \mathbf{q}_{\text{BE}} + \frac{\epsilon}{2}\,\mathbf{T}_{\text{BE}}^{\text{B}} \otimes \mathbf{q}_{\text{BE}} = \mathbf{q}_{\text{EB}}^* - \frac{\epsilon}{2}\,(\mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}) \otimes \mathbf{q}_{\text{EB}}^*\\
&= \mathbf{q}_{\text{EB}}^* - \frac{\epsilon}{2}\,\mathbf{q}_{\text{EB}}^* \otimes \mathbf{T}_{\text{EB}}^{\text{E}} = \mathbf{q}_{\text{r}}^* + \mathbf{q}_{\text{d}}^*\,\epsilon && (365)
\end{aligned}$$

The concatenation of transformations is straightforward based on (319):

$$\begin{aligned}
\boldsymbol{\zeta}_{\text{EB}} = \boldsymbol{\zeta}_{\text{EN}} \otimes \boldsymbol{\zeta}_{\text{NB}} &= \Big(\mathbf{q}_{\text{EN}} + \frac{\epsilon}{2}\,\mathbf{T}_{\text{EN}}^{\text{E}} \otimes \mathbf{q}_{\text{EN}}\Big) \otimes \Big(\mathbf{q}_{\text{NB}} + \frac{\epsilon}{2}\,\mathbf{T}_{\text{NB}}^{\text{N}} \otimes \mathbf{q}_{\text{NB}}\Big)\\
&= \mathbf{q}_{\text{EN}} \otimes \mathbf{q}_{\text{NB}} + \frac{\epsilon}{2}\,(\mathbf{q}_{\text{EN}} \otimes \mathbf{T}_{\text{NB}}^{\text{N}} \otimes \mathbf{q}_{\text{NB}} + \mathbf{T}_{\text{EN}}^{\text{E}} \otimes \mathbf{q}_{\text{EN}} \otimes \mathbf{q}_{\text{NB}})\\
&= \mathbf{q}_{\text{EB}} + \frac{\epsilon}{2}\,(\mathbf{q}_{\text{EN}} \otimes \mathbf{T}_{\text{NB}}^{\text{N}} \otimes \mathbf{q}_{\text{EN}}^* + \mathbf{T}_{\text{EN}}^{\text{E}}) \otimes \mathbf{q}_{\text{EB}} = \mathbf{q}_{\text{EB}} + \frac{\epsilon}{2}\,\mathbf{T}_{\text{EB}}^{\text{E}} \otimes \mathbf{q}_{\text{EB}}\\
&= (\mathbf{q}_{\text{ENr}} + \mathbf{q}_{\text{ENd}}\,\epsilon) \otimes (\mathbf{q}_{\text{NBr}} + \mathbf{q}_{\text{NBd}}\,\epsilon)\\
&= (\mathbf{q}_{\text{ENr}} \otimes \mathbf{q}_{\text{NBr}}) + (\mathbf{q}_{\text{ENr}} \otimes \mathbf{q}_{\text{NBd}} + \mathbf{q}_{\text{ENd}} \otimes \mathbf{q}_{\text{NBr}})\,\epsilon && (366)
\end{aligned}$$

Unit dual quaternions comply with the orthogonality and handedness conditions required in section 5.1 for rigid body motions, and hence their space $\mathbb{SE}(3) = \{\boldsymbol{\zeta} = \mathbf{q}_{\text{r}} + \mathbf{q}_{\text{d}}\,\epsilon \in \mathbb{H}_{\text{d}} \mid \|\mathbf{q}_{\text{r}}\| = 1, \langle \mathbf{q}_{\text{r}},\, \mathbf{q}_{\text{d}} \rangle = 0\}$ possesses group structure under dual quaternion multiplication $\{\otimes : \mathbb{H}_{\text{d}} \times \mathbb{H}_{\text{d}} \to \mathbb{H}_{\text{d}} \mid \boldsymbol{\zeta}_{\text{a}} \otimes \boldsymbol{\zeta}_{\text{b}} \in \mathbb{H}_{\text{d}}, \forall\, \boldsymbol{\zeta}_{\text{a}},\ \boldsymbol{\zeta}_{\text{b}} \in \mathbb{H}_{\text{d}}\}$. Because of the (352) and (353) constraints, although they have dimension eight, the special euclidean group $\mathbb{SE}(3)$ defined by means of unit dual quaternions constitutes a six dimensional manifold to euclidean space $\mathbb{E}^6$ called the *image space of spatial displacements*, which can be visualized in $\mathbb{R}^4$ as follows. Expression (352) defines a unit hypersphere of three dimensions, while (353) defines the three dimensional hyperplane orthogonal to the normal at the point $\mathbf{q}_{\text{r}}$ on the hypersphere. Thus, the image space consists of the hypersphere and all of its tangent spaces, which have been translated to contain the origin. Note that in this group $\boldsymbol{\zeta}_{\mathbf{1}}$ constitutes the identity and $\boldsymbol{\zeta}^*$ the inverse.

The map from the affine representation (or homogeneous matrix) to the unit dual quaternion is surjective but not injective for the same reasons as that between the rotation matrix and the unit quaternion described in section 5.4.3, this is, the double covering of the $\mathbb{SO}(3)$ by the unit quaternion.

6.6 Half Transform Vector as Tangent Space

As indicated in section 4.3.1, the structure of the Lie algebra $\mathfrak{se}(3)$ can be obtained by time derivating its $\mathbb{SE}(3)$ Lie group constraint, this is, $\boldsymbol{\zeta} \otimes \boldsymbol{\zeta}^* = \boldsymbol{\zeta}^* \otimes \boldsymbol{\zeta} = \boldsymbol{\zeta}_{\mathbf{1}}$, leading to $\boldsymbol{\zeta}^* \otimes \dot{\boldsymbol{\zeta}} = -\big(\boldsymbol{\zeta}^* \otimes \dot{\boldsymbol{\zeta}}\big)^*$, which indicates that $\boldsymbol{\zeta}^* \otimes \dot{\boldsymbol{\zeta}}$

is in fact a pure dual quaternion, as is $\dot{\boldsymbol{\zeta}} \otimes \boldsymbol{\zeta}^*$. This results in the following particularizations of (128) and (129):

$$\boldsymbol{\Upsilon}_{\rm EB}^{\rm E\wedge} = \dot{\boldsymbol{\zeta}}_{\rm EB} \otimes \boldsymbol{\zeta}_{\rm EB}^* = -\boldsymbol{\zeta}_{\rm EB} \otimes \dot{\boldsymbol{\zeta}}_{\rm EB}^* \tag{367}$$

$$\boldsymbol{\Upsilon}_{\rm EB}^{\rm B\wedge} = \boldsymbol{\zeta}_{\rm EB}^* \otimes \dot{\boldsymbol{\zeta}}_{\rm EB} = -\dot{\boldsymbol{\zeta}}_{\rm EB}^* \otimes \boldsymbol{\zeta}_{\rm EB} \tag{368}$$

The Lie algebra velocity $\mathbf{v}^\wedge$ is known as the *half twist* $\boldsymbol{\Upsilon}^\wedge$, and as shown in (367) and (368), has the structure of a pure dual quaternion because its negative coincides with its conjugate:

$$\boldsymbol{\Upsilon}^\wedge(\mathrm{t}) = [0^\diamond + \boldsymbol{\Upsilon}_{\rm v}^\diamond(\mathrm{t})] \in \mathbb{H}_{\rm dp} \tag{369}$$

Inverting the previous equations results in the unit dual quaternion time derivative, which is linear:

$$\dot{\boldsymbol{\zeta}}_{\rm EB} = \boldsymbol{\Upsilon}_{\rm EB}^{\rm E\wedge} \otimes \boldsymbol{\zeta}_{\rm EB} = \boldsymbol{\zeta}_{\rm EB} \otimes \boldsymbol{\Upsilon}_{\rm EB}^{\rm B\wedge} \tag{370}$$

Notice that if $\boldsymbol{\zeta}(\mathrm{t}_0) = \boldsymbol{\zeta}_{\mathbf{1}}$, then $\dot{\boldsymbol{\zeta}}(\mathrm{t}_0) = \boldsymbol{\Upsilon}(\mathrm{t}_0)$, and hence the pure dual quaternion $\boldsymbol{\Upsilon}^\wedge(\mathrm{t}_0)$ provides a first order approximation of the unit dual quaternion around the identity $\boldsymbol{\zeta}_{\mathbf{1}}$:

$$\boldsymbol{\zeta}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \boldsymbol{\zeta}_1 + \boldsymbol{\Upsilon}^\wedge(\mathrm{t}_0)\ \Delta\mathrm{t} \tag{371}$$

The *space of pure dual quaternions* $\mathfrak{se}(3) = \{\boldsymbol{\Upsilon}^\wedge \in \mathbb{H}_{\rm dp} \mid \boldsymbol{\Upsilon}^\diamond \in \mathbb{D}^3\}$ is hence the *tangent space* of the unit dual quaternions at the identity $\boldsymbol{\zeta}_1$, denoted as $\mathrm{T}_{\boldsymbol{\zeta}_1}\mathcal{M}$. The *hat* $\left\{\cdot^\wedge : \mathbb{R}^6 \to \mathbb{D}^3 \to \mathfrak{se}(3) \mid \boldsymbol{\Upsilon} \to \boldsymbol{\Upsilon}^\wedge\right\}$ and *vee* $\left\{\cdot^\vee : \mathfrak{se}(3) \to \mathbb{D}^3 \to \mathbb{R}^6 \mid \left(\boldsymbol{\Upsilon}^\wedge\right)^\vee \to \boldsymbol{\Upsilon}\right\}$ operators convert the half twist vector into its pure dual quaternion form, and viceversa.

If $\boldsymbol{\zeta}(\mathrm{t}_0) \neq \boldsymbol{\zeta}_1$, the tangent space needs to be transported right multiplying by $\boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0)$ (in the case of space tangent space), or left multiplying for the local space:

$$\boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) + [\boldsymbol{\Upsilon}_{\rm EB}^{\rm E\wedge}(\mathrm{t}_0)\ \Delta\mathrm{t}] \otimes \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) = [\boldsymbol{\zeta}_1 + \boldsymbol{\Upsilon}_{\rm EB}^{\rm E\wedge}(\mathrm{t}_0)\ \Delta\mathrm{t}] \otimes \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) \tag{372}$$

$$\boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0 + \Delta\mathrm{t}) \approx \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) + \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) \otimes [\boldsymbol{\Upsilon}_{\rm EB}^{\rm B\wedge}(\mathrm{t}_0)\ \Delta\mathrm{t}] = \boldsymbol{\zeta}_{\rm EB}(\mathrm{t}_0) \otimes [\boldsymbol{\zeta}_1 + \boldsymbol{\Upsilon}_{\rm EB}^{\rm B\wedge}(\mathrm{t}_0)\ \Delta\mathrm{t}] \tag{373}$$

Note that the solution to the ordinary differential equation $\dot{\mathbf{x}}(\mathrm{t}) = \mathbf{x}(\mathrm{t}) \otimes \boldsymbol{\Upsilon}^\wedge,\ \mathbf{x}(\mathrm{t}) \in \mathbb{R}^8$, where $\boldsymbol{\Upsilon}^\wedge$ is constant, is $\mathbf{x}(\mathrm{t}) = \mathbf{x}(0)\ \mathrm{e}^{\boldsymbol{\Upsilon}^\wedge\mathrm{t}}$. Based on it, assuming $\boldsymbol{\zeta}(0) = \boldsymbol{\zeta}_{\mathbf{1}}$ as initial condition, and considering for the time being that $\boldsymbol{\Upsilon}$ is constant,

$$\boldsymbol{\zeta}(\mathrm{t}) = \mathrm{e}^{\boldsymbol{\Upsilon}^\wedge\mathrm{t}} = \boldsymbol{\zeta}_{\mathbf{1}} + \boldsymbol{\Upsilon}^\wedge\mathrm{t} + \frac{\left(\boldsymbol{\Upsilon}^\wedge\mathrm{t}\right)^2}{2!} + \cdots + \frac{\left(\boldsymbol{\Upsilon}^\wedge\mathrm{t}\right)^{\rm n}}{\mathrm{n}!} + \ldots \tag{374}$$

which is indeed a pure dual quaternion and coincides with half the twist defined in section 6.4, as proven next based on (269), (354), (367), (425), and $\mathbf{T}_{\rm EB}^{\rm E}$ being a pure quaternion:

$$\begin{aligned}
\boldsymbol{\Upsilon}_{\rm EB}^{\rm E\wedge} &= \dot{\boldsymbol{\zeta}}_{\rm EB} \otimes \left[\mathbf{q}_{\rm EB}^* + \frac{\epsilon}{2}\left(\mathbf{T}_{\rm EB}^{\rm E} \otimes \mathbf{q}_{\rm EB}\right)^*\right] = \dot{\boldsymbol{\zeta}}_{\rm EB} \otimes \left[\mathbf{q}_{\rm EB}^* - \frac{\epsilon}{2}\,\mathbf{q}_{\rm EB}^* \otimes \mathbf{T}_{\rm EB}^{\rm E}\right] \\
&= \frac{\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge}}{2} + \frac{\epsilon}{2}\left[\dot{\mathbf{T}}_{\rm EB}^{\rm E} + \mathbf{T}_{\rm EB}^{\rm E} \otimes \frac{\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge}}{2} - \frac{\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge}}{2} \otimes \mathbf{T}_{\rm EB}^{\rm E}\right] = \frac{\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge}}{2} + \frac{\epsilon}{2}\left[\dot{\mathbf{T}}_{\rm EB}^{\rm E} - \frac{\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge}}{2} \otimes \mathbf{T}_{\rm EB}^{\rm E}\right] \\
&= \frac{1}{2}\left(\boldsymbol{\omega}_{\rm EB}^{\rm E\wedge} + \epsilon\ \boldsymbol{\nu}_{\rm EB}^{\rm E\wedge}\right) = \frac{\boldsymbol{\xi}_{\rm EB}^{\rm E\wedge}}{2}
\end{aligned} \tag{375}$$

A similar process employing (368) and (424) leads to:

$$\boldsymbol{\Upsilon}_{\rm EB}^{\rm B\wedge} = \frac{1}{2}\left(\boldsymbol{\omega}_{\rm EB}^{\rm B\wedge} + \epsilon\ \boldsymbol{\nu}_{\rm EB}^{\rm B\wedge}\right) = \frac{\boldsymbol{\xi}_{\rm EB}^{\rm B\wedge}}{2} \tag{376}$$

Concept	**Lie Theory**	$\mathbb{SE}^3$
Tangent space element	$\boldsymbol{\tau}^\wedge$	$\boldsymbol{\Psi}^\wedge = [0^\diamond + \boldsymbol{\Psi}_{\rm v}^\diamond]$
Velocity element	$\mathbf{v}^\wedge$	$\boldsymbol{\Upsilon}^\wedge = [0^\diamond + \boldsymbol{\Upsilon}_{\rm v}^\diamond]$
Structure	$\wedge$	pure dual quaternion

Table 15: Comparison between generic $\mathbb{SE}(3)$ and half transform vector as tangent space

Remembering that so far $\boldsymbol{\Upsilon}^\wedge$ is constant, (374) means that any rigid body motion $\boldsymbol{\zeta}(\mathrm{t}) = e^{\boldsymbol{\Upsilon}^\wedge \mathrm{t}}$ can be realized by maintaining a constant half twist $\boldsymbol{\Upsilon}^\wedge \in \mathbb{H}_{dp}$ for a given time t. The vectors $\mathbf{n} = \boldsymbol{\omega}\,\mathrm{t}/\|\boldsymbol{\omega}\,\mathrm{t}\| = \mathbf{r}/\|\mathbf{r}\|$ and $\mathbf{k} = \boldsymbol{\nu}\,\mathrm{t}/\|\boldsymbol{\nu}\,\mathrm{t}\| = \mathbf{s}/\|\mathbf{s}\|$ indicate the half twist directions, while $\theta = \phi/2 = \|\mathbf{r}\|/2$ and $\rho/2 = \|\mathbf{s}\|/2$ represent the half twist magnitudes, respectively. This enables the definition of the *half transform vector* $\boldsymbol{\Psi}$, also known as the *exponential coordinates* of the $\mathcal{M}$ motion, as

$$\boldsymbol{\Psi} = \boldsymbol{\Upsilon}\,\mathrm{t} = \frac{1}{2}\,[\boldsymbol{\nu}\,\mathrm{t},\, \boldsymbol{\omega}\,\mathrm{t}]^T = \frac{1}{2}\,[\mathbf{s},\, \mathbf{r}]^T = \frac{1}{2}\,[\mathbf{k}\,\rho,\, \mathbf{n}\,\phi]^T = \frac{\boldsymbol{\tau}}{2} \in \mathbb{R}^6 \tag{377}$$

Note that the half transform vector $\boldsymbol{\Psi}$ belongs to the tangent space as it is a multiple of the half twist $\boldsymbol{\Upsilon} \in \mathfrak{se}(3)$, and hence tends to coincide with it as time tends to zero. The *exponential map* $\{\exp() : \mathfrak{se}(3) \rightarrow \mathbb{SE}(3) \mid \mathcal{M} = \exp(\boldsymbol{\Psi}^\wedge)\}$ and its capitalized form $\{\mathrm{Exp}() : \mathbb{R}^6 \rightarrow \mathbb{SE}(3) \mid \mathcal{M} = \mathrm{Exp}(\boldsymbol{\Psi})\}$ wrap the half transform vector around the special Euclidean group. However, the half twist $\boldsymbol{\Upsilon}^\wedge(\mathrm{t})$ in fact is not required to be constant. Given a rigid body motion represented by its unit dual quaternion $\boldsymbol{\zeta} \in \mathbb{SE}(3)$, it can be proved that there exists a not necessarily unique half transform vector $\boldsymbol{\Psi} = \boldsymbol{\tau}/2 = [\mathbf{s},\, \mathbf{r}]^T/2 = [\mathbf{k}\,\rho,\, \mathbf{n}\,\phi]^T/2$ such that $\boldsymbol{\zeta} = e^{\boldsymbol{\Psi}^\wedge}$. The exponential map is made up by a combination of (264), (335), (336), and (354).

The *logarithmic map* $\{\log() : \mathbb{SE}(3) \rightarrow \mathfrak{se}(3) \mid \boldsymbol{\Psi}^\wedge = \log(\mathcal{M})\}$ and its capitalized version $\{\mathrm{Log}() : \mathbb{SE}(3) \rightarrow \mathbb{R}^6 \mid \boldsymbol{\Psi} = \mathrm{Log}(\mathcal{M})\}$ convert unit dual quaternions into half transform vectors. It is composed by (355), (265), and (266) for the rotation part, and (356) together with (337) for the translation part.

As the vector $\boldsymbol{\Upsilon}$ represents half the twist $\boldsymbol{\xi}$, it is possible to adjust expressions (367), (368), and (370):

$$\boldsymbol{\xi}_{EB}^{E\wedge} = 2\,\dot{\boldsymbol{\zeta}}_{EB} \otimes \boldsymbol{\zeta}_{EB}^* \tag{378}$$

$$\boldsymbol{\xi}_{EB}^{B\wedge} = 2\,\boldsymbol{\zeta}_{EB}^* \otimes \dot{\boldsymbol{\zeta}}_{EB} \tag{379}$$

$$\dot{\boldsymbol{\zeta}}_{EB} = \frac{1}{2}\,\boldsymbol{\xi}_{EB}^{E\wedge} \otimes \boldsymbol{\zeta}_{EB} = \frac{1}{2}\,\boldsymbol{\zeta}_{EB} \otimes \boldsymbol{\xi}_{EB}^{B\wedge} \tag{380}$$

6.7 Screw as Tangent Space

In the section 5 analysis of rotational motion there exists two different representations for the tangent space $\mathfrak{so}(3)$. The first is the skew-symmetric angular velocity $\boldsymbol{\omega}^\wedge = \boldsymbol{\omega}^\wedge$, which converts into the rotation vector $\mathbf{r}^\wedge$ when applied during a certain amount of time (section 5.3), and represents the origin of the exponential map $\exp(\mathbf{r}^\wedge)$ (237) that transforms it into the rotation matrix $\mathbf{R}$ (section 5.2). The second is the pure quaternion half angular velocity $\boldsymbol{\Omega}^\wedge = [0,\, \boldsymbol{\omega}/2]^T$, which converts into the half rotation vector $\mathbf{h}^\wedge = [0,\, \mathbf{r}/2]^T$ (section 5.5), and represents the origin of the exponential map $\exp(\mathbf{h}^\wedge) = \exp(\mathbf{r}^\wedge/2)$ (264) that transforms it into the unit quaternion $\mathbf{q}$ (section 5.4). Note that both representations of the tangent space $\mathfrak{so}(3)$ are so similar that for all practical purposes they are considered the same, resulting in two versions of the exponential map that convert the rotation vector $\mathbf{r}$ into either the rotation matrix $\mathbf{R}$ or the unit quaternion $\mathbf{q}$.

So far the rigid body motion looks similar. There are two $\mathfrak{se}(3)$ velocities, the twist $\boldsymbol{\xi}^\wedge$ and the pure dual quaternion half twist $\boldsymbol{\Upsilon}^\wedge$, which convert into the transform vector $\boldsymbol{\tau}^\wedge$ and the half transform vector $\boldsymbol{\Psi}^\wedge$ (sections 6.4 and 6.6) when applied during an amount of time, and constitute the origins of the exponential maps that transform them into the homogeneous matrix $\mathbf{M}$ (section 6.3), the affine representation (section 6.2), or the unit dual quaternion $\boldsymbol{\zeta}$ (section 6.5). As in the rotation case, both $\mathfrak{se}(3)$ representations are so similar that for all practical purposes they are considered the same and can be interchanged in the various exponential maps.

There exists however an additional $\mathbb{SE}(3)$ representation, which also belongs to its tangent space $\mathfrak{se}(3)$, that enables the definition of a different exponential map into the unit dual quaternion that explicitly separates the influence of the motion direction from that of its magnitude, and is also indispensable for the rigid body motion powers, linear interpolation, and perturbations introduced in section 6.8.

The origin of the screw $\mathbf{S}$ lies in the fact that every rigid body motion can be realized by a rotation about an axis combined with a translation parallel to that same axis [27]. The rotation and translation can be executed simultaneously or one after another without modifying the result. It is however necessary to remark that in this case the axis, defined in section 4.2, does not necessarily pass though the origin of the frame characterizing the rigid body, in contrast to previous representations of rigid body motions in which the rotating axis, more appropriately called rotating direction, in all cases passed through the origin. Note that according to section 4.2, an axis is represented by $(\mathbf{n},\, \mathbf{m})$ and has four degrees of freedom, while if restricted to passing through the origin $\mathbf{m} = \mathbf{0}$ and the degrees of freedom are two. The line point closest to the origin responds to $\mathbf{p}_{\perp} = \widehat{\mathbf{n}}\,\mathbf{m}$.

A *screw* $\{\mathbf{S} = [\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \phi]^{\mathrm{T}} \in \mathbb{R}^8 \mid \mathbf{n}, \mathbf{m} \in \mathbb{R}^3, \mathrm{h}, \phi \in \mathbb{R}\}$ consists of an axis $(\mathbf{n},\, \mathbf{m})$, a pitch h, and a magnitude ϕ [27]. As all other $\mathbb{SE}(3)$ representations, it contains six degrees of freedom as it shares the line redundancies described in section 4.2, this is, $\|\mathbf{n}\| = 1$ and $\mathbf{n}^{\mathrm{T}}\,\mathbf{m} = 0$. It represents a rotation by an amount ϕ about the axis $(\mathbf{n},\, \mathbf{m})$ combined by a translation by an amount $\mathrm{d} = \mathrm{h}\,\phi$ parallel to axis $(\mathbf{n},\, \mathbf{m})$. If $\mathrm{h} = \infty$, the corresponding screw motion consists of a pure translation along the axis of the screw by a distance ϕ. Note that $\mathbf{n}$ and ϕ are indeed the direction and magnitude of the rotation vector $\mathbf{r} = \mathbf{n}\,\phi$ defined by (236).

Given a rigid body motion represented by the combination of rotation and translation vectors $(\mathbf{r} = \mathbf{n}\,\phi,\, \mathbf{T})$, the map converting it into a screw has two versions:

- $\mathbf{r} \neq \mathbf{0}$. If the motion contains both rotation and translation components:

$$\phi = \|\mathbf{r}\| = \phi \tag{381}$$
$$\mathbf{n} = \frac{\mathbf{r}}{\|\mathbf{r}\|} \tag{382}$$
$$\mathbf{m} = \widehat{\mathbf{p}}_{\perp}\,\mathbf{n} = \frac{1}{2}\left[\widehat{\mathbf{T}}\,\mathbf{n} + \cot\frac{\phi}{2}\,(\widehat{\mathbf{n}}\,\mathbf{T}) \times \mathbf{n}\right] \tag{383}$$
$$\mathrm{d} = \mathbf{T}^{\mathrm{T}}\,\mathbf{n} \tag{384}$$
$$\mathrm{h} = \frac{\mathrm{d}}{\phi} \tag{385}$$

- $\mathbf{r} = \mathbf{0}$. If the motion is only a translation, the screw definition changes so it contains an ∞ pitch, a magnitude equal to the translation amount, and an axis in the direction of $\mathbf{T}$ that passes through the origin. The displacement definition does not change though.

$$\mathrm{h} = \infty \tag{386}$$
$$\phi = \|\mathbf{T}\| \tag{387}$$
$$\mathbf{n} = \mathbf{T}/\phi \tag{388}$$
$$\mathbf{m} = \mathbf{0} \tag{389}$$
$$\mathrm{d} = \mathbf{T}^{\mathrm{T}}\,\mathbf{n} \tag{390}$$

The opposite map, which provides the rotation and translation vectors from a screw, also has two versions:

- $\mathrm{h} \neq \infty$. The screw contains both translation and rotation components:

$$\mathbf{T} = \mathbf{p}_{\perp} - \sin\phi\;\widehat{\mathbf{n}}\,\mathbf{p}_{\perp} - \cos\phi\,\mathbf{p}_{\perp} + \mathrm{d}\,\mathbf{n} \tag{391}$$
$$\mathbf{r} = \mathbf{n}\,\phi \tag{392}$$

- $\mathrm{h} = \infty$. The screw does not rotate:

$$\mathbf{T} = \mathbf{n}\,\phi \tag{393}$$
$$\mathbf{r} = \mathbf{0} \tag{394}$$

It is however the *exponential map* between the screw and the unit dual quaternion the one that provides a different perspective to the motion of a rigid body. It is built based on the expressions for the axis moment (383) and the unit dual quaternion (354), first as the sum of two quaternions (395) and next as that of a dual number plus a dual vector (396):

$$\begin{aligned}
\boldsymbol{\zeta} &= \left(\cos\frac{\phi}{2} + \sin\frac{\phi}{2}\,\mathbf{n}\right) + \left[-\frac{\mathrm{d}}{2}\,\sin\frac{\phi}{2} + \sin\frac{\phi}{2}\,\mathbf{m} + \frac{\mathrm{d}}{2}\,\cos\frac{\phi}{2}\,\mathbf{n}\right]\,\epsilon \qquad (395)\\
&= \mathbf{q}_{\mathrm{r}} + \mathbf{q}_{\mathrm{d}}\,\epsilon = \left[\mathrm{q}_{0\mathrm{r}}, \mathbf{q}_{\mathrm{vr}}\right]^{\mathrm{T}} + \left[\mathrm{q}_{0\mathrm{d}}, \mathbf{q}_{\mathrm{vd}}\right]^{\mathrm{T}}\,\epsilon\\
&= \left[\cos\frac{\phi}{2} - \frac{\mathrm{d}}{2}\,\sin\frac{\phi}{2}\,\epsilon\right] + \left[\sin\frac{\phi}{2}\,\mathbf{n} + \left(\sin\frac{\phi}{2}\,\mathbf{m} + \frac{\mathrm{d}}{2}\,\cos\frac{\phi}{2}\,\mathbf{n}\right)\,\epsilon\right] \qquad (396)
\end{aligned}$$

This last expression can be modified based the application of the dual number Taylor expansion (341) to the sine and cosine:

$$\begin{aligned}
\cos\mathrm{d}^{\diamond} &= \cos\left(\mathrm{x} + \mathrm{y}\,\epsilon\right) = \cos\mathrm{x} - \mathrm{y}\,\epsilon\,\sin\mathrm{x} \qquad (397)\\
\sin\mathrm{d}^{\diamond} &= \sin\left(\mathrm{x} + \mathrm{y}\,\epsilon\right) = \sin\mathrm{x} + \mathrm{y}\,\epsilon\,\cos\mathrm{x} \qquad (398)
\end{aligned}$$

This results in:

$$\boldsymbol{\zeta} = \exp\left(\mathbf{S}\right) = \cos\frac{\phi + \mathrm{d}\,\epsilon}{2} + \left(\mathbf{n} + \mathbf{m}\,\epsilon\right)\cdot\sin\frac{\phi + \mathrm{d}\,\epsilon}{2} = \cos\frac{\phi^{\diamond}}{2} + \mathbf{nm}^{\diamond}\cdot\sin\frac{\phi^{\diamond}}{2} = \cos\theta^{\diamond} + \mathbf{nm}^{\diamond}\cdot\sin\theta^{\diamond} \qquad (399)$$

where $\mathbf{nm}^{\diamond} = \mathbf{n} + \mathbf{m}\,\epsilon$ is a unit dual vector, this is, one in which its real part is a unit vector that its orthogonal to its dual part, and $\theta^{\diamond} = \frac{\phi^{\diamond}}{2} = \frac{\phi + \mathrm{d}\,\epsilon}{2}$ is a dual number. In fact unit dual quaternions can always be written as (399), which is the equivalent to (249) for unit quaternions.

A process in some aspects similar to that described in section 5.5 proves that (399) is indeed the exponential map $\{\exp() : \mathfrak{se}\,(3) \rightarrow \mathbb{SE}\,(3)\mid \mathbf{S}\in\mathbb{R}^8 \rightarrow \exp\left(\mathbf{S}\right)\in\mathbb{H}_{\mathrm{d}}\}$, which transforms screws into unit dual quaternions. Note that to do so, it is first necessary to represent the screw as the combination of a unit dual vector $\mathbf{nm}^{\diamond} = \mathbf{n} + \mathbf{m}\,\epsilon$ representing the screw axis and a dual number $\theta^{\diamond} = \frac{\phi^{\diamond}}{2} = \frac{\phi + \mathrm{d}\,\epsilon}{2}$ containing the rotation angle and translation distance about the screw axis. This map algebraically separates the line information of the screw axis from the pitch and angle values, where the dual vector $\mathbf{nm}^{\diamond}$ represents the axis of the screw motion with its direction vector and the dual angle $\theta^{\diamond} = \frac{\phi^{\diamond}}{2}$ contains both the translation length and the angle of rotation [33]. If there is no rotation ($\mathrm{h} = \infty$), the resulting unit dual quaternion responds to $\boldsymbol{\zeta} = \mathbf{q}_{\mathrm{r}} + \mathbf{q}_{\mathrm{d}}\,\epsilon = \mathbf{q}_1 + \frac{\phi}{2}\,\mathbf{n}\,\epsilon$.

Obtainment of the logarithmic map $\{\log() : \mathbb{SE}\,(3) \rightarrow \mathfrak{se}\,(3)\mid \boldsymbol{\zeta}\in\mathbb{H}_{\mathrm{d}} \rightarrow \mathbf{S}\in\mathbb{R}^8\}$ is now straightforward, resulting in expressions (400) through (403) for the case of rotation plus translation ($\mathbf{q}_{\mathrm{r}} \neq \mathbf{q}_1$):

$$\begin{aligned}
\phi &= 2\,\arctan\frac{\|\mathbf{q}_{\mathrm{vr}}\|}{\mathrm{q}_{0\mathrm{r}}} \qquad (400)\\
\mathbf{n} &= \frac{\mathbf{q}_{\mathrm{vr}}}{\|\mathbf{q}_{\mathrm{vr}}\|} \qquad (401)\\
\mathrm{d} &= -2\,\frac{\mathrm{q}_{0\mathrm{d}}}{\|\mathbf{q}_{\mathrm{vr}}\|} \qquad (402)\\
\mathbf{m} &= \left(\mathbf{q}_{\mathrm{vd}} - \frac{\mathrm{d}\;\mathrm{q}_{0\mathrm{r}}}{2}\,\mathbf{n}\right)\,\|\mathbf{q}_{\mathrm{vr}}\|^{-1} \qquad (403)
\end{aligned}$$

In the no rotation case ($\mathbf{q}_{\mathrm{r}} = \mathbf{q}_1$), the logarithmic map changes as described above. Note that both the exponential and logarithmic maps share the same surjective traits as those between the rotation vector and unit quaternion described in section 5.4.3.

Although inverting the motion by means of the screw is straightforward,

$$\mathbf{S}^{-1} = \left[\mathbf{n},\,\mathbf{m},\,\mathrm{h},\,-\phi\right]^{\mathrm{T}} \qquad (404)$$

the different $\mathbb{SE}(3)$ actions (concatenation, point rotation, vector rotation) are complex and rarely used.

6.8 Rigid Body Motion Algebraic Operations

As in the case of pure rotations described in section 5.7, the basic algebraic operations of addition, subtraction, multiplication, division, and exponentiation are not defined for objects of the special Euclidean group $\mathbb{SE}(3)$. However, all rigid body motion representations are closed under a given operation that represents the concatenation of transformations, and define not only an identity transformation that represents the lack of motion, but also an inverse operation representing the opposite movement. The concatenation of transformations and the identity and inverse operations enable the definition of the power, exponential and logarithmic operators (section 6.8.1), the screw linear interpolation (section 6.8.2), and the perturbations together with the plus and minus operators (section 6.8.3).

6.8.1 Powers, Exponentials and Logarithms

Any rigid body motion can be executed by rotating an angle ϕ about a certain fixed axis $(\mathbf{n},\, \mathbf{m})$ combined with a translation of a distance $\mathrm{d} = \mathrm{h} \cdot \phi$ along that same axis, resulting in the screw $\mathbf{S} = [\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \phi]^{\mathrm{T}}$ (section 6.7). As the rotation and translation can be executed simultaneously or one after another, taking a fraction of a screw results in $\mathrm{t}\,\mathbf{S} = \mathrm{t}\ [\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \phi]^{\mathrm{T}} = [\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \mathrm{t}\,\phi]^{\mathrm{T}}\ \forall\, \mathrm{t} \in \mathbb{R}, \mathbf{S} \in \mathfrak{se}(3)$.

The exponential map defined in (399) is named that way because it complies with the behavior of the real exponential function $\exp^{\mathrm{b}}(\mathrm{a}) = \exp(\mathrm{a} \cdot \mathrm{b})\ \forall\, \mathrm{a},\, \mathrm{b} \in \mathbb{R}$. As such, the exponential function $\{\exp() : \mathfrak{se}(3) \times \mathbb{R} \rightarrow \mathbb{SE}(3)\ \mid \mathbf{S} \in \mathfrak{se}(3), \mathrm{t} \in \mathbb{R} \rightarrow \mathcal{M}^{\mathrm{t}} = \exp(\mathrm{t}\,\mathbf{S}) \in \mathbb{SE}(3)\}$ is defined as:

$$\boldsymbol{\zeta}^{\mathrm{t}}(\mathbf{S}) = \boldsymbol{\zeta}(\mathrm{t}\,\mathbf{S}) = \exp(\mathrm{t}\,\mathbf{S}) = \cos\frac{\mathrm{t}\,\phi^{\diamond}}{2} + \mathbf{nm}^{\diamond} \cdot \sin\frac{\mathrm{t}\,\theta^{\diamond}}{2} = \cos\frac{\mathrm{t}\,\phi + \mathrm{t}\,\mathrm{d}\,\epsilon}{2} + (\mathbf{n} + \mathbf{m}\,\epsilon) \cdot \sin\frac{\mathrm{t}\,\phi + \mathrm{t}\,\mathrm{d}\,\epsilon}{2} \tag{405}$$

In a similar way, the logarithmic map defined in (400) through (403) also complies with the behavior of the real logarithmic function $\mathrm{b} \cdot \log(\mathrm{a}) = \log\left(\mathrm{a}^{\mathrm{b}}\right)\ \forall\, \mathrm{a},\, \mathrm{b} \in \mathbb{R}$. As such, the logarithmic function $\{\log : \mathbb{SE}(3) \times \mathbb{R} \rightarrow \mathfrak{se}(3)\ \mid \mathcal{M} \in \mathbb{SE}(3), \mathrm{t} \in \mathbb{R} \rightarrow \mathrm{t}\,\mathbf{S} = \log\left(\mathcal{M}^{\mathrm{t}}\right) \in \mathfrak{se}(3)\}$ is defined as:

$$\log\left(\boldsymbol{\zeta}^{\mathrm{t}}(\mathbf{S})\right) = \log(\exp(\mathrm{t}\,\mathbf{S})) = \mathrm{t}\log(\boldsymbol{\zeta}(\mathbf{S})) = \mathrm{t}\log(\exp(\mathbf{S})) = \mathrm{t}\,\mathbf{S} \tag{406}$$

It is important to remark that although other exponential maps have been defined, with inputs either the transform vector (335) or the half transform vector, they can not be employed in the exponential function, as the multiple of a transform vector $\mathrm{t}\,\boldsymbol{\tau} = (\mathrm{t}\,\mathbf{s},\, \mathrm{t}\,\mathbf{r})$ does not result in a uniform movement and hence its associated motion does not coincide with that of the same multiple of the screw $\mathrm{t}\,\mathbf{S}$.

6.8.2 Screw Linear Interpolation

Given two rigid body motions $\mathcal{M}_0,\, \mathcal{M}_1 \in \mathbb{SE}(3)$, *screw linear interpolation* (`ScLERP`) is an extension of `SLERP` (section 5.7.2) that obtains a motion function $\mathcal{M}(\mathrm{t}), \mathrm{t} \in \mathbb{R}$ that linearly interpolates from $\mathcal{M}(0) = \mathcal{M}_0$ to $\mathcal{M}(1) = \mathcal{M}_1$ in such a way that the motion occurs with constant rotation and translation velocities [24].

If employing unit dual quaternions, $\Delta\boldsymbol{\zeta}$ is according to (366) the full motion required to go from $\boldsymbol{\zeta}_0$ to $\boldsymbol{\zeta}_1$, such that $\boldsymbol{\zeta}_1 = \boldsymbol{\zeta}_0 \otimes \Delta\boldsymbol{\zeta}$, from where $\Delta\boldsymbol{\zeta} = \boldsymbol{\zeta}_0^* \otimes \boldsymbol{\zeta}_1$. The corresponding screw is then $\Delta\mathbf{S} = [\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \Delta\phi]^{\mathrm{T}} = \log(\Delta\boldsymbol{\zeta})$. Let's take a fraction of the full screw magnitude $\delta\phi = \mathrm{t}\,\Delta\phi$ and obtain the corresponding unit dual quaternion:

$$\begin{aligned}\delta\boldsymbol{\zeta} &= \exp\bigl(\mathbf{S}(\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \delta\phi)\bigr) = \exp\bigl(\mathrm{t} \cdot \mathbf{S}(\mathbf{n},\, \mathbf{m},\, \mathrm{h},\, \Delta\phi)\bigr) = \exp(\mathrm{t}\,\Delta\mathbf{S}) \\ &= \exp\bigl(\mathrm{t}\log(\Delta\boldsymbol{\zeta})\bigr) = \exp\bigl(\mathrm{t}\log(\boldsymbol{\zeta}_0^* \otimes \boldsymbol{\zeta}_1)\bigr) = (\boldsymbol{\zeta}_0^* \otimes \boldsymbol{\zeta}_1)^{\mathrm{t}}\end{aligned} \tag{407}$$

The interpolated unit dual quaternion is hence the following, which relies on (405) for its solution:

$$\boldsymbol{\zeta}(\mathrm{t}) = \boldsymbol{\zeta}_0 \otimes (\boldsymbol{\zeta}_0^* \otimes \boldsymbol{\zeta}_1)^{\mathrm{t}} \tag{408}$$

The restrictions described in section 5.7.2 intended to ensure that the rotation is executed following the shortest path are also applicable in this case.

6.8.3 Plus and Minus Operators

A perturbed rigid body motion $\widetilde{\mathcal{M}} \in \mathbb{SE}(3)$ can always be expressed as the composition of the unperturbed motion $\mathcal{M}$ with a (usually) small perturbation $\Delta\mathcal{M}$. Perturbations can be specified either at the local or body frame $\mathrm{F_B}$, this is, at the local vector space tangent to $\mathbb{SE}(3)$ at the actual pose, in which case they are known as *local perturbations*. They can also be specified at the global frame $\mathrm{F_E}$, which coincides with the vector space tangent to $\mathbb{SE}(3)$ at the origin; in this case they are known as *global perturbations*. Local perturbations appear on the right hand side of the motion composition, resulting in $\widetilde{\mathcal{M}} = \mathcal{M} \circ \Delta\mathcal{M}^{\mathrm{B}}$, while global ones appear to the left, hence $\widetilde{\mathcal{M}} = \Delta\mathcal{M}^{\mathrm{E}} \circ \mathcal{M}$.

The *plus* and *minus operators* are introduced in section 4.3 and enable operating with increments of the nonlinear $\mathbb{SE}(3)$ manifold expressed in the linear tangent vector space $\mathfrak{se}(3)$. There exist right ($\oplus$, $\ominus$) or left ($\boxplus$, $\boxminus$) versions depending on whether the increments are viewed in the local frame (right) or the global one (left). It is important to remark that although perturbations and the plus and left operators are best suited to work with small motion changes (perturbations), the expressions below are generic and work just the same no matter the size of the perturbation.

The right plus operator $\{\oplus : \mathbb{SE}(3) \times \mathfrak{se}(3) \rightarrow \mathbb{SE}(3) \mid \widetilde{\mathcal{M}} = \mathcal{M} \oplus \ \Delta\boldsymbol{\tau}^{\mathrm{B}} = \mathcal{M} \circ \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{B}})\}$ produces a motion element $\widetilde{\mathcal{M}}$ resulting from the composition of a reference motion $\mathcal{M}$ with an often small motion $\Delta\boldsymbol{\tau}^{\mathrm{B}}$, contained in the tangent space to the reference motion $\mathcal{M}$, this is, in the local space. The left plus operator $\{\boxplus : \mathfrak{se}(3) \times \mathbb{SE}(3) \rightarrow \mathbb{SE}(3) \mid \widetilde{\mathcal{M}} = \Delta\boldsymbol{\tau}^{\mathrm{E}} \boxplus \mathcal{M} = \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{E}}) \circ \mathcal{M}\}$ is similar but the often small motion $\Delta\boldsymbol{\tau}^{\mathrm{E}}$ is contained in the tangent space at the identify or global space. The expressions shown below are valid up to the first coverage of $\mathbb{SE}(3)$, this is, $\phi < \pi$. In the cases of homogeneous matrix and unit dual quaternion, the plus operator is defined as:

$$
\begin{aligned}
\widetilde{\mathbf{M}} &= \mathbf{M} \oplus \Delta\boldsymbol{\tau}^{\mathrm{B}} = \mathbf{M}\ \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{B}}) = \mathbf{M}\ \Delta\mathbf{M}^{\mathrm{B}} && (409)\\
\widetilde{\boldsymbol{\zeta}} &= \boldsymbol{\zeta} \oplus \Delta\boldsymbol{\tau}^{\mathrm{B}} = \boldsymbol{\zeta} \otimes \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{B}}/2) = \boldsymbol{\zeta} \otimes \Delta\boldsymbol{\zeta}^{\mathrm{B}} && (410)\\
\widetilde{\mathbf{M}} &= \Delta\boldsymbol{\tau}^{\mathrm{E}} \boxplus \mathbf{M} = \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{E}})\ \mathbf{M} = \Delta\mathbf{M}^{\mathrm{E}}\ \mathbf{M} && (411)\\
\widetilde{\boldsymbol{\zeta}} &= \Delta\boldsymbol{\tau}^{\mathrm{E}} \boxplus \boldsymbol{\zeta} = \mathrm{Exp}(\Delta\boldsymbol{\tau}^{\mathrm{E}}/2) \otimes \boldsymbol{\zeta} = \Delta\boldsymbol{\zeta}^{\mathrm{E}} \otimes \boldsymbol{\zeta} && (412)
\end{aligned}
$$

The right minus operator $\{\ominus : \mathbb{SE}(3) \times \mathbb{SE}(3) \rightarrow \mathfrak{se}(3) \mid \Delta\boldsymbol{\tau}^{\mathrm{B}} = \widetilde{\mathcal{M}} \ominus \mathcal{M} = \mathrm{Log}\big(\mathcal{M}^{-1} \circ \widetilde{\mathcal{M}}\big)\}$, as well as the left $\{\boxminus : \mathbb{SE}(3) \times \mathbb{SE}(3) \rightarrow \mathfrak{se}(3) \mid \Delta\boldsymbol{\tau}^{\mathrm{E}} = \widetilde{\mathcal{M}} \boxminus \mathcal{M} = \mathrm{Log}\big(\widetilde{\mathcal{M}} \circ \mathcal{M}^{-1}\big)\}$, represent the inverse operations, returning the transform vector difference $\Delta\boldsymbol{\tau}$ between two motions $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ expressed in either the local or global tangent spaces to $\mathcal{M}$.

$$
\begin{aligned}
\Delta\boldsymbol{\tau}^{\mathrm{B}} &= \widetilde{\mathbf{M}} \ominus \mathbf{M} = \mathrm{Log}\left(\mathbf{M}^{-1}\ \widetilde{\mathbf{M}}\right) = \mathrm{Log}(\Delta\mathbf{M}^{\mathrm{B}}) && (413)\\
\Delta\boldsymbol{\tau}^{\mathrm{B}} &= \widetilde{\boldsymbol{\zeta}} \ominus \boldsymbol{\zeta} = 2\ \mathrm{Log}\left(\boldsymbol{\zeta}^{*} \otimes \widetilde{\boldsymbol{\zeta}}\right) = 2\ \mathrm{Log}(\Delta\boldsymbol{\zeta}^{\mathrm{B}}) && (414)\\
\Delta\boldsymbol{\tau}^{\mathrm{E}} &= \widetilde{\mathbf{M}} \boxminus \mathbf{M} = \mathrm{Log}\left(\widetilde{\mathbf{M}}\ \mathbf{M}^{-1}\right) = \mathrm{Log}(\Delta\mathbf{M}^{\mathrm{E}}) && (415)\\
\Delta\boldsymbol{\tau}^{\mathrm{E}} &= \widetilde{\boldsymbol{\zeta}} \boxminus \boldsymbol{\zeta} = 2\ \mathrm{Log}\left(\widetilde{\boldsymbol{\zeta}} \otimes \boldsymbol{\zeta}^{*}\right) = 2\ \mathrm{Log}(\Delta\boldsymbol{\zeta}^{\mathrm{E}}) && (416)
\end{aligned}
$$

If the $\Delta\boldsymbol{\tau}$ perturbation is small, the (333) and (374) Taylor expansions can be truncated, resulting in the following expressions, valid for both the body frame ($\Delta\boldsymbol{\tau}^{\mathrm{B}}$) or the global one ($\Delta\boldsymbol{\tau}^{\mathrm{E}}$):

$$
\begin{aligned}
\Delta\mathbf{M} = \mathrm{Exp}(\Delta\boldsymbol{\tau}) &\approx \mathbf{I}_4 + \Delta\boldsymbol{\tau}^{\wedge} = \mathbf{I}_4 + [\mathbf{k}\,\Delta\rho,\ \mathbf{n}\,\Delta\phi]^{\wedge} && (417)\\
\Delta\boldsymbol{\zeta} = \exp(\Delta\boldsymbol{\tau}/2) &\approx \boldsymbol{\zeta}_{\mathbf{1}} + \Delta\boldsymbol{\tau}^{\wedge}/2 = \big[[1, \mathbf{n}\,\Delta\phi/2]^{\mathrm{T}} + \epsilon\,\mathbf{k}\,\Delta\rho/2\big]^{\wedge} && (418)
\end{aligned}
$$

6.9 Rigid Body Motion Time Derivative and Twist

Let's consider a moving rigid body $\mathcal{M}(\mathrm{t}) \in \mathbb{SE}(3), \mathrm{t} \in \mathbb{R}$ and compute its derivative with time, which belongs to neither $\mathbb{SE}(3)$ nor $\mathfrak{se}(3)$ but to the Euclidean space of the chosen motion representation, $\mathbb{R}^{4\mathrm{x}4}$ for the homogeneous matrix and $\mathbb{H}_d$ for the unit dual quaternion:

$$\dot{\mathcal{M}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\mathcal{M}(\mathrm{t}+\Delta \mathrm{t}) - \mathcal{M}(\mathrm{t})}{\Delta \mathrm{t}} \tag{419}$$

Considering the time modified motion $\mathcal{M}(\mathrm{t}+\Delta \mathrm{t})$ as the perturbed state (section 6.8.3), the resulting time derivatives for the homogeneous matrix and unit dual quaternion representations are the following:

$$\begin{aligned} \dot{\mathbf{M}}(\mathrm{t}) &= \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{M}\,\Delta\mathbf{M}^{\mathrm{B}} - \mathbf{M}}{\Delta \mathrm{t}} \approx \lim_{\Delta \mathrm{t} \to 0} \frac{\mathbf{M}\left[\left(\mathbf{I}_4 + [\mathbf{k}^{\mathrm{B}}\,\Delta\rho,\,\mathbf{n}^{\mathrm{B}}\,\Delta\phi]^{\wedge}\right) - \mathbf{I}_4\right]}{\Delta \mathrm{t}} \\ &= \mathbf{M} \lim_{\Delta \mathrm{t} \to 0} \frac{[\Delta\rho\,\mathbf{k}^{\mathrm{B}},\,\Delta\phi\,\mathbf{n}^{\mathrm{B}}]^{\wedge}}{\Delta \mathrm{t}} \end{aligned} \tag{420}$$

$$\begin{aligned} \dot{\boldsymbol{\zeta}}(\mathrm{t}) &= \lim_{\Delta \mathrm{t} \to 0} \frac{\boldsymbol{\zeta} \otimes \Delta\boldsymbol{\zeta}^{\mathrm{B}} - \boldsymbol{\zeta}}{\Delta \mathrm{t}} \approx \lim_{\Delta \mathrm{t} \to 0} \frac{\boldsymbol{\zeta} \otimes \left[\left([1, \mathbf{n}^{\mathrm{B}}\,\Delta\phi/2]^{\mathrm{T}} + \epsilon\,\mathbf{k}^{\mathrm{B}}\,\Delta\rho/2\right)^{\wedge} - \boldsymbol{\zeta}_{\mathbf{1}}\right]}{\Delta \mathrm{t}} \\ &= \boldsymbol{\zeta} \otimes \lim_{\Delta \mathrm{t} \to 0} \frac{[\mathbf{n}^{\mathrm{B}}\,\Delta\phi/2 + \epsilon\,\mathbf{k}^{\mathrm{B}}\,\Delta\rho/2]^{\wedge}}{\Delta \mathrm{t}} \end{aligned} \tag{421}$$

Similar expressions based on $\boldsymbol{\tau}^{\mathrm{E}} = [\Delta\rho\,\mathbf{k}^{\mathrm{E}},\,\Delta\phi\,\mathbf{n}^{\mathrm{E}}]^{\mathrm{T}}$ can be found if left multiplying by the perturbation instead of right multiplying. The $\dot{\mathbf{M}}(\mathrm{t})$ and $\dot{\boldsymbol{\zeta}}(\mathrm{t})$ expressions (329) and (380) are then directly obtained when defining the *body twist* $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}}$ as the time derivative of the transform vector $\boldsymbol{\tau}^{\mathrm{B}}$ when viewed in local or body frame $\mathrm{F_B}$, and the *spatial twist* $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}}$ as the time derivative of the transform vector $\boldsymbol{\tau}^{\mathrm{E}}$ when viewed in global or spatial frame $\mathrm{F_E}$:

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}}(\mathrm{t}) = \Delta\dot{\boldsymbol{\tau}}^{\mathrm{B}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\Delta\boldsymbol{\tau}^{\mathrm{B}}}{\Delta \mathrm{t}} = \lim_{\Delta \mathrm{t} \to 0} \frac{[\mathbf{k}^{\mathrm{B}}\,\Delta\rho,\,\mathbf{n}^{\mathrm{B}}\,\Delta\phi]^{\mathrm{T}}}{\Delta \mathrm{t}} \tag{422}$$

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}}(\mathrm{t}) = \Delta\dot{\boldsymbol{\tau}}^{\mathrm{E}}(\mathrm{t}) = \lim_{\Delta \mathrm{t} \to 0} \frac{\Delta\boldsymbol{\tau}^{\mathrm{E}}}{\Delta \mathrm{t}} = \lim_{\Delta \mathrm{t} \to 0} \frac{[\mathbf{k}^{\mathrm{E}}\,\Delta\rho,\,\mathbf{n}^{\mathrm{E}}\,\Delta\phi]^{\mathrm{T}}}{\Delta \mathrm{t}} \tag{423}$$

The twist $\boldsymbol{\xi} = [\boldsymbol{\nu},\,\boldsymbol{\omega}]^{\mathrm{T}}$ represents the motion velocity and is composed by the angular velocity $\boldsymbol{\omega}$ defined in section 5.8 and the linear velocity $\boldsymbol{\nu}$. The twist physical meaning is revealed by obtaining its expressions when viewed in both the local and spatial frames. The body twist $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge} \in \mathfrak{se}(3)$ corresponding to the rigid body motion $\mathbf{M}_{\mathrm{EB}}(\mathrm{t}) \in \mathbb{SE}(3)$ responds to (327):

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge} = \begin{bmatrix} \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{B}} & \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}} \\ 0 & 0 \end{bmatrix} = \mathbf{M}_{\mathrm{EB}}^{-1}\,\dot{\mathbf{M}}_{\mathrm{EB}} = \begin{bmatrix} \mathbf{R}_{\mathrm{EB}}^{\mathrm{T}} & -\mathbf{R}_{\mathrm{EB}}^{\mathrm{T}}\,\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\mathbf{R}}_{\mathrm{EB}} & \dot{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathrm{EB}}^{\mathrm{T}}\,\dot{\mathbf{R}}_{\mathrm{EB}} & \mathbf{R}_{\mathrm{EB}}^{\mathrm{T}}\,\dot{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 0 \end{bmatrix} \tag{424}$$

Its physical interpretation is that the angular component $\boldsymbol{\omega}_{\mathrm{EB}}^{\mathrm{B}}$ is indeed the (230) angular velocity $\boldsymbol{\omega}_{\mathrm{EB}}$ as viewed from the body frame, and the linear component $\boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}}$ is the linear velocity of the body frame origin $\dot{\mathbf{T}}_{\mathrm{EB}}$ also viewed from the body frame (226, 227) [27]. The global twist $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge} \in \mathfrak{se}(3)$ is determined by means of (326):

$$\begin{aligned} \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge} &= \begin{bmatrix} \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{E}} & \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 0 \end{bmatrix} = \dot{\mathbf{M}}_{\mathrm{EB}}\,\mathbf{M}_{\mathrm{EB}}^{-1} = \begin{bmatrix} \dot{\mathbf{R}}_{\mathrm{EB}} & \dot{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\mathrm{EB}}^{\mathrm{T}} & -\mathbf{R}_{\mathrm{EB}}^{\mathrm{T}}\,\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \dot{\mathbf{R}}_{\mathrm{EB}}\,\mathbf{R}_{\mathrm{EB}}^{\mathrm{T}} & \dot{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}} - \dot{\mathbf{R}}_{\mathrm{EB}}\,\mathbf{R}_{\mathrm{EB}}^{\mathrm{T}}\,\mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{425}$$

The $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}$ physical interpretation is not intuitive, however. While the angular component $\boldsymbol{\omega}_{\mathrm{EB}}^{\mathrm{E}}$ is the (229) angular velocity $\boldsymbol{\omega}_{\mathrm{EB}}$ as viewed from the spatial frame, its linear component $\boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}}$ is not the velocity of the body frame origin $\dot{\mathbf{T}}_{\mathrm{EB}}$ viewed in the $\mathrm{F_E}$ frame ($\dot{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}}$), but the velocity, viewed in the spatial frame, of a possibly imaginary point of the rigid body which at time t is traveling through the origin of the spatial frame [27].

Note that the transformation or motion of the twist (relationship between $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}}$ and $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}}$), and hence that of the linear velocities $\boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}}$ and $\boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}}$, is not given by the motion action $\mathbf{g}_{\mathcal{M}*}$ (314) but by the adjoint map $\mathbf{Ad}_{\mathcal{M}}$ described in section 6.11. Unlike in the case of rotations, these two maps do not coincide.

6.10 Rigid Body Motion Point Velocity

There exists a direct relationship between the velocity of a point belonging to a rigid body and the elements of its tangent space, this is, the twist in $\mathfrak{se}(3)$. This relationship is independent of the $\mathbb{SE}(3)$ representation, although the homogeneous matrix is employed in the expressions below. If $\bar{\mathbf{p}}^{\mathrm{B}} = [\mathbf{p}^{\mathrm{B}},\, 1]^{\mathrm{T}}$ are the fixed coordinates of a point belonging to the $\mathrm{F_B}$ rigid body, the point spatial coordinates $\bar{\mathbf{p}}^{\mathrm{E}} = [\mathbf{p}^{\mathrm{E}},\, 1]^{\mathrm{T}}$ can be obtained by means of (321):

$$\bar{\mathbf{p}}^{\mathrm{E}}(\mathrm{t}) = \mathbf{g}_{\mathcal{M}_{\mathrm{EB}}(\mathrm{t})}\left(\bar{\mathbf{p}}^{\mathrm{B}}\right) = \mathbf{M}_{\mathrm{EB}}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{B}} \tag{426}$$

The velocity of a point is the time derivative of its spatial or global coordinates. As $\bar{\mathbf{p}}$ is fixed to $\mathrm{F_B}$, its time derivative is zero $\left(\dot{\bar{\mathbf{p}}}^{\mathrm{B}} = \mathbf{0}\right)$, so its velocity viewed in the spatial frame responds to:

$$\bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t}) = \dot{\bar{\mathbf{p}}}^{\mathrm{E}}(\mathrm{t}) = \dot{\mathbf{M}}_{\mathrm{EB}}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{B}} \tag{427}$$

Although $\dot{\mathbf{M}}_{\mathrm{EB}}$ maps the point body coordinates to its spatial velocity per (427), its high dimension makes it inefficient. By making use of the spatial and body twists $(\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge},\, \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge})$ introduced in (329), the velocity of a point $\bar{\mathbf{p}}^{\mathrm{B}}$ viewed in $\mathrm{F_E}$ can be obtained as follows:

$$\begin{aligned} \bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t}) &= \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}(\mathrm{t})\ \mathbf{M}_{\mathrm{EB}}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{B}} = \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{E}}(\mathrm{t}) && (428)\\ \bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t}) &= \mathbf{M}_{\mathrm{EB}}(\mathrm{t})\ \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{B}} && (429) \end{aligned}$$

The velocity of $\bar{\mathbf{p}}^{\mathrm{B}}$ viewed in $\mathrm{F_B}$ can then be obtained by means of the vector action map:

$$\bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{B}}(\mathrm{t}) = \mathbf{g}_{\mathcal{M}_{\mathrm{EB}(\mathrm{t})*}}^{-1}\left(\bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t})\right) = \mathbf{M}_{\mathrm{EB}}^{-1}(\mathrm{t})\ \bar{\mathbf{v}}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t}) = \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}(\mathrm{t})\ \bar{\mathbf{p}}^{\mathrm{B}} \tag{430}$$

Returning to cartesian coordinates and introducing the angular and linear components of the twist results in:

$$\begin{aligned} \mathbf{v}_{\mathrm{P}}^{\mathrm{E}}(\mathrm{t}) &= \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{E}}(\mathrm{t})\ \mathbf{p}^{\mathrm{E}}(\mathrm{t}) + \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}}(\mathrm{t}) && (431)\\ \mathbf{v}_{\mathrm{P}}^{\mathrm{B}}(\mathrm{t}) &= \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{B}}(\mathrm{t})\ \mathbf{p}^{\mathrm{B}} + \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}}(\mathrm{t}) && (432) \end{aligned}$$

The point velocity is hence the result of the sum of the linear velocity and the cross product between the angular velocity and the point coordinates.

6.11 Rigid Body Motion Adjoint

The *adjoint map* of a Lie group is defined in section 4.3.3 as an action of the Lie group on its own Lie algebra that converts between the local tangent space and that at the identity. In the case of rigid body motion, both the transform vector and the twist belong to the tangent space, so $\left\{\mathbf{Ad}\,() : \mathbb{SE}(3) \times \mathfrak{se}(3) \to \mathfrak{se}(3) \mid \mathbf{Ad}_{\mathcal{M}}(\boldsymbol{\tau}^{\wedge}) = \mathcal{M} \circ \boldsymbol{\tau}^{\wedge} \circ \mathcal{M}^{-1},\ \mathbf{Ad}_{\mathcal{M}}\left(\boldsymbol{\xi}^{\wedge}\right) = \mathcal{M} \circ \boldsymbol{\xi}^{\wedge} \circ \mathcal{M}^{-1}\right\}$. This is equivalent to $\boldsymbol{\zeta} \otimes \boldsymbol{\xi}^{\wedge} \otimes \boldsymbol{\zeta}^{*}$ for unit dual quaternions or $\mathbf{M}\,\boldsymbol{\xi}^{\wedge}\,\mathbf{M}^{-1}$ for homogeneous matrices, which represents the similarity transformation[35] between the spatial and body twists $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge}$ and $\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}$:

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}\wedge} = \mathbf{M}_{\mathrm{EB}}\ \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}\wedge}\ \mathbf{M}_{\mathrm{EB}}^{-1} \to \begin{cases} \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{E}} = \mathbf{R}_{\mathrm{EB}}\ \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{B}}\ \mathbf{R}_{\mathrm{EB}}^{\mathrm{T}} = \mathbf{Ad}_{\mathcal{R}_{\mathrm{EB}}}\left(\widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{B}}\right) \\ \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}} = \mathbf{R}_{\mathrm{EB}}\ \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}} - \widehat{\boldsymbol{\omega}}_{\mathrm{EB}}^{\mathrm{E}}\ \mathbf{T}_{\mathrm{EB}}^{\mathrm{E}} = \mathbf{R}_{\mathrm{EB}}\ \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}} + \widehat{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}}\ \boldsymbol{\omega}_{\mathrm{EB}}^{\mathrm{E}} \end{cases} \tag{433}$$

The application of the vee operator results in the adjoint matrix:

$$\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{E}} = \begin{bmatrix} \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{E}} \\ \boldsymbol{\omega}_{\mathrm{EB}}^{\mathrm{E}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathrm{EB}} & \widehat{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}}\ \mathbf{R}_{\mathrm{EB}} \\ 0 & \mathbf{R}_{\mathrm{EB}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\nu}_{\mathrm{EB}}^{\mathrm{B}} \\ \boldsymbol{\omega}_{\mathrm{EB}}^{\mathrm{B}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathrm{EB}} & \widehat{\mathbf{T}}_{\mathrm{EB}}^{\mathrm{E}}\ \mathbf{R}_{\mathrm{EB}} \\ 0 & \mathbf{R}_{\mathrm{EB}} \end{bmatrix} \boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}} = \mathbf{Ad}_{\mathcal{M}_{\mathrm{EB}}}\,\boldsymbol{\xi}_{\mathrm{EB}}^{\mathrm{B}} \tag{434}$$

[35] Two square matrices $\mathbf{A}$ and $\mathbf{B}$ are called similar if $\mathbf{B} = \mathbf{P}^{-1}\ \mathbf{A}\ \mathbf{P}$ for some invertible square matrix $\mathbf{P}$.

As stated above, note that the adjoint map (434) is different than the vector action $\mathbf{g}_{\mathcal{M}*}$ (314), unlike the case of rotational motion described in section 5.10, in which they coincide.

A similar process leads to the inverse adjoint matrix ($\mathbf{Ad}_{\mathcal{M}}^{-1}\,\boldsymbol{\xi} = \mathbf{Ad}_{\mathcal{M}^{-1}}\,\boldsymbol{\xi}$):

$$\boldsymbol{\xi}_{\text{EB}}^{\text{B}} = \mathbf{Ad}_{\mathcal{M}_{\text{EB}}}^{-1}\,\boldsymbol{\xi}_{\text{EB}}^{\text{E}} = \begin{bmatrix} \mathbf{R}_{\text{EB}}^{\text{T}} & -\mathbf{R}_{\text{EB}}^{\text{T}}\,\widehat{\mathbf{T}}_{\text{EB}}^{\text{E}} \\ 0 & \mathbf{R}_{\text{EB}}^{\text{T}} \end{bmatrix} \boldsymbol{\xi}_{\text{EB}}^{\text{E}} \tag{435}$$

6.12 Rigid Body Motion Uncertainty and Covariance

Following the analysis of uncertainty on Lie groups presented in section 4.3.5, the definitions of local and global autocovariances for $\mathbb{SE}(3)$ elements around a nominal or expected rotation $\text{E}\,[\mathcal{M}] = \boldsymbol{\mu}_{\mathcal{M}} \in \mathbb{SE}(3)$ are the following:

$$\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{B}} = \text{E}\left[\Delta\boldsymbol{\tau}^{\text{B}}\,\Delta\boldsymbol{\tau}^{\text{BT}}\right] = \text{E}\left[(\mathcal{M} \ominus \boldsymbol{\mu}_{\mathcal{M}})\,(\mathcal{M} \ominus \boldsymbol{\mu}_{\mathcal{M}})^{\text{T}}\right] \quad \in \mathbb{R}^{6\text{x}6} \tag{436}$$

$$\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{E}} = \text{E}\left[\Delta\boldsymbol{\tau}^{\text{E}}\,\Delta\boldsymbol{\tau}^{\text{ET}}\right] = \text{E}\left[(\mathcal{M} \boxminus \boldsymbol{\mu}_{\mathcal{M}})\,(\mathcal{M} \boxminus \boldsymbol{\mu}_{\mathcal{M}})^{\text{T}}\right] \quad \in \mathbb{R}^{6\text{x}6} \tag{437}$$

Note that although the notation refers to the covariance of the rigid body motion manifold $\mathcal{M} \in \mathbb{SE}(3)$, the definition in fact refers to the covariance of the transform vectors $\Delta\boldsymbol{\tau}^{\text{B}}$ or $\Delta\boldsymbol{\tau}^{\text{E}}$ located in the tangent space, with its dimension (6) matching the number of degrees of freedom of the $\mathbb{SE}(3)$ manifold. The relationship between the local and global autocovariances responds to:

$$\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{E}} = \mathbf{Ad}_{\mathcal{M}_{\text{EB}}}\,\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{B}}\,\mathbf{Ad}_{\mathcal{M}_{\text{EB}}}^{\text{T}} \tag{438}$$

Given a function $\{\text{f} : \mathcal{M} \to \mathcal{N} \mid \mathcal{N} = \text{f}\,(\mathcal{M}) \in \mathbb{SE}(3),\, \forall \mathcal{M} \in \mathbb{SE}(3)\}$ between two rigid body motions, the covariances are propagated as follows:

$$\mathbf{C}_{\mathcal{N}\mathcal{N}}^{\text{B}} = \mathbf{J}_{\oplus\,\mathcal{M}}^{\oplus\,\text{f}\,(\mathcal{M})}\,\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{B}}\,\mathbf{J}_{\oplus\,\mathcal{M}}^{\oplus\,\text{f}\,(\mathcal{M}),\text{T}} \quad \in \mathbb{R}^{6\text{x}6} \tag{439}$$

$$\mathbf{C}_{\mathcal{N}\mathcal{N}}^{\text{E}} = \mathbf{J}_{\boxplus\,\mathcal{M}}^{\boxplus\,\text{f}\,(\mathcal{M})}\,\mathbf{C}_{\mathcal{M}\mathcal{M}}^{\text{E}}\,\mathbf{J}_{\boxplus\,\mathcal{M}}^{\boxplus\,\text{f}\,(\mathcal{M}),\text{T}} \quad \in \mathbb{R}^{6\text{x}6} \tag{440}$$

6.13 Rigid Body Motion Jacobians

Lie group jacobians are introduced in section 4.4 based on the right and left Lie group derivatives of section 4.3.4, and in this section are customized for the $\mathbb{SE}(3)$ case, with table 16 representing the particularization of table 1 to the case of rigid body motions. The various jacobians listed in table 16 have been obtained by means of the chain rule, the expressions already introduced in this article, and those of section 4.3. Note that although in many cases the results internally include the rotation matrix, all jacobians are generic and do not depend on the specific $\mathbb{SE}(3)$ parameterization.

In addition to the adjoint matrix, two other jacobians are of particular importance as they appear repeatedly in table 16. These are the right and left jacobians of the capitalized exponential function, also known as simply the *right jacobian* $\text{J}_{\text{R}}\,(\boldsymbol{\tau})$ and the *left jacobian* $\text{J}_{\text{L}}\,(\boldsymbol{\tau})$, and they evaluate the variation of the $\mathfrak{se}(3)$ tangent space provided by the output of the $\text{Exp}\,(\boldsymbol{\tau})$ map (locally for J_{R} and globally for J_{L}) while moving along the $\mathbb{SE}\,(3)$ manifold with respect to the (Euclidean) variations within the original tangent space provided by $\boldsymbol{\tau}$. Their closed forms as well as those of their inverses are included in table 16, and have been obtained from [34]; they are based on the $\mathbf{Q}\,(\boldsymbol{\tau})$ matrix:

$$\begin{aligned}
\mathbf{Q}\,(\boldsymbol{\tau}) &= \mathbf{Q}\,(\mathbf{s},\,\mathbf{r}) = \mathbf{Q}\,(\mathbf{k}\,\rho,\,\mathbf{n}\,\phi) = \frac{\widehat{\mathbf{s}}}{2} + \frac{\phi - \sin\phi}{\phi^3}\,(\widehat{\mathbf{r}}\,\widehat{\mathbf{s}} + \widehat{\mathbf{s}}\,\widehat{\mathbf{r}} + \widehat{\mathbf{r}}\,\widehat{\mathbf{s}}\,\widehat{\mathbf{r}}) \\
&- \frac{1 - \phi^2/2 - \cos\phi}{\phi^4}\left(\widehat{\mathbf{r}}^2\,\widehat{\mathbf{s}} + \widehat{\mathbf{s}}\,\widehat{\mathbf{r}}^2 - 3\,\widehat{\mathbf{r}}\,\widehat{\mathbf{s}}\,\widehat{\mathbf{r}}\right) \\
&- \frac{1}{2}\left[\frac{1 - \phi^2/2 - \cos\phi}{\phi^4} - 3\,\frac{\phi - \sin\phi - \phi^3/6}{\phi^5}\right]\left(\widehat{\mathbf{r}}\,\widehat{\mathbf{s}}\,\widehat{\mathbf{r}}^2 + \widehat{\mathbf{r}}^2\,\widehat{\mathbf{s}}\,\widehat{\mathbf{r}}\right) \quad \in \mathbb{R}^{3\text{x}3}
\end{aligned} \tag{441}$$

It is also worth noting the special importance of the $\mathbf{J}^{+\,\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{p})}_{+\,\boldsymbol{\tau}}$ jacobian present at the bottom of table 16, which represents the derivative of a transformed point with respect to perturbations in the Euclidean tangent space (not on the curved manifold) that generates the motion, as it enables tangent space optimization by calculus methods designed exclusively for Euclidean spaces.

Jacobian		**Table 1**		**Expression**	**Size**
$\mathbf{J}^{\oplus\,\mathcal{M}^{-1}}_{\oplus\,\mathcal{M}}$	$=$	$-\mathbf{Ad}_{\mathcal{M}}$	$=$	$-\left[\mathbf{R}\quad \widehat{\mathbf{T}}\,\mathbf{R};\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\mathcal{M}^{-1}}_{\boxplus\,\mathcal{M}}$	$=$	$-\mathbf{Ad}_{\mathcal{M}}^{-1}$	$=$	$-\left[\mathbf{R}^{\mathrm{T}}\quad -\mathbf{R}^{\mathrm{T}}\,\widehat{\mathbf{T}};\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}^{\mathrm{T}}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\oplus\,\mathcal{M}\circ\mathcal{N}}_{\oplus\,\mathcal{M}}$	$=$	$\mathbf{Ad}_{\mathcal{N}}^{-1}$	$=$	$\left[\mathbf{R}_{\mathcal{N}}^{\mathrm{T}}\quad -\mathbf{R}_{\mathcal{N}}^{\mathrm{T}}\,\widehat{\mathbf{T}}_{\mathcal{N}};\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}_{\mathcal{N}}^{\mathrm{T}}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\mathcal{M}\circ\mathcal{N}}_{\boxplus\,\mathcal{M}}$	$=$	$\mathbf{I}$	$=$	$\mathbf{I}_{6\mathrm{x}6}$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\oplus\,\mathcal{M}\circ\mathcal{N}}_{\oplus\,\mathcal{N}}$	$=$	$\mathbf{I}$	$=$	$\mathbf{I}_{6\mathrm{x}6}$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\mathcal{M}\circ\mathcal{N}}_{\boxplus\,\mathcal{N}}$	$=$	$\mathbf{Ad}_{\mathcal{M}}$	$=$	$\left[\mathbf{R}_{\mathcal{M}}\quad \widehat{\mathbf{T}}_{\mathcal{M}}\,\mathbf{R}_{\mathcal{M}};\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}_{\mathcal{M}}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\oplus\,\mathrm{Exp}(\boldsymbol{\tau})}_{+\,\boldsymbol{\tau}}$	$=$	$\mathrm{J}_{\mathrm{R}}(\boldsymbol{\tau})$	$=$	$\mathbf{J}_{\mathrm{L}}(-\boldsymbol{\tau}) = \mathbf{J}_{\mathrm{L}}(-\mathbf{s},\ -\mathbf{r})$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathrm{J}_{\mathrm{R}}^{-1}(\boldsymbol{\tau})$			$=$	$\mathbf{J}_{\mathrm{L}}^{-1}(-\boldsymbol{\tau}) = \mathbf{J}_{\mathrm{L}}^{-1}(-\mathbf{s},\ -\mathbf{r})$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\mathrm{Exp}(\boldsymbol{\tau})}_{+\,\boldsymbol{\tau}}$	$=$	$\mathrm{J}_{\mathrm{L}}(\boldsymbol{\tau})$	$=$	$\left[\mathbf{J}_{\mathrm{L}}(\mathbf{r})\quad \mathbf{Q}(\boldsymbol{\tau});\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{J}_{\mathrm{L}}(\mathbf{r})\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathrm{J}_{\mathrm{L}}^{-1}(\boldsymbol{\tau})$			$=$	$\left[\mathbf{J}_{\mathrm{L}}^{-1}(\mathbf{r})\quad -\mathbf{J}_{\mathrm{L}}^{-1}(\mathbf{r})\ \mathbf{Q}(\boldsymbol{\tau})\ \mathbf{J}_{\mathrm{L}}^{-1}(\mathbf{r});\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{J}_{\mathrm{L}}^{-1}(\mathbf{r})\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{Log}(\mathcal{M})}_{\oplus\,\mathcal{M}}$	$=$	$\mathrm{J}_{\mathrm{R}}^{-1}\big(\mathrm{Log}(\mathcal{M})\big)$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{Log}(\mathcal{M})}_{\boxplus\,\mathcal{M}}$	$=$	$\mathrm{J}_{\mathrm{L}}^{-1}\big(\mathrm{Log}(\mathcal{M})\big)$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\oplus\,\mathcal{M}\oplus\boldsymbol{\tau}}_{\oplus\,\mathcal{M}}$	$=$	$\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}^{-1}$	$=$	$\left[\mathbf{R}(\mathbf{r})^{\mathrm{T}}\quad -\mathbf{R}(\mathbf{r})^{\mathrm{T}}\,\widehat{\mathbf{T}}(\mathbf{s},\mathbf{r});\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}(\mathbf{r})^{\mathrm{T}}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{M}}_{\boxplus\,\mathcal{M}}$	$=$	$\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}$	$=$	$\left[\mathbf{R}(\mathbf{r})\quad \widehat{\mathbf{T}}(\mathbf{s},\mathbf{r})\,\mathbf{R}(\mathbf{r});\quad \mathbf{0}_{3\mathrm{x}3}\quad \mathbf{R}(\mathbf{r})\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\oplus\,\mathcal{M}\oplus\boldsymbol{\tau}}_{+\,\boldsymbol{\tau}}$	$=$	$\mathrm{J}_{\mathrm{R}}(\boldsymbol{\tau})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{\boxplus\,\boldsymbol{\tau}\boxplus\mathcal{M}}_{+\,\boldsymbol{\tau}}$	$=$	$\mathrm{J}_{\mathrm{L}}(\boldsymbol{\tau})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathcal{N}\ominus\mathcal{M}}_{\oplus\,\mathcal{M}}$	$=$	$-\mathrm{J}_{\mathrm{L}}^{-1}(\mathcal{N}\ominus\mathcal{M})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathcal{N}\boxminus\mathcal{M}}_{\boxplus\,\mathcal{M}}$	$=$	$-\mathrm{J}_{\mathrm{R}}^{-1}(\mathcal{N}\boxminus\mathcal{M})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathcal{N}\ominus\mathcal{R}}_{\oplus\,\mathcal{N}}$	$=$	$\mathrm{J}_{\mathrm{R}}^{-1}(\mathcal{N}\ominus\mathcal{M})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathcal{N}\boxminus\mathcal{R}}_{\boxplus\,\mathcal{N}}$	$=$	$\mathrm{J}_{\mathrm{L}}^{-1}(\mathcal{N}\boxminus\mathcal{M})$			$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}(\mathbf{p})}_{\oplus\,\mathcal{M}}$			$=$	$[\mathbf{R}\quad -\mathbf{R}\,\widehat{\mathbf{p}}]$	$\in \mathbb{R}^{3\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}(\mathbf{p})}_{\boxplus\,\mathcal{M}}$			$=$	$\left[\mathbf{I}_{3\mathrm{x}3}\quad -(\mathbf{R}\,\mathbf{p})^{\wedge} - \widehat{\mathbf{T}}\right]$	$\in \mathbb{R}^{3\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}(\mathbf{p})}_{+\,\mathbf{p}}$			$=$	$\mathbf{R}$	$\in \mathbb{R}^{3\mathrm{x}3}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}^{-1}(\mathbf{p})}_{\oplus\,\mathcal{M}}$			$=$	$\left[-\mathbf{I}_{3\mathrm{x}3}\quad \big(\mathbf{R}^{\mathrm{T}}(\mathbf{p}-\mathbf{T})\big)^{\wedge}\right]$	$\in \mathbb{R}^{3\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}^{-1}(\mathbf{p})}_{\boxplus\,\mathcal{M}}$			$=$	$\left[-\mathbf{R}^{\mathrm{T}}\quad \mathbf{R}^{\mathrm{T}}\,\widehat{\mathbf{p}}\right]$	$\in \mathbb{R}^{3\mathrm{x}6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathcal{M}}^{-1}(\mathbf{p})}_{+\,\mathbf{p}}$			$=$	$\mathbf{R}^{\mathrm{T}}$	$\in \mathbb{R}^{3\mathrm{x}3}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{M}}(\boldsymbol{\xi})}_{\oplus\,\mathcal{M}}$	$=$			$\left[-(\mathbf{R}\,\boldsymbol{\omega})^{\wedge}\,\mathbf{R}\quad -\mathbf{R}\,\widehat{\boldsymbol{\nu}} - \widehat{\mathbf{T}}\,\mathbf{R}\,\widehat{\boldsymbol{\omega}};\quad \mathbf{0}_{3\mathrm{x}3}\quad -\mathbf{R}\,\widehat{\boldsymbol{\omega}}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{M}}(\boldsymbol{\xi})}_{\boxplus\,\mathcal{M}}$	$=$			$\left[-(\mathbf{R}\,\boldsymbol{\omega})^{\wedge}\quad -(\mathbf{R}\,\boldsymbol{\nu})^{\wedge} - \widehat{\mathbf{T}}\,(\mathbf{R}\,\boldsymbol{\omega})^{\wedge} + (\mathbf{R}\,\boldsymbol{\omega})^{\wedge}\,\widehat{\mathbf{T}};\quad \mathbf{0}_{3\mathrm{x}3}\quad -(\mathbf{R}\,\boldsymbol{\omega})^{\wedge}\right]$	$\in \mathbb{R}^{6\mathrm{x}6}$

Jacobian		Table 1		Expression	Size
$\mathbf{J}^{+\,\mathbf{Ad}_{\mathcal{M}}(\boldsymbol{\xi})}_{+\,\boldsymbol{\xi}}$	$=$	$\mathbf{Ad}_{\mathcal{M}}$	$=$	$\Big[\mathbf{R} \;\; \widehat{\mathbf{T}}\,\mathbf{R};\;\; \mathbf{0}_{3x3} \;\; \mathbf{R}\Big]$	$\in \mathbb{R}^{6x6}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{M}}(\boldsymbol{\xi})}_{\oplus\,\mathcal{M}}$	$=$			$\Big[\mathbf{R}^{\mathrm{T}}\widehat{\boldsymbol{\omega}}\,\mathbf{R} \;\; \Big(\mathbf{R}^{\mathrm{T}}\boldsymbol{\nu}\Big)^{\wedge} - \Big(\mathbf{R}^{\mathrm{T}}\widehat{\mathbf{T}}\,\boldsymbol{\omega}\Big)^{\wedge};\;\; \mathbf{0}_{3x3} \;\; \Big(\mathbf{R}^{\mathrm{T}}\boldsymbol{\omega}\Big)^{\wedge}\Big]$	$\in \mathbb{R}^{6x6}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{M}}(\boldsymbol{\xi})}_{\boxplus\,\mathcal{M}}$	$=$			$\Big[\mathbf{R}^{\mathrm{T}}\widehat{\boldsymbol{\omega}} \;\; \mathbf{R}^{\mathrm{T}}\widehat{\boldsymbol{\nu}} - \mathbf{R}^{\mathrm{T}}\widehat{\mathbf{T}}\,\widehat{\boldsymbol{\omega}};\;\; \mathbf{0}_{3x3} \;\; \mathbf{R}^{\mathrm{T}}\widehat{\boldsymbol{\omega}}\Big]$	$\in \mathbb{R}^{6x6}$
$\mathbf{J}^{+\,\mathbf{Ad}^{-1}_{\mathcal{M}}(\boldsymbol{\xi})}_{+\,\boldsymbol{\xi}}$	$=$	$\mathbf{Ad}^{-1}_{\mathcal{M}}$	$=$	$\Big[\mathbf{R}^{\mathrm{T}} \;\; -\mathbf{R}^{\mathrm{T}}\,\widehat{\mathbf{T}};\;\; \mathbf{0}_{3x3} \;\; \mathbf{R}^{\mathrm{T}}\Big]$	$\in \mathbb{R}^{6x6}$
$\mathbf{J}^{+\,\mathrm{g}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{p})}_{+\,\boldsymbol{\tau}}$			$=$	$\Big[\mathbf{J}_{\mathrm{L}}\,(\mathbf{r}) \;\; \mathbf{Q}\,(\boldsymbol{\tau}) - [\mathbf{R}\,\mathbf{p}]^{\wedge}\,\mathbf{J}_{\mathrm{L}}\,(\mathbf{r}) - \widehat{\mathbf{T}}\,\mathbf{J}_{\mathrm{L}}\,(\mathbf{r})\Big]$	$\in \mathbb{R}^{3x6}$
$\mathbf{J}^{+\,\mathrm{g}^{-1}_{\mathrm{Exp}(\boldsymbol{\tau})}(\mathbf{p})}_{+\,\boldsymbol{\tau}}$			$=$	$\Big[-\mathbf{J}_{\mathrm{L}}\,(-\mathbf{r}) \;\; -\mathbf{Q}\,(-\boldsymbol{\tau}) - \Big[\mathbf{R}^{\mathrm{T}}\;(\mathbf{T}-\mathbf{p})\Big]^{\wedge}\,\mathbf{J}_{\mathrm{L}}\,(-\mathbf{r})\Big]$	$\in \mathbb{R}^{3x6}$

Table 16: Rigid body motion jacobians

6.14 Rigid Body Motion Discrete Integration

The discrete integration with time of an element of a Lie group based on its Lie algebra is discussed in detail in section 4.5, which includes expressions for the Euler, Heun and Runge-Kutta methods. In the case of rigid body motion, the state vector includes the motion element $\mathcal{M} \in \mathbb{SE}(3)$ and its twist $\boldsymbol{\xi} \in \mathbb{R}^6$ contained in the tangent space, viewed either in the local ($\boldsymbol{\xi}^{\mathrm{B}}_{\mathrm{EB}}$) or global ($\boldsymbol{\xi}^{\mathrm{E}}_{\mathrm{EB}}$) frames. The Euler method expressions equivalent to (168) and (170) are shown below. Expressions for other integration schemes can easily be derived from those in section 4.5:

$$\mathcal{M}_{\mathrm{k+1}} \approx \mathcal{M}_{\mathrm{k}} \oplus [\Delta \mathrm{t}\; \boldsymbol{\xi}^{\mathrm{B}}_{\mathrm{EBk}}] = \mathcal{M}_{\mathrm{k}} \circ \mathrm{Exp}\,(\Delta \mathrm{t}\; \boldsymbol{\xi}^{\mathrm{B}}_{\mathrm{EBk}}) \tag{442}$$

$$\mathcal{M}_{\mathrm{k+1}} \approx [\Delta \mathrm{t}\; \boldsymbol{\xi}^{\mathrm{E}}_{\mathrm{EBk}}] \boxplus \mathcal{M}_{\mathrm{k}} = \mathrm{Exp}\,(\Delta \mathrm{t}\; \boldsymbol{\xi}^{\mathrm{E}}_{\mathrm{EBk}}) \circ \mathcal{M}_{\mathrm{k}} \tag{443}$$

6.15 Rigid Body Motion Gauss-Newton Optimization

The minimization by means of the Gauss-Newton iterative method of the Euclidean norm of a non linear function whose input is a Lie group element is presented in section 4.6. In the case of rigid body motion, the resulting expressions for perturbations $\Delta\boldsymbol{\tau}^{\mathrm{E}}_{\mathrm{EB}} \in \mathfrak{se}(3)$ to an input motion $\mathcal{M} \in \mathbb{SE}(3)$ viewed in the global frame $\mathrm{F_E}$ are shown in (444) and (445), which are equivalent to the generic (189) and (192). Refer to section 4.6 for the meaning of the function jacobian $\mathbf{J}$ and to section 6.13 for that of the left jacobian $\mathbf{J}_{\mathrm{L}}$.

$$\mathcal{M}_{\mathrm{k+1}} \longleftarrow \Delta\boldsymbol{\tau}^{\mathrm{E}}_{\mathrm{EBk}} \boxplus \mathcal{M}_{\mathrm{k}} = \Delta\boldsymbol{\tau}^{\mathrm{E}}_{\mathrm{EBk}} \circ \mathrm{Exp}\,(\boldsymbol{\tau}_{\mathrm{EBk}}) \tag{444}$$

$$\Delta\boldsymbol{\tau}^{\mathrm{E}}_{\mathrm{EBk}} = -\Big[\mathbf{J}^{-\mathrm{T}}_{\mathrm{Lk}}\,\mathbf{J}^{\mathrm{T}}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{k}}\,\mathbf{J}^{-1}_{\mathrm{Lk}}\Big]^{-1}\,\mathbf{J}^{-\mathrm{T}}_{\mathrm{Lk}}\,\mathbf{J}^{\mathrm{T}}_{\mathrm{k}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}} \tag{445}$$

If the perturbation is viewed in the local frame $\mathrm{F_B}$, (193) and (194) are customized as follows, making use of the right jacobian $\mathbf{J}_{\mathrm{R}}$ defined in section 6.13:

$$\mathcal{M}_{\mathrm{k+1}} \longleftarrow \mathcal{M}_{\mathrm{k}} \oplus \Delta\boldsymbol{\tau}^{\mathrm{B}}_{\mathrm{EBk}} = \mathrm{Exp}\,(\boldsymbol{\tau}_{\mathrm{EBk}}) \circ \Delta\boldsymbol{\tau}^{\mathrm{B}}_{\mathrm{EBk}} \tag{446}$$

$$\Delta\boldsymbol{\tau}^{\mathrm{B}}_{\mathrm{EBk}} = -\Big[\mathbf{J}^{-\mathrm{T}}_{\mathrm{Rk}}\,\mathbf{J}^{\mathrm{T}}_{\mathrm{k}}\,\mathbf{J}_{\mathrm{k}}\,\mathbf{J}^{-1}_{\mathrm{Rk}}\Big]^{-1}\,\mathbf{J}^{-\mathrm{T}}_{\mathrm{Rk}}\,\mathbf{J}^{\mathrm{T}}_{\mathrm{k}}\,\boldsymbol{\mathcal{E}}_{\mathrm{k}} \tag{447}$$

6.16 Rigid Body Motion State Estimation

The adaptation of the `EKF` state estimation introduced in section 3.3 to the case in which Lie group elements and their velocities are present is discussed in detail in section 4.7. For rigid body motion with local perturbations, it is necessary to replace $\mathcal{X} \in \mathcal{G}$ by $\mathcal{M} \in \mathbb{SE}(3)$, $\Delta\boldsymbol{\tau}^{\mathcal{X}} \in \mathrm{T}_{\mathcal{X}}\mathcal{G}$ by $\Delta\boldsymbol{\tau}^{\mathrm{B}} \in \mathfrak{se}(3)$, $\mathbf{v}^{\mathcal{X}} \in \mathbb{R}^{\mathrm{m}}$ by $\boldsymbol{\xi}^{\mathrm{B}} \in \mathbb{R}^6$, $\mathbf{C}^{\mathcal{X}}_{\mathcal{X}\mathcal{X}} \in \mathbb{R}^{\mathrm{mxm}}$ by $\mathbf{C}^{\mathrm{B}}_{\mathcal{M}\mathcal{M}} \in \mathbb{R}^{6x6}$, and $\mathbf{J}^{\oplus\,\mathcal{X}\,\oplus\,\boldsymbol{\tau}}_{\oplus\,\mathcal{X}}$ by $\mathbf{J}^{\oplus\,\mathcal{M}\,\oplus\,\boldsymbol{\tau}}_{\oplus\,\mathcal{M}}$. The particularizations for global perturbations are similar.

6.17 Applications of the Various Motion Representations

This section discusses six different representations of the rigid body motion or special euclidean group $\mathbb{SE}(3)$: the affine representation, the homogeneous matrix, the transform vector, the unit dual quaternion, the half transform vector, and the screw. Although in theory all of them can be employed for each of the purposes described in this article, and the required expressions derived, each representation has its own advantages and disadvantages, being suited for certain purposes but not recommended for others.

- The affine representation $(\mathcal{R}, \mathbf{T})$ based on either the rotation matrix or the unit quaternion is the most natural representation for rigid body motion. It can be employed to track the motion over its manifold, although other options are preferred. Many difficulties arise from its complex nature as a composition between an $\mathbb{SO}(3)$ rotation and a $\mathbb{R}^3$ vector, such as the complex inverse and concatenation, the different nature of the point and vector actions, the lack of simple plus and minus operators to deal with perturbations, and the need to continuously keep track of both components when moving over the manifold.
- The homogeneous matrix $\mathbf{M}$ is a generalization of the rotation matrix for the case of rigid body motion that not only linearizes the transformation of coordinates at the expense of bigger size, but also adopts matrix algebra for the inversion and concatenation of transformations. A second advantage is that the transformations of vectors and points share the same map when employing homogeneous coordinates. Additionally, it provides a clear connection with the tangent space, together with the exponential and logarithmic maps, and plus and minus operators, which are not complex. Its main inconvenients are the huge size (16), the expense of maintaining the internal rotation matrix orthogonal if allowed to deviate from the manifold, and the need to work with homogeneous coordinates for both points and vectors. Its high cost precludes its use to track the motion over its manifold, although most implementations continuously compute its components (the rotation matrix and the translation vector) if the adjoint matrix or the jacobian blocks are required.
- The unit dual quaternion $\boldsymbol{\zeta}$ is the preferred representation to track the motion over its manifold, even if it is necessary to obtain the rotation matrix and the translation vector for the adjoint and jacobian blocks. It possesses a significant size advantage (8) with respect to the homogeneous matrix, although it is not cheap to recover its structure if allowed to deviate from the manifold. Unit dual quaternions are the least natural of the rigid body motion representations, being necessary to convert to a different $\mathbb{SE}(3)$ representation for visualization. While the inverse and concatenation are simple and linear, the motion actions for points and vectors are bilinear and require slightly different expressions when inverting them, which presents a disadvantage with the homogeneous matrix. A significant advantage is given by its direct relationship with the screw and associated `ScLERP` capabilities. Unit dual quaternion expressions are significantly more complex than those of the homogeneous matrix, and present a slightly less obvious connection with the tangent space.
- The main advantage of the transform vector $\boldsymbol{\tau}$ is that it belongs to the $\mathfrak{se}(3)$ tangent space while simultaneously being an $\mathbb{SE}(3)$ representation. It is hence indicated for those uses related with incremental motion changes by means of the exponential map together with the plus and minus operators (periodically adding the perturbations to the unit dual quaternion tracking the motion), such as discrete integration, optimization, and state estimation. The norms of its angular and linear components (ϕ, ρ) are the most adequate metrics for evaluating the distance (or estimation error) between two rigid bodies. Although it benefits from its straightforward inverse when used as a perturbation, its geometric appeal, and its small dimension (6), its usage for other applications is discouraged by its complex non linear kinematics, coordinate transformation, and composition, which are not shown in this article.
- The half transform vector $\boldsymbol{\Psi}$ is so similar (half) to the transform vector that its usage is not recommended in order to avoid confusion. Its only real application as the tangent space of the unit dual quaternion is in practice solved by dividing the transform vector by two when necessary.

- The screw $\mathbf{S}$, in addition to simultaneously belonging to the $\mathfrak{se}(3)$ tangent space and the $\mathbb{SE}(3)$ manifold, has the advantage that it clearly separates the influence of the motion direction from that of its magnitude, and as such it enables the definition of powers and `ScLERP`, which can not be obtained with any of the other representations. The dimension is not big (8) and the inversion is straightforward, but all other possible expressions, including motion and concatenation, are very complex and not shown in this article.

References

[1] J. L. Blanco, "A Tutorial on SE(3) Transformation Parameterizations and On-Manifold Optimization," 2020. `arXiv:2103.15980v2 [cs.RO]`, `https://doi.org/10.48550/arXiv.2103.15980`.

[2] J. Sola, J. Deray, and D. Atchuthan, "A Micro Lie Theory for State Estimation in Robotics," 2018. `arXiv:1812.01537v9 [cs.RO]`, `https://doi.org/10.48550/arXiv.1812.01537`.

[3] O. C. Ibe, *Fundamentals of Applied Probability and Random Processes.* Elsevier, 2nd ed., 2005. ISBN 0-080-49270-3.

[4] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 4th ed., 2002. ISBN 0-07-122661-3.

[5] J. A. Farrell, *Aided Navigation, GPS with High Rate Sensors.* McGraw-Hill, Electronic Engineering Series, 2008. ISBN 0-071-49329-8.

[6] D. Simon, *Optimal State Estimation.* John Wiley & Sons, 2006. ISBN 0-471-70858-5.

[7] Brian D. Ripley, *Stochastic Simulation.* John Wiley & Sons, 1987. ISBN 9780471818847.

[8] Shiomo S. Sawilowsky, "You Think You've Got Trivials?," *Journal of Modern Applied Statistical Methods*, 2003. `https://doi.org/10.22237/jmasm/1051748460`.

[9] M. H. Shojaeefard, A. Khalkhali, and S. Yarmohammadistri, "An Efficient Sensitivity Analysis Method for Modified Geometry of Macpherson Suspension based on Pearson Correlation Coefficient," *Vehicle System Dynamics*, 2017. `https://doi.org/10.1080/00423114.2017.1283046`.

[10] F. Frishman, "On the Arithmetic Means and Variances of Products and Ratios of Random Variables," *Army Research Office*, 1971. `https://doi.org/10.1007/978-94-010-1842-5_32`.

[11] P. G. Hoel, S. C. Port, and C. J. Stone, *Introduction to Stochastic Processes.* Waveland Press Inc., 1972. ISBN 1-478-60899-4.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++, The Art of Scientific Computing.* Cambridge University Press, 2nd ed., 2002. ISBN 0-521-75033-4.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2nd ed., 2003. ISBN 0-521-54051-8.

[14] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A Unifying Framework," *International Journal of Computer Vision*, 2004. `https://doi.org/10.1023/B:VISI.0000011205.11775.fd`.

[15] C. C. Pinter, *A Book of Abstract Algebra.* Dover Publications, 2nd ed., 1990. ISBN 0-486-47417-8.

[16] M. A. Armstrong, *Basic Topology.* Springer, 1983. ISBN 978-0-387-90839-7.

[17] M. D. Shuster, "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, 1993. DOI not available.

[18] S. Roman, *Advanced Linear Algebra.* Springer, 2nd ed., 2005. ISBN 0-387-24766-1.

[19] V. Bryant, *Metric Spaces: Iteration and Application.* Cambridge University Press, 1985. ISBN 0-521-31897-1.

[20] P. K. Jain, K. Ahmad, and O. P. Ahuja, *Functional Analysis.* New Age International, 1995. ISBN 8-122-40801-X.

[21] E. Artin, *Geometric Algebra.* John Wiley & Sons, 1988 ed. ISBN 0-471-60839-4.

[22] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision, From Images to Geometric Models.* Springer, 2001. ISBN 978-1-4419-1846-8.

[23] G. Strang, *Linear Algebra and its Applications.* Thomson Brooks/Cole, 4^{th} ed., 2006. ISBN 0-03-010567-6.

[24] Y. B. Jia, "Dual Quaternions," 2013. DOI not available.

[25] T. W. Gamelin and R. E. Greene, *Introduction to Topology.* Courier Corporation, 1999. ISBN 0-486-40680-6.

[26] J. Sola, "Quaternion Kinematics for the Error-State Kalman Filter," 2017. `arXiv:1711.02508v1 [cs.RO]`, `https://doi.org/10.48550/arXiv.1711.02508`.

[27] Richard M. Murray and Zexiang li and S. Shankar Sastry, *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1994. ISBN 0-849-37981-4.

[28] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups*, vol. 2. Birkhauser, 2012. ISBN 978-0-8176-4943-2.

[29] I. Y. Baritzhack and K. A. Fegley, "Orthogonalization Techniques of a Direction Cosine Matrix," *IEEE Transactions of Aerospace and Electronic Systems*, 1969. `https://doi.org/10.1109/TAES.1969.309878`.

[30] B. Kenwright, "Dual Quaternions: From Classical Mechanics to Computer Graphics and Beyond," 2012. DOI not available.

[31] A. Valverde and P. Tsiotras, "Dual Quaternion Framework for Modeling of Spacecraft Mounted Multibody Robotic Systems," *Frontiers in Robotics and AI*, 2018. `https://doi.org/10.3389/frobt.2018.00128`.

[32] K. Daniilidis, "Hand-Eye Calibration Using Dual Quaternions," 1999. `https://doi.org/10.1177/02783649922066213`.

[33] B. Busam, T. Birdal, and N. Navab, "Camera Pose Filtering with Local Regression Geodesics on the Riemannian Manifold of Dual Quaternions," 2017. `arXiv:1704.07072v4 [cs.CV]`, `https://doi.org/10.48550/arXiv.1704.07072`.

[34] T. D. Barfoot and P. T. Furgale, "Associating Uncertainty With Three-Dimensional Poses for Use in Estimation Problems," *IEEE Transactions in Robotics*, 2014. `https://doi.org/10.1109/TRO.2014.2298059`.