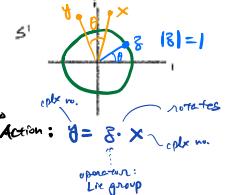


Lie Theory for Robotists (Intuition)

S^1 : The unit complex numbers (an example of lie group)



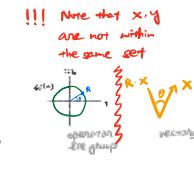
$$\text{constraint: } \bar{z} \cdot z = 1$$

topology: unit circle S^1

elements: $z = \cos \theta + i \sin \theta$

inverse: \bar{z} (conjugate)

composition: $z_1 \cdot z_2$



$$\text{constraint: } R^T R = I$$

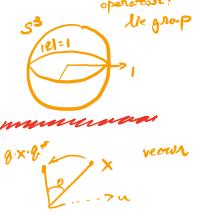
topology: "circle" $SO(2)$

elements: $R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$

inverse: R^T

composition: $R_1 \cdot R_2$

S^3 : unit quaternion in \mathbb{R}^4



Pose of a robot in the plane: $SE(2)$

$$X(t) = \begin{bmatrix} R(t) & P(t) \\ 0 & I \end{bmatrix}$$

act on each other
Lie
Transform objects
invert
compose
interpolate

integrate
differentiate
uncertainty

Lie group definition

What is a group? 4 group axioms

Group: set G of elements $\{x, y, z, \dots\}$

with an operation " \cdot " s.t.:

- composition: $X \cdot Y$ is in G
- Identity: $E \cdot X = X \cdot E = X$ is in G
- Inverse: $X^{-1} \cdot X = X \cdot X^{-1} = E$ is in G
- operation is associative: $X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$

!!! non-commutative $X \cdot Y \neq Y \cdot X$

What is a lie group?

- A group that is also a smooth manifold



- A lie group satisfies (4 axioms + smooth manifold)

Group Action

- A Group can act another set V to transform its element

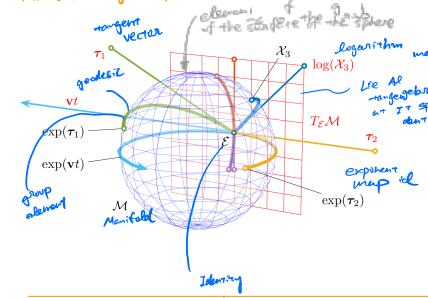
- Given X, Y, v in G action " \cdot " is such that:

- v in V
- Identity is the Null action: $E \cdot v = v$
- compatible w/ composition: $(X \cdot Y) \cdot v = X \cdot (Y \cdot v)$

- Continuous Transformation Groups

Topology of Lie Theory

Manifold, tangent space & exponential map



tangent space & lie algebra

- $T_x M$ on M is a tangent space
- is a vector space
- can do calculus
- dimension of tangent space no. of D.F. of the M
- Tangent space $\otimes E$ is Lie Algebra

e.g.:

$$S^1: \dot{\theta} = 1$$

$$\dot{\theta} \cdot \dot{\theta} = 0$$

$$\dot{\theta} \cdot \dot{\theta} = -(\dot{\theta})^2$$

Lie Algebra

$$W = i\dot{\theta} \in i\mathbb{R}$$

• Lie Algebra

$$W \cdot W = i\dot{\theta} \cdot i\dot{\theta} = -\dot{\theta}^2$$

• Isomorphism

$$H: W \mapsto i\dot{\theta}$$

• Vect: $W = W^a \partial_a$

$$W = W^a \partial_a$$

e.g.:

$$SO(3): \dot{\theta}^1 = \omega_1 \in \mathbb{R}$$

$$\dot{\theta}^2 = \omega_2 \in \mathbb{R}$$

$$\dot{\theta}^3 = \omega_3 \in \mathbb{R}$$

• Lie algebra

$$W = i\omega_1 \hat{x} + i\omega_2 \hat{y} + i\omega_3 \hat{z} \in \mathbb{R}^3$$

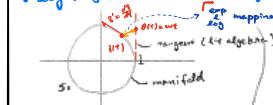
• Lie algebra when $R = I$

$$R = Wx \equiv \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \in SO(3)$$

exponential mapping

• exp: from tangent to manifold (wrap on the geodesic)

• log: from manifold to tangent (unwrap)



• exponential mapping

$$- \dot{\theta} \cdot \dot{\theta} = i\omega \rightarrow \dot{\theta}' = \dot{\theta} \cdot i\omega \text{ (ODE)}$$

$$\dot{\theta}(t) = \dot{\theta}_0 \exp(i\omega t)$$

$$- \text{if } \dot{\theta}_0 = \dot{\theta}(0) = I \text{ and}$$

$$i\omega t = i\theta$$

$$\Rightarrow \dot{\theta}(t) = \exp(i\theta)$$

- Taylor expansion

$$I^2 = I \quad I^3 = -I$$

$$\exp(I\theta) = I + I\theta + (I\theta)^2/2! + \dots$$

$$= I + I\theta - \theta^2/2! - I\theta^3/3! + I\theta^5/5! + \dots$$

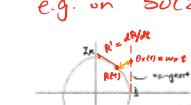
$$= (I - \theta^2/2! + \dots) + I\theta \cdot (I - \theta^2/2! + \dots)$$

$$= \cos \theta + i \sin \theta$$

- exp map: $\dot{\theta} = \exp(i\theta)$

log map: $i\theta = \log(\dot{\theta})$

e.g. on $SO(2)$



• exponential mapping

$$- \dot{\theta} = R \cdot W_x \quad W_x = \begin{bmatrix} 0 & -w \\ w & 0 \end{bmatrix}$$

$$\Rightarrow R(t) = R \exp(W_x t)$$

$$- R(0) = R = I$$

$$\Rightarrow R(t) = \exp(W_x t)$$

$$\text{let } W_x = \theta x = \exp(i\theta x)$$

$$= \exp(i\theta x)$$

- similarly, do Taylor Expansion

$$I^2 = I \quad I^3 = -I$$

$$\Rightarrow \exp(W_x) = I + W_x + (W_x)^2/2! + (W_x)^3/3! + \dots$$

$$= I(1 - \theta^2/2! + \dots) + Ix \cdot (1 - \theta^2/2! + \dots)$$

$$= I \cos \theta + Ix \sin \theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

- exp map: $R(t) = \exp(W_x t)$

log map: $\theta x = \log(R)$

- Lie algebra $\mathfrak{so}(2)$ $W_x = \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix} \leftrightarrow \theta \in \mathbb{R}$

isomorphism

$$R = \exp(W_x t) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ expressed in } \mathbb{R}^2, \mathbb{R}^3$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & -\theta t \\ \theta t & 1 \end{pmatrix} \text{ in Lie algebra}$$

<

Lie Theory (Mathematics)

- △ Motivation to exploit Lie Theory in robotics application
- we're dealing with non-Euclidean components
 - when doing operations (estimation/optimization) on those non-Euclidean vectors, HARD to impose constraints
 - e.g. when adapting SE(3) as a variable, $SE(3) \rightarrow$ orthogonal, and determinant = 1 hard to impose this during optimization

△ Algebra structure: combination of + set operation

FACT A set "A" has group structure under operation $f: A \times A \rightarrow A$

& properties & axioms: $\forall a, b, c \in A$

1. Closure: $a * b \in A$
2. Associativity: $(a * b) * c = a * (b * c)$
3. Identity: $\exists e \in A | e * a = a * e = a$
4. Inverse: $\exists f \in A | f * a = a * f = e$

group \rightarrow abelian group
ring (two operations) \rightarrow abelian ring
field (+, *) operation



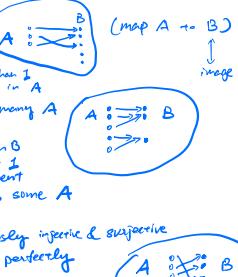
FACT map or morphism

a rule: \rightarrow each element in a set A
 \downarrow assigns unique element
in a set B

$$\{f: A \rightarrow B | f(a) = b \in B, \forall a \in A\}$$

LEMMA

- Injective:
• each element in A
• has no more than 1 mapped element in B
- Surjective:
• each element in B
• has at least 1 mapped element
• every B has some A
- Bijective:
• simultaneously injective & surjective
• $A \leftrightarrow B$, perfectly



FACT homomorphism

a map/morphism that preserve the group is called homomorphism

FACT isomorphism

bijection homomorphism is called isomorphism

FACT group action

$$\{g: G \times V \rightarrow V | g_1(u) = v \in V \text{ } \forall u \in V \text{ } \forall v \in V\}$$

△ Lie group & Lie algebra

FACT manifold

• a manifold is a topological space

• at local point, resembles Euclidean Space

• embedded in a higher-dimension space:

- curved
- smooth (spike & edge-less)
- defined by constraints

FACT Lie group

$\triangleleft G, \circ \triangleright$ a smooth manifold under group actions

• G has group multiplication operation

• example of manifolds: S^1, S^2, \dots, S^n

• $O(n)$: $\{x \in \mathbb{R}^n | x_i \geq 0, \sum x_i^2 = 1\}$

• $SO(n)$: $\{x \in O(n) | \det(x) = 1\}$

• $SE(3)$: $\{x \in \mathbb{R}^7 | x_7 = 1\}$

• $SE(2)$: $\{x \in \mathbb{R}^3 | x_3 = 1\}$

• $SO(3)$: $\{x \in \mathbb{R}^3 | x_3 = 1\}$

• $SO(2)$: $\{x \in \mathbb{R}^2 | x_2 = 1\}$

• $U(n)$: $\{x \in \mathbb{C}^{n \times n} | x^H x = I_n\}$

• $U(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1\}$

• $Sp(2n)$: $\{x \in \mathbb{C}^{2n \times 2n} | x^H x = I_{2n}, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(10)$: $\{x \in \mathbb{C}^{10 \times 10} | x^H x = I_{10}, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(11)$: $\{x \in \mathbb{C}^{11 \times 11} | x^H x = I_{11}, x_1^H x_2 = 0\}$

• $Sp(10)$: $\{x \in \mathbb{C}^{10 \times 10} | x^H x = I_{10}, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(12)$: $\{x \in \mathbb{C}^{12 \times 12} | x^H x = I_{12}, x_1^H x_2 = 0\}$

• $Sp(11)$: $\{x \in \mathbb{C}^{11 \times 11} | x^H x = I_{11}, x_1^H x_2 = 0\}$

• $Sp(10)$: $\{x \in \mathbb{C}^{10 \times 10} | x^H x = I_{10}, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(13)$: $\{x \in \mathbb{C}^{13 \times 13} | x^H x = I_{13}, x_1^H x_2 = 0\}$

• $Sp(12)$: $\{x \in \mathbb{C}^{12 \times 12} | x^H x = I_{12}, x_1^H x_2 = 0\}$

• $Sp(11)$: $\{x \in \mathbb{C}^{11 \times 11} | x^H x = I_{11}, x_1^H x_2 = 0\}$

• $Sp(10)$: $\{x \in \mathbb{C}^{10 \times 10} | x^H x = I_{10}, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} | x^H x = I_6, x_1^H x_2 = 0\}$

• $Sp(5)$: $\{x \in \mathbb{C}^{5 \times 5} | x^H x = I_5, x_1^H x_2 = 0\}$

• $Sp(4)$: $\{x \in \mathbb{C}^{4 \times 4} | x^H x = I_4, x_1^H x_2 = 0\}$

• $Sp(3)$: $\{x \in \mathbb{C}^{3 \times 3} | x^H x = I_3, x_1^H x_2 = 0\}$

• $Sp(2)$: $\{x \in \mathbb{C}^{2 \times 2} | x^H x = I_2, x_1^H x_2 = 0\}$

• $Sp(1)$: $\{x \in \mathbb{C}^{1 \times 1} | |x| = 1, x_1^H x_2 = 0\}$

• $Sp(14)$: $\{x \in \mathbb{C}^{14 \times 14} | x^H x = I_{14}, x_1^H x_2 = 0\}$

• $Sp(13)$: $\{x \in \mathbb{C}^{13 \times 13} | x^H x = I_{13}, x_1^H x_2 = 0\}$

• $Sp(12)$: $\{x \in \mathbb{C}^{12 \times 12} | x^H x = I_{12}, x_1^H x_2 = 0\}$

• $Sp(11)$: $\{x \in \mathbb{C}^{11 \times 11} | x^H x = I_{11}, x_1^H x_2 = 0\}$

• $Sp(10)$: $\{x \in \mathbb{C}^{10 \times 10} | x^H x = I_{10}, x_1^H x_2 = 0\}$

• $Sp(9)$: $\{x \in \mathbb{C}^{9 \times 9} | x^H x = I_9, x_1^H x_2 = 0\}$

• $Sp(8)$: $\{x \in \mathbb{C}^{8 \times 8} | x^H x = I_8, x_1^H x_2 = 0\}$

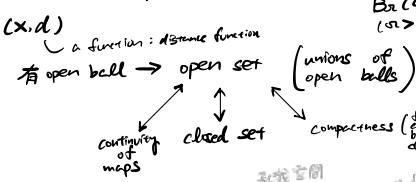
• $Sp(7)$: $\{x \in \mathbb{C}^{7 \times 7} | x^H x = I_7, x_1^H x_2 = 0\}$

• $Sp(6)$: $\{x \in \mathbb{C}^{6 \times 6} |$

What is Topology

- recall (metric spaces)

down open balls



FACT

A topological space $X = (X, \mathcal{T}_X)$

consists of - a set X (called the underlying space of X)
- a family \mathcal{T}_X of subsets of X (i.e. $\mathcal{T}_X \subseteq \mathcal{P}(X)$)

s.t. (1) $X \notin \mathcal{T}_X \in \mathcal{T}_X$ (open or?)

(2) $\bigcup_{U \in \mathcal{T}_X} U \in \mathcal{T}_X$ (closed)
if U is open ↓ closed should be open

(3) $U, V \in \mathcal{T}_X \Rightarrow U \cap V \in \mathcal{T}_X$

*** \mathcal{T}_X is called a topology on X**

Convention: we usually use X + indicate the set X & omit the subscript X in \mathcal{T}_X

topological structure by saying "a topological space (X, \mathcal{T}) "

e.g. metric spaces

$(X, d) \rightsquigarrow (X, \mathcal{T}_d)$

- different distance functions might determine the same topology

$$d(x, y) := \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}$$

ex. V : finite diml vector space over $K (= \mathbb{R} \text{ or } \mathbb{C})$

All norms on V determine the same topology on V

\mathcal{T}_{all}

Set X ,
 T is a topology on X
if:
1) $\emptyset, X \in T$
2) $A, B \in T \Rightarrow A \cap B \in T$
3) $A \in T \Rightarrow \bigcup_{x \in A} x \in T$
e.g. $X = \{a, b, c\}$
 $T = \{\emptyset, \{a\}, \{a, b\}\}$
1) $\emptyset \in T, X \in T$
2) $\{a\} \cap \{b\} = \emptyset \in T$
 $\{a\} \cap \{c\} = \emptyset \in T$
 $\{b\} \cap \{c\} = \emptyset \in T$
3) $\{a\} \in T \Rightarrow \{a\} \in T$
 $\{a\} \cup \{b\} = \{a, b\} \in T$
 $\{a\} \cup \{c\} = \{a, c\} \in T$
 $\{b\} \cup \{c\} = \{b, c\} \in T$
 $\{a, b\} \cup \{a, c\} = \{a, b, c\} \in T$
 $\therefore T$ is a topology on X
e.g. $X = \{a, b, c\}$
 $\{a, b\} \in T \in T$
2) \emptyset
3) X

A micro Lie theory for state estimation in robotics

Joan Solà, Jeremie Deray, Dinesh Atchuthan

Abstract—A Lie group is an old mathematical abstract object dating back to the XIX century, when mathematician Sophus Lie laid the foundations of the theory of continuous transformation groups. Its influence has spread over diverse areas of science and technology many years later. In robotics, we are recently experiencing an important trend in its usage, at least in the fields of estimation, and particularly in motion estimation for navigation. Yet for a vast majority of roboticians, Lie groups are highly abstract constructions and therefore difficult to understand and to use.

In estimation for robotics it is often not necessary to exploit the full capacity of the theory, and therefore an effort of selection of materials is required. In this paper, we will walk through the most basic principles of the Lie theory, with the aim of conveying clear and useful ideas, and leave a significant corpus of the Lie theory behind. Even with this mutilation, the material included here has proven to be extremely useful in modern estimation algorithms for robotics, especially in the fields of SLAM, visual odometry, and the like.

Alongside this micro Lie theory, we provide a chapter with a few application examples, and a vast reference of formulas for the major Lie groups used in robotics, including most Jacobian matrices and the way to easily manipulate them. We also present a new C++ template-only library implementing all the functionality described here.

I. INTRODUCTION

There has been a remarkable effort in the last years in the robotics community to formulate estimation problems properly. This is motivated by an increasing demand for precision, consistency and stability of the solutions. Indeed, proper modeling of the states and measurements, the functions relating them, and their uncertainties, is crucial to achieving these goals. This has led to designs involving what has been known as ‘manifolds’, which in this context are no less than the (smooth topologic surfaces) of the Lie groups where the state representations evolve. Relying on the Lie theory (LT) we are able to construct a rigorous calculus corpus to handle uncertainties, derivatives and integrals with precision and ease. Typically, these works have focused on the well-known manifolds of rotation $\text{SO}(3)$ and rigid motion $\text{SE}(3)$.

When being introduced to Lie groups for the first time, it is important to try to regard them from different points of view. The topological viewpoint, see Fig. 1, involves the shape of the manifold and conveys powerful intuitions of its relation to the tangent space and the exponential map. The algebraic viewpoint involves the group operations and their concrete realization, allowing the exploitation of algebraic properties to develop closed-form formulas or to simplify them. The geometrical viewpoint, particularly useful in robotics, associates group elements to the position, velocity, orientation,

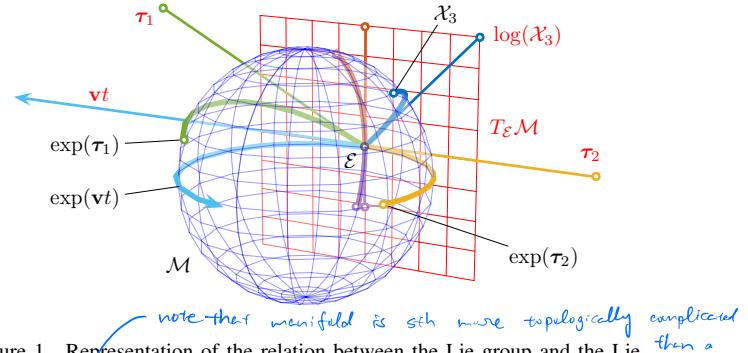


Figure 1. Representation of the relation between the Lie group and the Lie algebra. The Lie algebra $T_{\mathcal{E}} \mathcal{M}$ (red plane) is the tangent space to the Lie group’s manifold \mathcal{M} (here represented as a blue sphere) at the identity \mathcal{E} . Through the exponential map, each straight path v_t through the origin on the Lie algebra produces a path $\exp(vt)$ around the manifold which runs along the respective geodesic. Conversely, each element of the group has an equivalent in the Lie algebra. This relation is so profound that (nearly) all operations in the group, which is curved and nonlinear, have an exact equivalent in the Lie algebra, which is a linear vector space. Though the sphere in \mathbb{R}^3 is not a Lie group (we just use it as a representation that can be drawn on paper), that in \mathbb{R}^4 is, and describes the group of unit quaternions —see Fig. 4 and Ex. 5.

and/or other modifications of bodies or reference frames. The origin frame may be identified with the group’s identity, and any other point on the manifold represents a certain ‘local’ frame. By resorting to these analogies, many mathematical abstractions of the LT can be brought closer to intuitive notions in vector spaces, geometry, kinematics, and other more classical fields.

Lie theory is by no means simple. To grasp a minimum idea of what LT can be, we may consider the following three references. First, Abbaspour’s “*Basic Lie theory*” [1] comprises more than 400 pages. With a similar title, Howe’s “*Very basic Lie theory*” [2] comprises 24 (dense) pages, and is sometimes considered a must-read introduction. Finally, the more modern and often celebrated Stillwell’s “*Naive Lie theory*” [3] comprises more than 200 pages. With such precedents labeled as ‘basic’, ‘very basic’ and ‘naive’, the aim of this paper at merely 17 pages is to simplify Lie theory even more (thus our adjective ‘micro’ in the title). This we do in two ways. First, we select a small subset of material from the LT. This subset is so small that it merely explores the potential of LT. However, it appears very useful for uncertainty management in the kind of estimation problems we deal with in robotics (e.g. inertial pre-integration, odometry and SLAM, visual servoing, and the like), thus enabling elegant and rigorous designs of optimal optimizers. Second, we explain it in a didactical way, with plenty of redundancy so as to

reduce the entry gap to LT even more, which we believe is still needed. That is, we insist on the efforts in this direction of, to name a paradigmatic title, Stillwell's [3], and provide yet a more simplified version. The main text body is generic, though we try to keep the abstraction level to a minimum. Inserted examples serve as a grounding base for the general concepts when applied to known groups (rotation and motion matrices, quaternions, etc.). Also, plenty of figures with very verbose captions re-explain the same concepts once again. We put special attention to the computation of Jacobians (a topic that is not treated in [3]), which are essential for most optimal estimators and the source of much trouble when designing new algorithms. We provide a chapter with some applicative examples for robot localization and mapping, implementing EKF and nonlinear optimization algorithms based on LT. And finally, several appendices contain ample reference for the most relevant details of the most commonly used groups in robotics: unit complex numbers, quaternions, 2D and 3D rotation matrices, 2D and 3D rigid motion matrices, and the trivial translation groups.

Yet our most important simplification to Lie theory is in terms of scope. The following passage from Howe [2] may serve us to illustrate what we leave behind: “*The essential phenomenon of Lie theory is that one may associate in a natural way to a Lie group \mathcal{G} its Lie algebra \mathfrak{g} . The Lie algebra \mathfrak{g} is first of all a vector space and secondly is endowed with a bilinear nonassociative product called the Lie bracket [...]. Amazingly, the group \mathcal{G} is almost completely determined by \mathfrak{g} and its Lie bracket. Thus for many purposes one can replace \mathcal{G} with \mathfrak{g} . Since \mathcal{G} is a complicated nonlinear object and \mathfrak{g} is just a vector space, it is usually vastly simpler to work with \mathfrak{g} . [...] This is one source of the power of Lie theory.*” In [3], Stillwell even speaks of “*the miracle of Lie theory*”. In this work, we will effectively relegate the Lie algebra to a second plane in favor of its equivalent vector space \mathbb{R}^n , and will not introduce the Lie bracket at all. Therefore, the connection between the Lie group and its Lie algebra will not be made here as profound as it should. Our position is that, given the target application areas that we foresee, this material is often not necessary. Moreover, if included, then we would fail in the objective of being clear and useful, because the reader would have to go into mathematical concepts that, by their abstraction or subtleness, are unnecessarily complicated.

Our effort is in line with other recent works on the subject [4], [5], [6], which have also identified this need of bringing the LT closer to the roboticist. Our approach aims at appearing familiar to the target audience of this paper: an audience that is skilled in state estimation (Kalman filtering, graph-based optimization, and the like), but not yet familiar with the theoretical corpus of the Lie theory. We have for this taken some initiatives concerning notation, especially in the definition of the derivative, bringing it close to the vectorial counterparts, thus making the chain rule clearly visible. As said, we opted to practically avoid the material proper to the Lie algebra, and prefer instead to work on its isomorphic tangent vector space \mathbb{R}^n , which is where we ultimately represent uncertainty or (small) state increments. All these steps are undertaken with absolutely no loss in precision or exactness,

remark: Jacobian can be directly calculated based on chain rule: yet tedious & error-prone

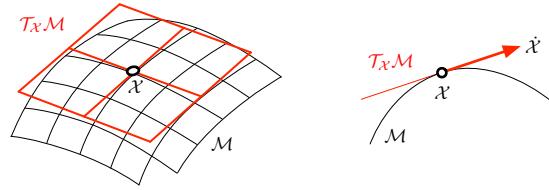


Figure 2. A manifold \mathcal{M} and the vector space $T_X \mathcal{M}$ (in this case $\cong \mathbb{R}^2$) tangent at the point X , and a convenient side-cut. The velocity element, $\dot{\mathcal{X}} = \partial \mathcal{X} / \partial t$, does not belong to the manifold \mathcal{M} but to the tangent space $T_X \mathcal{M}$.

and we believe they make the understanding of the LT and the manipulation of its tools easier.

This paper is accompanied by a new open-source C++ header-only library, called `manif` [7], which can be found at <https://github.com/artivis/manif>. `manif` implements the widely used groups $SO(2)$, $SO(3)$, $SE(2)$ and $SE(3)$, with support for the creation of analytic Jacobians. The library is designed for ease of use, flexibility, and performance.

II. A MICRO LIE THEORY

A. The Lie group

The Lie group encompasses the concepts of *group* and *smooth manifold* in a unique body: a Lie group \mathcal{G} is a smooth manifold whose elements satisfy the group axioms. We briefly present these two concepts before joining them together.

On one hand, a differentiable or *smooth manifold* is a topological space that locally resembles linear space. The reader should be able to visualize the idea of manifold (Fig. 2): it is like a curved, smooth (hyper)-surface, with no edges or spikes, embedded in a space of higher dimension. In robotics, we say that our state vector evolves on this surface, that is, the manifold describes or is defined by the constraints imposed on the state. For example, vectors with the unit norm constraint define a spherical manifold of radius one. The smoothness of the manifold implies the existence of a unique tangent space at each point. This space is a linear or vector space on which we are allowed to do calculus.

On the other hand, a *group* (\mathcal{G}, \circ) is a set, \mathcal{G} , with a composition operation, \circ , that, for elements $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{G}$, satisfies the following axioms,

$$\text{Closure under } \circ : \mathcal{X} \circ \mathcal{Y} \in \mathcal{G} \quad (1)$$

$$\text{Identity } \mathcal{E} : \mathcal{E} \circ \mathcal{X} = \mathcal{X} \circ \mathcal{E} = \mathcal{X} \quad (2)$$

$$\text{Inverse } \mathcal{X}^{-1} : \mathcal{X}^{-1} \circ \mathcal{X} = \mathcal{X} \circ \mathcal{X}^{-1} = \mathcal{E} \quad (3)$$

$$\text{Associativity} : (\mathcal{X} \circ \mathcal{Y}) \circ \mathcal{Z} = \mathcal{X} \circ (\mathcal{Y} \circ \mathcal{Z}) . \quad (4)$$

In a *Lie group*, the manifold looks the same at every point (like *e.g.* in the surface of a sphere, see Exs. 1 and 2), and therefore all tangent spaces at any point are alike. The group structure imposes that the composition of elements of the manifold remains on the manifold, (1), and that each element has an inverse also in the manifold, (3). A special one of these elements is the identity, (2), and thus a special one of the tangent spaces is the tangent at the identity, which we call the Lie algebra of the Lie group. Lie groups join the local properties of smooth manifolds, allowing us to do calculus,

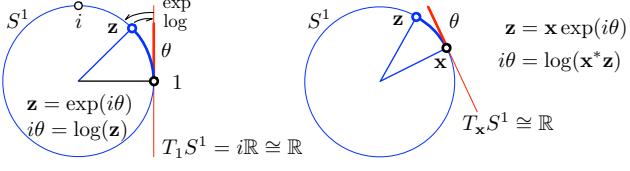


Figure 3. The S^1 manifold is a unit circle (blue) in the plane \mathbb{C} , where the unit complex numbers $z^* z = 1$ live. The Lie algebra $\mathfrak{s}^1 = T_{\mathcal{E}} S^1$ is the line of imaginary numbers $i\mathbb{R}$ (red), and any tangent space $T S^1$ is isomorphic to the line \mathbb{R} (red). Tangent vectors (red segment) wrap the manifold creating the arc of circle (blue arc). Mappings \exp and \log (arrows) map (wrap and unwrap) elements of $i\mathbb{R}$ to/from elements of S^1 (blue arc). Increments between unit complex numbers are expressed in the tangent space via composition and the exponential map (and we will define special operators \oplus, \ominus for this). See the text for explanations, and Fig. 4 for a similar group.

Example 1: The unit complex numbers group S^1

Our first example of Lie group, which is the easiest to visualize, is the group of unit complex numbers under complex multiplication (Fig. 3). Unit complex numbers take the form $z = \cos \theta + i \sin \theta$.

– *Action*: Vectors $x = x + iy$ rotate in the plane by an angle θ , through complex multiplication, $x' = zx$.

– *Group facts*: The product of unit complex numbers is a unit complex number, the identity is 1, and the inverse is the conjugate z^* .

– *Manifold facts*: The unit norm constraint defines the unit circle in the complex plane (which can be viewed as the 1-sphere, and hence the name S^1). This is a 1-DoF curve in 2-dimensional space. Unit complex numbers evolve with time on this circle. The group (the circle) resembles the linear space (the tangent line) locally, but not globally.

with the global properties of groups, enabling the nonlinear composition of distant objects.

B. The group actions

Importantly, Lie groups come with the power to transform elements of other sets, producing *e.g.* rotations, translations, scalings, and combinations of them. These are extensively used in robotics, both in 2D and 3D.

Given a Lie group \mathcal{M} and a set \mathcal{V} , we note $\mathcal{X} \cdot v$ the *action* of $\mathcal{X} \in \mathcal{M}$ on $v \in \mathcal{V}$,

$$\cdot : \mathcal{M} \times \mathcal{V} \rightarrow \mathcal{V} ; (\mathcal{X}, v) \mapsto \mathcal{X} \cdot v . \quad (5)$$

For \cdot to be a group action, it must satisfy the axioms,

$$\text{Identity} : \quad \mathcal{E} \cdot v = v \quad (6)$$

$$\text{Compatibility} : \quad (\mathcal{X} \circ \mathcal{Y}) \cdot v = \mathcal{X} \cdot (\mathcal{Y} \cdot v) . \quad (7)$$

Common examples are the groups of rotation matrices $SO(n)$, the group of unit quaternions, and the groups of rigid

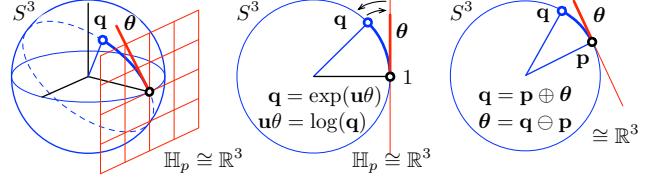


Figure 4. The S^3 manifold is a unit 3-sphere (blue) in the 4-space of quaternions \mathbb{H} , where the unit quaternions $q^* q = 1$ live. The Lie algebra is the space of pure imaginary quaternions $ix + jy + kz \in \mathbb{H}_p$, isomorphic to the hyperplane \mathbb{R}^3 (red grid), and any other tangent space $T S^3$ is also isomorphic to \mathbb{R}^3 . Tangent vectors (red segment) wrap the manifold over the great arc or *geodesic* (dashed). The centre and right figures show a side-cut through this geodesic (notice how it resembles S^1 in Fig. 3). Mappings \exp and \log (arrows) map (wrap and unwrap) elements of \mathbb{H}_p to/from elements of S^3 (blue arc). Increments between quaternions are expressed in the tangent space via the operators \oplus, \ominus (see text).

Example 2: The unit quaternions group S^3

A second example of Lie group, which is also relatively easy to visualize, is the group of unit quaternions under quaternion multiplication (Fig. 4). Unit quaternions take the form $q = \cos(\theta/2) + \mathbf{u} \sin(\theta/2)$, with $\mathbf{u} = iu_x + ju_y + ku_z$ a unitary axis and θ a rotation angle.

– *Action*: Vectors $x = ix + jy + kz$ rotate in 3D space by an angle θ around the unit axis \mathbf{u} through the double quaternion product $x' = q x q^*$.

– *Group facts*: The product of unit quaternions is a unit quaternion, the identity is 1, and the inverse is the conjugate q^* .

– *Manifold facts*: The unit norm constraint defines the 3-sphere S^3 , a spherical 3-dimensional surface or *manifold* in 4-dimensional space. Unit quaternions evolve with time on this surface. The group (the sphere) resembles the linear space (the tangent hyperplane $\mathbb{R}^3 \subset \mathbb{R}^4$) locally, but not globally.

motion $SE(n)$. Their respective actions on vectors satisfy

$SO(n)$: rotation matrix	$\mathbf{R} \cdot \mathbf{x} \triangleq \mathbf{R}\mathbf{x}$
$SE(n)$: Euclidean matrix	$\mathbf{H} \cdot \mathbf{x} \triangleq \mathbf{R}\mathbf{x} + \mathbf{t}$
S^1 : unit complex	$\mathbf{z} \cdot \mathbf{x} \triangleq \mathbf{z}\mathbf{x}$
S^3 : unit quaternion	$\mathbf{q} \cdot \mathbf{x} \triangleq \mathbf{q}\mathbf{x}\mathbf{q}^*$

See Table I for a more detailed exposition, and the appendices.

The group composition (1) may be viewed as an action of the group on itself, $\circ : \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{M}$. Another interesting action is the *adjoint action*, which we will see in Section II-F.

C. The tangent spaces and the Lie algebra

Given $\mathcal{X}(t)$ a point moving on a Lie group's manifold \mathcal{M} , its velocity $\dot{\mathcal{X}} = \partial \mathcal{X} / \partial t$ belongs to the space tangent to \mathcal{M} at \mathcal{X} (Fig. 2), which we note $T_{\mathcal{X}} \mathcal{M}$. The smoothness of the manifold, *i.e.*, the absence of edges or spikes, implies the existence of a unique tangent space at each point. The structure of such tangent spaces is the same everywhere.

Table I
TYPICAL LIE GROUPS USED IN 2D AND 3D MOTION, INCLUDING THE TRIVIAL \mathbb{R}^n . SEE THE APPENDICES FOR FULL REFERENCE

Lie group \mathcal{M}, \circ	size	dim	$\mathcal{X} \in \mathcal{M}$	Constraint	$\tau^\wedge \in \mathfrak{m}$	$\tau \in \mathbb{R}^m$	Exp(τ)	Comp.	Action	
n -D vector	$\mathbb{R}^n, +$	n	n	$\mathbf{v} \in \mathbb{R}^n$	$\mathbf{v} - \mathbf{v} = \mathbf{0}$	$\mathbf{v} \in \mathbb{R}^n$	$\mathbf{v} \in \mathbb{R}^n$	$\mathbf{v} = \exp(\mathbf{v})$	$\mathbf{v}_1 + \mathbf{v}_2$	$\mathbf{v} + \mathbf{x}$
circle Rotation	S^1, \cdot $SO(2), \cdot$	2 4	1 1	$\mathbf{z} \in \mathbb{C}$ \mathbf{R}	$\mathbf{z}^* \mathbf{z} = 1$ $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$	$i\theta \in i\mathbb{R}$ $[\theta]_\times \in \mathfrak{so}(2)$	$\theta \in \mathbb{R}$ $\theta \in \mathbb{R}$	$\mathbf{z} = \exp(i\theta)$ $\mathbf{R} = \exp([\theta]_\times)$	$\mathbf{z}_1 \mathbf{z}_2$ $\mathbf{R}_1 \mathbf{R}_2$	$\mathbf{z} \mathbf{x}$ $\mathbf{R} \mathbf{x}$
Rigid motion	$SE(2), \cdot$	9	3	$\mathbf{M} = [\begin{smallmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{smallmatrix}]$	$\mathbf{R}^\top \mathbf{R} = \mathbf{I}$	$[\begin{smallmatrix} [\theta]_\times & \rho \\ 0 & 0 \end{smallmatrix}] \in \mathfrak{se}(2)$	$[\begin{smallmatrix} \rho \\ \theta \end{smallmatrix}] \in \mathbb{R}^3$	$\exp([\begin{smallmatrix} [\theta]_\times & \rho \\ 0 & 0 \end{smallmatrix}])$	$\mathbf{M}_1 \mathbf{M}_2$	$\mathbf{R} \mathbf{x} + \mathbf{t}$
3-sphere Rotation	S^3, \cdot $SO(3), \cdot$	4 9	3 3	$\mathbf{q} \in \mathbb{H}$ \mathbf{R}	$\mathbf{q}^* \mathbf{q} = 1$ $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$	$\theta/2 \in \mathbb{H}_p$ $[\theta]_\times \in \mathfrak{so}(3)$	$\theta \in \mathbb{R}^3$ $\theta \in \mathbb{R}^3$	$\mathbf{q} = \exp(\mathbf{u}\theta/2)$ $\mathbf{R} = \exp([\theta]_\times)$	$\mathbf{q}_1 \mathbf{q}_2$ $\mathbf{R}_1 \mathbf{R}_2$	$\mathbf{q} \mathbf{x} \mathbf{q}^*$ $\mathbf{R} \mathbf{x}$
Rigid motion	$SE(3), \cdot$	16	6	$\mathbf{M} = [\begin{smallmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{smallmatrix}]$	$\mathbf{R}^\top \mathbf{R} = \mathbf{I}$	$[\begin{smallmatrix} [\theta]_\times & \rho \\ 0 & 0 \end{smallmatrix}] \in \mathfrak{se}(3)$	$[\begin{smallmatrix} \rho \\ \theta \end{smallmatrix}] \in \mathbb{R}^6$	$\exp([\begin{smallmatrix} [\theta]_\times & \rho \\ 0 & 0 \end{smallmatrix}])$	$\mathbf{M}_1 \mathbf{M}_2$	$\mathbf{R} \mathbf{x} + \mathbf{t}$

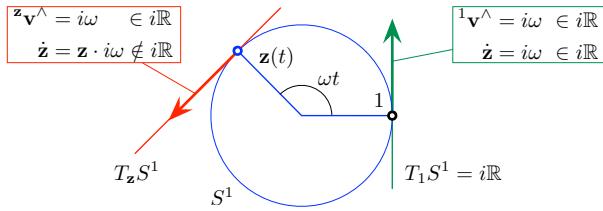


Figure 5. Let a point $\mathbf{z} \in S^1$ move at constant rotation rate ω , $\mathbf{z}(t) = \cos \omega t + i \sin \omega t$. Its velocities when passing through 1 and \mathbf{z} are in the respective tangent spaces, $T_1 S^1$ and $T_z S^1$. In the case of $T_z S^1$, the velocity is $\dot{\mathbf{z}} = \mathbf{z} \omega = -\omega \sin \omega t + i \omega \cos \omega t$ when expressed in the global coordinates, and $\hat{\mathbf{v}}^\wedge = i\omega$ when expressed locally. Their relation is given by $\hat{\mathbf{v}}^\wedge = \mathbf{z}^{-1} \dot{\mathbf{z}} = \mathbf{z}^* \dot{\mathbf{z}}$. In the case of $T_1 S^1$, this relation is the identity $\hat{\mathbf{v}}^\wedge = \dot{\mathbf{z}} = i\omega$. Clearly, the structure of all tangent spaces is $i\mathbb{R}$, which is the Lie algebra. This is also the structure of $\dot{\mathbf{z}}$ at the identity, and this is why the Lie algebra is defined as the tangent space at the identity.

1) *The Lie algebra \mathfrak{m} :* The tangent space at the identity, $T_{\mathcal{E}} \mathcal{M}$, is called the *Lie algebra* of \mathcal{M} , and noted \mathfrak{m} ,

$$\text{Lie algebra : } \mathfrak{m} \triangleq T_{\mathcal{E}} \mathcal{M}. \quad (8)$$

Every Lie group has an associated Lie algebra. We relate the Lie group with its Lie algebra through the following facts [5] (see Figs. 1 and 6):

- The Lie algebra \mathfrak{m} is a vector space.¹ As such, its elements can be *identified* with vectors in \mathbb{R}^m , whose dimension m is the number of degrees of freedom of \mathcal{M} .
- The *exponential map*, $\exp : \mathfrak{m} \rightarrow \mathcal{M}$, exactly converts elements of the Lie algebra into elements of the group. The log map is the inverse operation.
- *Vectors of the tangent space at \mathcal{X} can be transformed to the tangent space at the identity \mathcal{E} through a linear transform. This transform is called the adjoint.*

Lie algebras can be defined locally to a tangent point \mathcal{X} , establishing local coordinates for $T_{\mathcal{X}} \mathcal{M}$ (Fig. 5). We shall denote elements of the Lie algebras with a ‘hat’ decorator, such as \mathbf{v}^\wedge for velocities or $\tau^\wedge = (\mathbf{v}t)^\wedge = \mathbf{v}^\wedge t$ for general elements. A left superscript may also be added to specify the precise tangent space, e.g., ${}^\mathcal{X} \mathbf{v}^\wedge \in T_{\mathcal{X}} \mathcal{M}$ and ${}^\mathcal{E} \mathbf{v}^\wedge \in T_{\mathcal{E}} \mathcal{M}$.

The structure of the Lie algebra can be found (see Examples 3 and 5) by time-differentiating the group constraint (3).

¹In any Lie algebra, the vector space is endowed with a non-associative product called the Lie bracket. In this work, we will not make use of it.

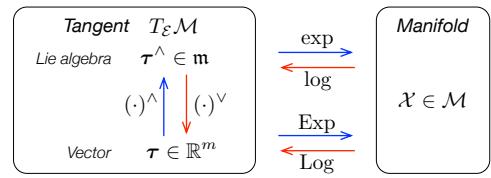
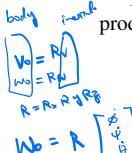


Figure 6. Mappings between the manifold \mathcal{M} and the representations of its tangent space at the origin $T_{\mathcal{E}} \mathcal{M}$ (Lie algebra \mathfrak{m} and Cartesian \mathbb{R}^m). Maps hat $(\cdot)^\wedge$ and vee $(\cdot)^\vee$ are the linear invertible maps or *isomorphisms* (10–11), $\exp(\cdot)$ and $\log(\cdot)$ map the Lie algebra to/from the manifold, and $\text{Exp}(\cdot)$ and $\text{Log}(\cdot)$ are shortcuts to map directly the vector space \mathbb{R}^m to/from \mathcal{M} .

For multiplicative groups this yields the new constraint $\mathcal{X}^{-1} \dot{\mathcal{X}} + \dot{\mathcal{X}}^{-1} \mathcal{X} = 0$, which applies to the elements tangent at \mathcal{X} (the term \mathcal{X}^{-1} is the derivative of the inverse). The elements of the Lie algebra are therefore of the form,²

$$\mathbf{v}^\wedge = \mathcal{X}^{-1} \dot{\mathcal{X}} = -\dot{\mathcal{X}}^{-1} \mathcal{X}. \quad (9)$$

2) *The Cartesian vector space \mathbb{R}^m :* The elements τ^\wedge of the Lie algebra have non-trivial structures (skew-symmetric matrices, imaginary numbers, pure quaternions, see Table I) but the key aspect for us is that they can be expressed as linear combinations of some base elements E_i , where E_i are called the *generators* of \mathfrak{m} (they are the derivatives of \mathcal{X} around the origin in the i -th direction). It is then handy to manipulate just the coordinates as vectors in \mathbb{R}^m , which we shall note simply τ . We may pass from \mathfrak{m} to \mathbb{R}^m and vice versa through two mutually inverse linear maps or *isomorphisms*, commonly called *hat* and *vee* (see Fig. 6),

$$\text{Hat : } \mathbb{R}^m \rightarrow \mathfrak{m}; \quad \tau \mapsto \tau^\wedge = \sum_{i=1}^m \tau_i E_i \quad (10)$$

$$\text{Vee : } \mathfrak{m} \rightarrow \mathbb{R}^m; \quad \tau^\wedge \mapsto (\tau^\wedge)^\vee = \tau = \sum_{i=1}^m \tau_i \mathbf{e}_i, \quad (11)$$

with \mathbf{e}_i the vectors of the base of \mathbb{R}^m (we have $\mathbf{e}_i^\wedge = E_i$). This means that \mathfrak{m} is isomorphic to the vector space \mathbb{R}^m — one writes $\mathfrak{m} \cong \mathbb{R}^m$, or $\tau^\wedge \cong \tau$. Vectors $\tau \in \mathbb{R}^m$ are handier for our purposes than their isomorphic $\tau^\wedge \in \mathfrak{m}$, since they can be stacked in larger state vectors, and more importantly,

²For additive Lie groups the constraint $\mathcal{X} - \dot{\mathcal{X}} = 0$ differentiates to $\dot{\mathcal{X}} = \dot{\mathcal{X}}$, that is, no constraint affects the tangent space. This means that the tangent space is the same as the group space. See App. E for more details.

Example 3: The rotation group $SO(3)$, its Lie algebra $\mathfrak{so}(3)$, and the vector space \mathbb{R}^3

In the rotation group $SO(3)$, of 3×3 rotation matrices \mathbf{R} , we have the orthogonality condition $\mathbf{R}^\top \mathbf{R} = \mathbf{I}$. The tangent space may be found by taking the time derivative of this constraint, that is $\mathbf{R}^\top \dot{\mathbf{R}} + \dot{\mathbf{R}}^\top \mathbf{R} = 0$, which we rearrange as

$$\mathbf{R}^\top \dot{\mathbf{R}} = -(\mathbf{R}^\top \dot{\mathbf{R}})^\top.$$

This expression reveals that $\mathbf{R}^\top \dot{\mathbf{R}}$ is a skew-symmetric matrix (the negative of its transpose). Skew-symmetric matrices are often noted $[\omega]_\times$ and have the form

$$[\omega]_\times = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}.$$

This gives $\mathbf{R}^\top \dot{\mathbf{R}} = [\omega]_\times$. When $\mathbf{R} = \mathbf{I}$ we have

$$\dot{\mathbf{R}} = [\omega]_\times,$$

that is, $[\omega]_\times$ is in the Lie algebra of $SO(3)$, which we name $\mathfrak{so}(3)$. Since $[\omega]_\times \in \mathfrak{so}(3)$ has 3 DoF, the dimension of $SO(3)$ is $m = 3$. The Lie algebra is a vector space whose elements can be decomposed into

$$[\omega]_\times = \omega_x \mathbf{E}_x + \omega_y \mathbf{E}_y + \omega_z \mathbf{E}_z$$

with $\mathbf{E}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$, $\mathbf{E}_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, $\mathbf{E}_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ the generators of $\mathfrak{so}(3)$, and where $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z) \in \mathbb{R}^3$ is the vector of angular velocities. The one-to-one linear relation above allows us to identify $\mathfrak{so}(3)$ with \mathbb{R}^3 — we write $\mathfrak{so}(3) \cong \mathbb{R}^3$. We pass from $\mathfrak{so}(3)$ to \mathbb{R}^3 and viceversa using the linear operators *hat* and *vee*,

$$\text{Hat : } \mathbb{R}^3 \rightarrow \mathfrak{so}(3); \quad \boldsymbol{\omega} \mapsto \boldsymbol{\omega}^\wedge = [\boldsymbol{\omega}]_\times$$

$$\text{Vee : } \mathfrak{so}(3) \rightarrow \mathbb{R}^3; \quad [\boldsymbol{\omega}]_\times \mapsto [\boldsymbol{\omega}]^\vee_\times = \boldsymbol{\omega}.$$

manipulated with linear algebra using matrix operators. In this work, we enforce this preference of \mathbb{R}^m over \mathfrak{m} , to the point that most of the operators and objects that we define (specifically: the adjoint, the Jacobians, the perturbations and their covariances matrices, as we will see soon) are on \mathbb{R}^m .

D. The exponential map

The exponential map $\exp()$ allows us to exactly transfer elements of the Lie algebra to the group (Fig. 1), an operation generically known as *retraction*. Intuitively, $\exp()$ wraps the tangent element around the manifold following the great arc or *geodesic* (as when wrapping a string around a ball, Figs. 1, 3 and 4). The inverse map is the $\log()$, i.e., the unwrapping operation. The $\exp()$ map arises naturally by considering the time-derivatives of $\mathcal{X} \in \mathcal{M}$ over the manifold, as follows. From (9) we have,

$$\dot{\mathcal{X}} = \mathcal{X} \mathbf{v}^\wedge. \quad (12)$$

For \mathbf{v} constant, this is an ordinary differential equation (ODE) whose solution is

$$\mathcal{X}(t) = \mathcal{X}(0) \exp(\mathbf{v}^\wedge t). \quad (13)$$

Example 4: The exponential map of $SO(3)$

We have seen in Ex. 3 that $\dot{\mathbf{R}} = \mathbf{R}[\omega]_\times \in T_{\mathbf{R}} SO(3)$. For $\boldsymbol{\omega}$ constant, this is an ordinary differential equation (ODE), whose solution is $\mathbf{R}(t) = \mathbf{R}_0 \exp([\boldsymbol{\omega}]_\times t)$. At the origin $\mathbf{R}_0 = \mathbf{I}$ we have the exponential map,

$$\mathbf{R}(t) = \exp([\boldsymbol{\omega}]_\times t) \in SO(3).$$

We now define the vector $\boldsymbol{\theta} \triangleq \mathbf{u}\theta \triangleq \boldsymbol{\omega}t \in \mathbb{R}^3$ as the integrated rotation in angle-axis form, with angle θ and unit axis \mathbf{u} . Thus $[\boldsymbol{\theta}]_\times \in \mathfrak{so}(3)$ is the total rotation expressed in the Lie algebra. We substitute it above. Then write the exponential as a power series,

$$\mathbf{R} = \exp([\boldsymbol{\theta}]_\times) = \sum_k \frac{\theta^k}{k!} ([\mathbf{u}]_\times)^k.$$

In order to find a closed-form expression, we write down a few powers of $[\mathbf{u}]_\times$,

$$\begin{aligned} [\mathbf{u}]_\times^0 &= \mathbf{I}, & [\mathbf{u}]_\times^1 &= [\mathbf{u}]_\times, \\ [\mathbf{u}]_\times^2 &= \mathbf{u}\mathbf{u}^\top - \mathbf{I}, & [\mathbf{u}]_\times^3 &= -[\mathbf{u}]_\times, \\ [\mathbf{u}]_\times^4 &= -[\mathbf{u}]_\times^2, & \dots \end{aligned}$$

and realize that all can be expressed as multiples of \mathbf{I} , $[\mathbf{u}]_\times$ or $[\mathbf{u}]_\times^2$. We thus rewrite the series as,

$$\begin{aligned} \mathbf{R} &= \mathbf{I} + [\mathbf{u}]_\times \left(\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 - \dots \right) \\ &\quad + [\mathbf{u}]_\times^2 \left(\frac{1}{2}\theta^2 - \frac{1}{4!}\theta^4 + \frac{1}{6!}\theta^6 - \dots \right), \end{aligned}$$

where we identify the series of $\sin \theta$ and $\cos \theta$, yielding the closed form,

$$\mathbf{R} = \exp([\mathbf{u}\theta]_\times) = \mathbf{I} + [\mathbf{u}]_\times \sin \theta + [\mathbf{u}]_\times^2 (1 - \cos \theta).$$

This expression is the well known Rodrigues rotation formula. It can be used as the capitalized exponential just by doing $\mathbf{R} = \text{Exp}(\mathbf{u}\theta) = \exp([\mathbf{u}\theta]_\times)$.

Since $\mathcal{X}(t)$ and $\mathcal{X}(0)$ are elements of the group, then $\exp(\mathbf{v}^\wedge t) = \mathcal{X}(0)^{-1} \mathcal{X}(t)$ must be in the group too, and so $\exp(\mathbf{v}^\wedge t)$ maps elements $\mathbf{v}^\wedge t$ of the Lie algebra to the group. This is known as the *exponential map*.

In order to provide a more generic definition of the exponential map, let us define the tangent increment $\boldsymbol{\tau} \triangleq \mathbf{v}t \in \mathbb{R}^m$ as velocity per time, so that we have $\boldsymbol{\tau}^\wedge = \mathbf{v}^\wedge t \in \mathfrak{m}$ a point in the Lie algebra. The exponential map, and its inverse the logarithmic map, can be now written as,

$$\exp : \mathfrak{m} \rightarrow \mathcal{M} ; \quad \boldsymbol{\tau}^\wedge \mapsto \mathcal{X} = \exp(\boldsymbol{\tau}^\wedge) \quad (14)$$

$$\log : \mathcal{M} \rightarrow \mathfrak{m} ; \quad \mathcal{X} \mapsto \boldsymbol{\tau}^\wedge = \log(\mathcal{X}) . \quad (15)$$

Closed forms of the exponential in multiplicative groups are obtained by writing the absolutely convergent Taylor series,

$$\exp(\boldsymbol{\tau}^\wedge) = \mathcal{E} + \boldsymbol{\tau}^\wedge + \frac{1}{2}\boldsymbol{\tau}^{\wedge 2} + \frac{1}{3!}\boldsymbol{\tau}^{\wedge 3} + \dots , \quad (16)$$

and taking advantage of the algebraic properties of the powers of $\boldsymbol{\tau}^\wedge$ (see Ex. 4 and 5 for developments of the exponential

Example 5: The unit quaternions group S^3 (cont.)

In the group S^3 (recall Ex. 2 and see e.g. [8]), the time derivative of the unit norm condition $\mathbf{q}^*\mathbf{q} = 1$ yields

$$\mathbf{q}^*\dot{\mathbf{q}} = -(\mathbf{q}^*\dot{\mathbf{q}})^*.$$

This reveals that $\mathbf{q}^*\dot{\mathbf{q}}$ is a pure quaternion (its real part is zero). Pure quaternions $uv \in \mathbb{H}_p$ have the form

$$uv = (iu_x + ju_y + ku_z)v = iv_x + jv_y + kv_z,$$

where $\mathbf{u} \triangleq iu_x + ju_y + ku_z$ is pure and unitary, v is the norm, and i, j, k are the generators of the Lie algebra $\mathfrak{s}^3 = \mathbb{H}_p$. Re-writing the condition above we have,

$$\dot{\mathbf{q}} = \mathbf{q} uv \quad \in T_{\mathbf{q}} S^3,$$

which integrates to $\mathbf{q} = \mathbf{q}_0 \exp(\mathbf{u}vt)$. Letting $\mathbf{q}_0 = 1$ and defining $\phi \triangleq \mathbf{u}\phi \triangleq \mathbf{u}vt$ we get the exponential map,

$$\mathbf{q} = \exp(\mathbf{u}\phi) \triangleq \sum \frac{\phi^k}{k!} \mathbf{u}^k \quad \in S^3.$$

The powers of \mathbf{u} follow the pattern $1, \mathbf{u}, -1, -\mathbf{u}, 1, \dots$. Thus we group the terms in 1 and \mathbf{u} and identify the series of $\cos \phi$ and $\sin \phi$. We get the closed form,

$$\mathbf{q} = \exp(\mathbf{u}\phi) = \cos(\phi) + \mathbf{u} \sin(\phi),$$

which is a beautiful extension of the Euler formula, $\exp(i\phi) = \cos \phi + i \sin \phi$. The elements of the Lie algebra $\phi = \mathbf{u}\phi \in \mathfrak{s}^3$ can be identified with the rotation vector $\theta \in \mathbb{R}^3$ through the mappings *hat* and *vee*,

$$\begin{aligned} \text{Hat : } & \mathbb{R}^3 \rightarrow \mathfrak{s}^3; & \theta \mapsto \theta^\wedge = \theta/2 \\ \text{Vee : } & \mathfrak{s}^3 \rightarrow \mathbb{R}^3; & \phi \mapsto \phi^\vee = 2\phi, \end{aligned}$$

where the factor 2 accounts for the double effect of the quaternion in the rotation action, $\mathbf{x}' = \mathbf{q} \mathbf{x} \mathbf{q}^*$. With this choice of Hat and Vee, the quaternion exponential

$$\mathbf{q} = \text{Exp}(\mathbf{u}\theta) = \cos(\theta/2) + \mathbf{u} \sin(\theta/2)$$

is equivalent to the rotation matrix $\mathbf{R} = \text{Exp}(\mathbf{u}\theta)$.

map in $SO(3)$ and S^3). These are then inverted to find the logarithmic map. Key properties of the exponential map are

$$\exp((t+s)\tau^\wedge) = \exp(t\tau^\wedge) \exp(s\tau^\wedge) \quad (17)$$

$$\exp(t\tau^\wedge) = \exp(\tau^\wedge)^t \quad (18)$$

$$\exp(-\tau^\wedge) = \exp(\tau^\wedge)^{-1} \quad (19)$$

$$\exp(\mathcal{X}\tau^\wedge \mathcal{X}^{-1}) = \mathcal{X} \exp(\tau^\wedge) \mathcal{X}^{-1}, \quad (20)$$

where (20), a surprising and powerful statement, can be proved easily by expanding the Taylor series and simplifying the many terms $\mathcal{X}^{-1}\mathcal{X}$.

1) *The capitalized exponential map:* The capitalized Exp and Log maps are convenient shortcuts to map vector elements $\tau \in \mathbb{R}^m$ ($\cong T_{\mathcal{E}}\mathcal{M}$) directly with elements $\mathcal{X} \in \mathcal{M}$. We have,

$$\text{Exp} : \quad \mathbb{R}^m \rightarrow \mathcal{M} \quad ; \quad \tau \mapsto \mathcal{X} = \text{Exp}(\tau) \quad (21)$$

$$\text{Log} : \quad \mathcal{M} \rightarrow \mathbb{R}^m \quad ; \quad \mathcal{X} \mapsto \tau = \text{Log}(\mathcal{X}). \quad (22)$$

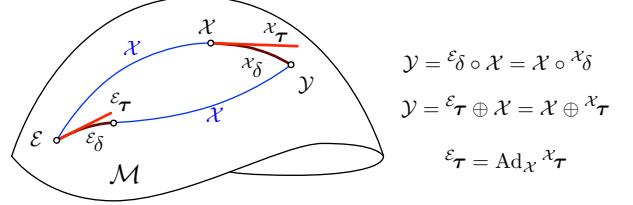


Figure 7. Two paths, $\mathcal{X} \circ x_\delta$ and $\epsilon_\tau \circ \mathcal{X}$, join the origin \mathcal{E} with the point \mathcal{Y} . They both compose the element \mathcal{X} with increments or ‘deltas’ expressed either in the local frame, x_δ , or in the origin, ϵ_τ . Due to non-commutativity, the elements x_δ and ϵ_τ are not equal. Their associated tangent vectors $x_\tau = \text{Log}(x_\delta)$ and $\epsilon_\tau = \text{Log}(\epsilon_\delta)$ are therefore unequal too. They are related by the linear transform $\epsilon_\tau = \text{Ad}_{\mathcal{X}} x_\tau$ where $\text{Ad}_{\mathcal{X}}$ is the adjoint of \mathcal{M} at \mathcal{X} .

$$\begin{bmatrix} \mathcal{L}[\cdot] \\ \mathcal{S}[\cdot] \end{bmatrix} = \begin{bmatrix} \mathcal{L}[\cdot] \\ \mathcal{S}[\cdot] \end{bmatrix} \bullet \begin{bmatrix} \mathcal{L}[x] \\ \mathcal{S}[x] \end{bmatrix}$$

Clearly from Fig. 6,

$$\mathcal{X} = \text{Exp}(\tau) \triangleq \exp(\tau^\wedge) \quad (23)$$

$$\tau = \text{Log}(\mathcal{X}) \triangleq \log(\mathcal{X})^\vee. \quad (24)$$

See the Appendices for details on the implementation of these maps for different manifolds.



E. Plus and minus operators

Plus and minus allow us to introduce increments between elements of a (curved) manifold, and express them in its (flat) tangent vector space. Denoted by \oplus and \ominus , they combine one Exp/Log operation with one composition. Because of the non-commutativity of the composition, they are defined in right- and left- versions depending on the order of the operands. The right operators are (see Fig. 4-right),

$$\text{right-}\oplus : \quad \mathcal{Y} = \mathcal{X} \oplus x_\tau \triangleq \mathcal{X} \circ \text{Exp}(x_\tau) \in \mathcal{M} \quad (25)$$

$$\text{right-}\ominus : \quad x_\tau = \mathcal{Y} \ominus \mathcal{X} \triangleq \text{Log}(\mathcal{X}^{-1} \circ \mathcal{Y}) \in T_{\mathcal{X}}\mathcal{M}. \quad (26)$$

Because in (25) $\text{Exp}(x_\tau)$ appears at the right hand side of the composition, x_τ belongs to the tangent space at \mathcal{X} (see (26)): we say by convention³ that x_τ is expressed in the *local* frame at \mathcal{X} — we note reference frames with a left superscript.

The left operators are,

$$\text{left-}\oplus : \quad \mathcal{Y} = \epsilon_\tau \oplus \mathcal{X} \triangleq \text{Exp}(\epsilon_\tau) \circ \mathcal{X} \in \mathcal{M} \quad (27)$$

$$\text{left-}\ominus : \quad \epsilon_\tau = \mathcal{Y} \ominus \mathcal{X} \triangleq \text{Log}(\mathcal{Y} \circ \mathcal{X}^{-1}) \in T_{\mathcal{E}}\mathcal{M}. \quad (28)$$

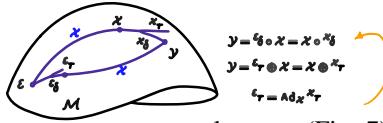
Now, in (27) $\text{Exp}(\epsilon_\tau)$ is on the left and we have $\epsilon_\tau \in T_{\mathcal{E}}\mathcal{M}$: we say that ϵ_τ is expressed in the *global* frame.

Notice that while left- and right- \oplus are distinguished by the operands order, the notation \ominus in (26) and (28) is ambiguous. In this work, we express perturbations locally by default and therefore we use the right- forms of \oplus and \ominus by default.

F. The adjoint, and the adjoint matrix

If we identify \mathcal{Y} in (25, 27), we arrive at $\epsilon_\tau \oplus \mathcal{X} = \mathcal{X} \oplus x_\tau$, which determines a relation between the local and global

³The convention sticks to that of frame transformation, e.g. $G_{\mathbf{x}} = \mathbf{R}^{L_{\mathbf{x}}}$, where the matrix $\mathbf{R} \in SO(3)$ transforms local vectors into global. Notice that this convention is not shared by all authors, and for example [9] uses the opposite, $L_{\mathbf{x}} = \mathbf{R}^{G_{\mathbf{x}}}$.



tangent elements (Fig. 7). We develop it with (20, 25, 27) as

$$\begin{aligned}\text{Exp}^{\mathcal{E}_\tau} \mathcal{X} &= \mathcal{X} \text{Exp}(\mathcal{X}_\tau) \\ \exp(\mathcal{E}_\tau^\wedge) &= \mathcal{X} \exp(\mathcal{X}_\tau^\wedge) \mathcal{X}^{-1} = \exp(\mathcal{X} \mathcal{X}_\tau^\wedge \mathcal{X}^{-1}) \\ \mathcal{E}_\tau^\wedge &= \mathcal{X} \boxed{\mathcal{X}_\tau^\wedge} \mathcal{X}^{-1}\end{aligned}$$

1) *The adjoint*: We thus define the *adjoint* of \mathcal{M} at \mathcal{X} , noted $\text{Ad}_{\mathcal{X}}$, to be

$$\text{Ad}_{\mathcal{X}} : \mathfrak{m} \rightarrow \mathfrak{m}; \quad \tau^\wedge \mapsto \text{Ad}_{\mathcal{X}}(\tau^\wedge) \triangleq \mathcal{X} \tau^\wedge \mathcal{X}^{-1}, \quad (29)$$

so that $\mathcal{E}_\tau^\wedge = \text{Ad}_{\mathcal{X}}(\mathcal{X}_\tau^\wedge)$. This defines the *adjoint action* of the group on its own Lie algebra. The adjoint has two interesting (and easy to prove) properties,

$$\text{Linear : } \text{Ad}_{\mathcal{X}}(a\tau^\wedge + b\sigma^\wedge) = a\text{Ad}_{\mathcal{X}}(\tau^\wedge) + b\text{Ad}_{\mathcal{X}}(\sigma^\wedge)$$

$$\text{Homomorphism : } \text{Ad}_{\mathcal{X}}(\text{Ad}_{\mathcal{Y}}(\tau^\wedge)) = \text{Ad}_{\mathcal{X}\mathcal{Y}}(\tau^\wedge).$$

2) *The adjoint matrix*: Since $\text{Ad}_{\mathcal{X}}()$ is linear, we can find an equivalent matrix operator $\mathbf{Ad}_{\mathcal{X}}$ that maps the Cartesian tangent vectors $\mathcal{E}_\tau \cong \mathcal{E}_\tau^\wedge$ and $\mathcal{X}_\tau \cong \mathcal{X}_\tau^\wedge$,

$$\mathbf{Ad}_{\mathcal{X}} : \mathbb{R}^m \rightarrow \mathbb{R}^m; \quad \mathcal{X}_\tau \mapsto \mathcal{E}_\tau = \mathbf{Ad}_{\mathcal{X}} \mathcal{X}_\tau, \quad (30)$$

which we call the *adjoint matrix*. This can be computed by applying ${}^\vee$ to (29), thus writing

$$\mathbf{Ad}_{\mathcal{X}} \tau = (\mathcal{X} \tau^\wedge \mathcal{X}^{-1})^\vee, \quad (31)$$

then developing the right hand side to identify the adjoint matrix (see Ex. 6 and the appendices). Additional properties of the adjoint matrix are,

$$\mathcal{X} \oplus \tau = (\mathbf{Ad}_{\mathcal{X}} \tau) \oplus \mathcal{X} \quad (32)$$

$$\mathbf{Ad}_{\mathcal{X}^{-1}} = \mathbf{Ad}_{\mathcal{X}}^{-1} \quad (33)$$

$$\mathbf{Ad}_{\mathcal{X}\mathcal{Y}} = \mathbf{Ad}_{\mathcal{X}} \mathbf{Ad}_{\mathcal{Y}}. \quad (34)$$

Notice in (33, 34) that the left parts of the equality are usually cheaper to compute than the right ones. We will use the adjoint matrix often as a way to linearly transform vectors of the tangent space at \mathcal{X} onto vectors of the tangent space at the origin, with $\mathcal{E}_\tau = \mathbf{Ad}_{\mathcal{X}} \mathcal{X}_\tau$, (30). In this work, the adjoint matrix will be referred to as simply the adjoint.

G. Derivatives on Lie groups

Among the different ways to define derivatives in the context of Lie groups, we concentrate on those in the form of Jacobian matrices mapping vector tangent spaces. This is sufficient here since in these spaces uncertainties and increments can be properly and easily defined. Using these Jacobians, the formulas for uncertainty management in Lie groups will largely resemble those in vector spaces.

The Jacobians described hereafter fulfill the chain rule, so that we can easily compute any Jacobian from the partial Jacobian blocks of *inversion*, *composition*, *exponentiation* and *action*. See Section III-A for details and proofs.

Example 6: The adjoint matrix of $SE(3)$

The $SE(3)$ group of rigid body motions (see App. D) has group, Lie algebra and vector elements,

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}, \quad \tau^\wedge = \begin{bmatrix} [\theta]_\times & \boldsymbol{\rho} \\ \mathbf{0} & 0 \end{bmatrix}, \quad \tau = \begin{bmatrix} \boldsymbol{\rho} \\ \theta \end{bmatrix}.$$

The adjoint matrix is identified by developing (31) as

$$\begin{aligned}\mathbf{Ad}_{\mathbf{M}} \tau &= (\mathbf{M} \tau^\wedge \mathbf{M}^{-1})^\vee = \dots = \\ &= \left(\begin{bmatrix} \mathbf{R} [\theta]_\times \mathbf{R}^\top & -\mathbf{R} [\theta]_\times \mathbf{R}^\top \mathbf{t} + \mathbf{R} \boldsymbol{\rho} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right)^\vee \\ &= \left(\begin{bmatrix} [\mathbf{R} \theta]_\times & [\mathbf{t}]_\times \mathbf{R} \theta + \mathbf{R} \boldsymbol{\rho} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right)^\vee \\ &= \begin{bmatrix} [\mathbf{t}]_\times \mathbf{R} \theta + \mathbf{R} \boldsymbol{\rho} \\ \mathbf{R} \theta \end{bmatrix} = \begin{bmatrix} \mathbf{R} & [\mathbf{t}]_\times \mathbf{R} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho} \\ \theta \end{bmatrix}\end{aligned}$$

where we used $[\mathbf{R} \theta]_\times = \mathbf{R} [\theta]_\times \mathbf{R}^\top$ and $[\mathbf{a}]_\times \mathbf{b} = -[\mathbf{b}]_\times \mathbf{a}$. So the adjoint matrix is

$$\mathbf{Ad}_{\mathbf{M}} = \begin{bmatrix} \mathbf{R} & [\mathbf{t}]_\times \mathbf{R} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \in \mathbb{R}^{6 \times 6}.$$

1) *Reminder: Jacobians on vector spaces*: For a multivariate function $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, the Jacobian matrix is defined as the $n \times m$ matrix stacking all partial derivatives,

$$\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_m} \end{bmatrix} \in \mathbb{R}^{n \times m}. \quad (35)$$

It is handy to define this matrix in the following form. Let us partition $\mathbf{J} = [\mathbf{j}_1 \dots \mathbf{j}_m]$, and let $\mathbf{j}_i = [\frac{\partial f_1}{\partial x_i} \dots \frac{\partial f_n}{\partial x_i}]^\top$ be its i -th column vector. This column vector responds to

$$\mathbf{j}_i = \frac{\partial f(\mathbf{x})}{\partial x_i} \triangleq \lim_{h \rightarrow 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h} \in \mathbb{R}^n, \quad (36)$$

where \mathbf{e}_i is the i -th vector of the natural basis of \mathbb{R}^m . Regarding the numerator, notice that the vector

$$\mathbf{v}_i(h) \triangleq f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x}) \in \mathbb{R}^n \quad (37)$$

is the variation of $f(\mathbf{x})$ when \mathbf{x} is perturbed in the direction of \mathbf{e}_i , and that the respective Jacobian column is just $\mathbf{j}_i = \partial \mathbf{v}_i(h)/\partial h|_{h=0} = \lim_{h \rightarrow 0} \mathbf{v}_i(h)/h$. In this work, for the sake of convenience, we introduce the compact form,

$$\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \triangleq \lim_{\mathbf{h} \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})}{\mathbf{h}} \in \mathbb{R}^{n \times m}, \quad (38)$$

with $\mathbf{h} \in \mathbb{R}^m$, which aglutinates all columns (36) to form the definition of (35). We remark that (38) is just a notation convenience (just as (35) is), since division by the vector \mathbf{h} is undefined and proper computation requires (36). However, this form may be used to calculate Jacobians by developing the numerator into a form linear in \mathbf{h} , and identifying the left hand side as the Jacobian, that is,

$$\lim_{\mathbf{h} \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})}{\mathbf{h}} = \dots = \lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{J}\mathbf{h}}{\mathbf{h}} \triangleq \frac{\partial \mathbf{J}\mathbf{h}}{\partial \mathbf{h}} = \mathbf{J}. \quad (39)$$

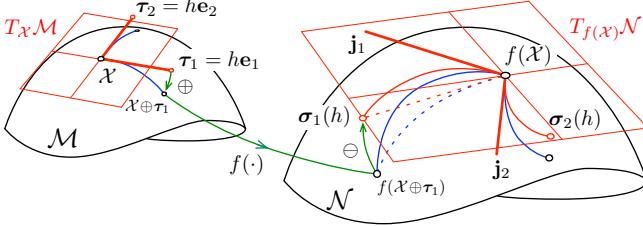


Figure 8. Right Jacobian of a function $f : \mathcal{M} \rightarrow \mathcal{N}$. The perturbation vectors in the canonical directions, $\tau_i = h \mathbf{e}_i \in T_{\mathcal{X}} \mathcal{M}$, are propagated to perturbation vectors $\sigma_i \in T_{f(\mathcal{X})} \mathcal{N}$ through the processes of plus, apply $f()$, and minus (green arrows), obtaining $\sigma_i(h) = f(\mathcal{X} \oplus h \mathbf{e}_i) \ominus f(\mathcal{X})$. For varying values of h , notice that in \mathcal{M} the perturbations $\tau_i(h) = h \mathbf{e}_i$ (thick red) produce paths in \mathcal{M} (blue) along the geodesic (recall Fig. 1). Notice also that in \mathcal{N} , due to the non-linearity of $f()$, the image paths (solid blue) are generally not in the geodesic (dashed blue). These image paths are lifted onto the tangent space $T_{f(\mathcal{X})} \mathcal{N}$, producing smooth curved paths (thin solid red). The column vectors j_i of J (thick red) are the derivatives of the lifted paths evaluated at $f(\mathcal{X})$, i.e., $j_i = \lim_{h \rightarrow 0} \sigma_i(h)/h$. Each $h \mathbf{e}_i \in T_{\mathcal{X}} \mathcal{M}$ gives place to a $j_i \in T_{f(\mathcal{X})} \mathcal{N}$, and thus the resulting Jacobian matrix $J = [j_1 \cdots j_m] \in \mathbb{R}^{n \times m}$ linearly maps vectors from $T_{\mathcal{X}} \mathcal{M} \cong \mathbb{R}^m$ to $T_{f(\mathcal{X})} \mathcal{N} \cong \mathbb{R}^n$.

Notice finally that for small values of h we have the linear approximation,

$$f(\mathbf{x} + \mathbf{h}) \xrightarrow{\mathbf{h} \rightarrow 0} f(\mathbf{x}) + \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \mathbf{h}. \quad (40)$$

2) *Right Jacobians on Lie groups:* Inspired by the standard derivative definition (38) above, we can now use our \oplus and \ominus operators to define Jacobians of functions $f : \mathcal{M} \rightarrow \mathcal{N}$ acting on manifolds (see Fig. 8). Using the right- $\{\oplus, \ominus\}$ in place of $\{+, -\}$ we obtain a form akin to the standard derivative,⁴

$$\frac{\mathcal{X} Df(\mathcal{X})}{D\mathcal{X}} \triangleq \lim_{\tau \rightarrow 0} \frac{f(\mathcal{X} \oplus \tau) \ominus f(\mathcal{X})}{\tau} \in \mathbb{R}^{n \times m} \quad (41a)$$

which develops as,

$$= \lim_{\tau \rightarrow 0} \frac{\text{Log}(f(\mathcal{X})^{-1} \circ f(\mathcal{X} \circ \text{Exp}(\tau)))}{\tau} \quad (41b)$$

$$= \left. \frac{\partial \text{Log}(f(\mathcal{X})^{-1} \circ f(\mathcal{X} \circ \text{Exp}(\tau)))}{\partial \tau} \right|_{\tau=0}. \quad (41c)$$

We call this Jacobian the *right Jacobian* of f . Notice that (41c) is just the standard derivative (38) of the rather complicated function $g(\tau) = \text{Log}(f(\mathcal{X})^{-1} \circ f(\mathcal{X} \circ \text{Exp}(\tau)))$. Writing it as in (41a) conveys much more intuition: it is the derivative of $f(\mathcal{X})$ with respect to \mathcal{X} , only that we expressed the infinitesimal variations in the tangent spaces! Indeed, thanks to the way right- \oplus and \ominus operate, variations in \mathcal{X} and $f(\mathcal{X})$ are now expressed as vectors in the local tangent spaces, i.e., tangent respectively at $\mathcal{X} \in \mathcal{M}$ and $f(\mathcal{X}) \in \mathcal{N}$. This derivative is then a proper Jacobian matrix $\mathbb{R}^{n \times m}$ linearly mapping the local tangent spaces $T_{\mathcal{X}} \mathcal{M} \rightarrow T_{f(\mathcal{X})} \mathcal{N}$ (and we mark the derivative with a local \mathcal{X} superscript). Just as in vector spaces, the columns of this matrix correspond to directional derivatives. That is, the vector

$$\sigma_i(h) = f(\mathcal{X} \oplus h \mathbf{e}_i) \ominus f(\mathcal{X}) \in \mathbb{R}^n \quad (42)$$

⁴The notation $\frac{D\mathcal{Y}}{D\mathcal{X}} = \frac{Df(\mathcal{X})}{D\mathcal{X}}$ is chosen in front of other alternatives in order to make the chain rule readable, i.e., $\frac{D\mathcal{Z}}{D\mathcal{X}} = \frac{D\mathcal{Z}}{D\mathcal{Y}} \frac{D\mathcal{Y}}{D\mathcal{X}}$. We will later introduce the lighter notation $\mathbf{J}_{\mathcal{X}}^{\mathcal{Y}} \triangleq \frac{D\mathcal{Y}}{D\mathcal{X}}$.

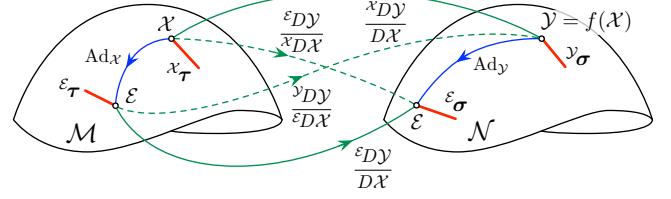


Figure 9. Linear maps between all tangent spaces involved in a function $\mathcal{Y} = f(\mathcal{X})$, from \mathcal{M} to \mathcal{N} . The linear maps $\varepsilon_\tau = \text{Ad}_{\mathcal{X}} x_\tau$, $\varepsilon_\sigma = \text{Ad}_{\mathcal{Y}} y_\sigma$, $\varepsilon_\tau = \frac{x_DY}{D\mathcal{X}} \varepsilon_\tau$, and $v_\sigma = \frac{x_DY}{D\mathcal{X}} x_\tau$, form a loop (solid) that leads to (46). The crossed Jacobians (dashed) form more mapping loops leading to (47,48).

(see Fig. 8 again, and compare σ_i in (42) with \mathbf{v}_i in (37)) is the variation of $f(\mathcal{X})$ when \mathcal{X} varies in the direction of \mathbf{e}_i . Its respective Jacobian column is $j_i = \partial \sigma_i(h)/\partial h|_{h=0}$.

As before, we use (41a) to actually find Jacobians by resorting to the same mechanism (39). For example, for a 3D rotation $f : SO(3) \rightarrow \mathbb{R}^3$; $f(\mathbf{R}) = \mathbf{R}\mathbf{p}$, we have $\mathcal{M} = SO(3)$ and $\mathcal{N} = \mathbb{R}^3$ and so (see App. B-C5),

$$\begin{aligned} \frac{R D R p}{D R} &= \lim_{\theta \rightarrow 0} \frac{(R \oplus \theta)p \ominus R p}{\theta} = \lim_{\theta \rightarrow 0} \frac{R \text{Exp}(\theta)p - R p}{\theta} \\ &= \lim_{\theta \rightarrow 0} \frac{R(I + [\theta]_x)p - R p}{\theta} = \lim_{\theta \rightarrow 0} \frac{R[\theta]_x p}{\theta} \\ &= \lim_{\theta \rightarrow 0} \frac{-R[p]_x \theta}{\theta} = -R[p]_x \in \mathbb{R}^{3 \times 3}. \end{aligned}$$

Many examples of this mechanism can be observed in Section III and the appendices. Remark that whenever the function f passes from one manifold to another, the plus and minus operators in (41a) must be selected appropriately: \oplus for the domain \mathcal{M} , and \ominus for the codomain or image \mathcal{N} .

For small values of τ , the following approximation holds,

$$f(\mathcal{X} \oplus x_\tau) \xrightarrow{x_\tau \rightarrow 0} f(\mathcal{X}) \oplus \frac{\mathcal{X} Df(\mathcal{X})}{D\mathcal{X}} x_\tau \in \mathcal{N}. \quad (43)$$

3) *Left Jacobians on Lie groups:* Derivatives can also be defined from the left- plus and minus operators, leading to,

$$\begin{aligned} \frac{\varepsilon Df(\mathcal{X})}{D\mathcal{X}} &\triangleq \lim_{\tau \rightarrow 0} \frac{f(\tau \oplus \mathcal{X}) \ominus f(\mathcal{X})}{\tau} \in \mathbb{R}^{n \times m} \quad (44) \\ &= \lim_{\tau \rightarrow 0} \frac{\text{Log}(f(\text{Exp}(\tau) \circ \mathcal{X}) \circ f(\mathcal{X})^{-1})}{\tau} \\ &= \left. \frac{\partial \text{Log}(f(\text{Exp}(\tau) \circ \mathcal{X}) \circ f(\mathcal{X})^{-1})}{\partial \tau} \right|_{\tau=0}, \end{aligned}$$

which we call the *left Jacobian* of f . Notice that now $\tau \in T_E \mathcal{M}$, and the numerator belongs to $T_E \mathcal{N}$, thus the left Jacobian is a $n \times m$ matrix mapping the global tangent spaces, $T_E \mathcal{M} \rightarrow T_E \mathcal{N}$, which are the Lie algebras of \mathcal{M} and \mathcal{N} (and we mark the derivative with a global or origin ‘ \mathcal{E} ’ superscript). For small values of τ the following holds,

$$f(\mathcal{E}_\tau \oplus \mathcal{X}) \xrightarrow{\mathcal{E}_\tau \rightarrow 0} \frac{\mathcal{E} Df(\mathcal{X})}{D\mathcal{X}} \mathcal{E}_\tau \oplus f(\mathcal{X}) \in \mathcal{N}. \quad (45)$$

We can show from (32, 43, 45) (see Fig. 9) that left and right Jacobians are related by the adjoints of \mathcal{M} and \mathcal{N} ,

$$\frac{\mathcal{E} Df(\mathcal{X})}{D\mathcal{X}} \text{Ad}_{\mathcal{X}} = \text{Ad}_{f(\mathcal{X})} \frac{\mathcal{X} Df(\mathcal{X})}{D\mathcal{X}}. \quad (46)$$

$\boxed{\mathcal{X}_{k+1}} = \boxed{\mathcal{E} \Phi(\mathcal{T})} \circ \boxed{\mathcal{X}_k}$ and then left Jacobian

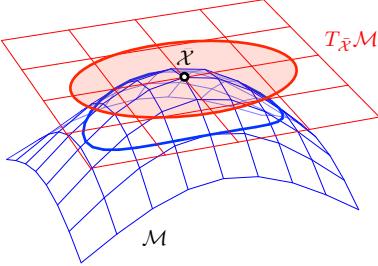


Figure 10. Uncertainty around a point $\bar{\mathcal{X}} \in \mathcal{M}$ is properly expressed as a covariance on the vector space tangent at the point (red). Using \oplus (51), the probability ellipses in the tangent space are wrapped over the manifold (blue), thus illustrating the probability concentration region on the group.

4) *Crossed right-left Jacobians*: One can also define Jacobians using right-plus but left-minus, or vice versa. Though improbable, these are sometimes useful, since they map local to global tangents or vice versa. To keep it short, we will just relate them to the other Jacobians through the adjoints,

$$\frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \text{Ad}_{\mathcal{X}} = \text{Ad}_{\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \quad (47)$$

$$\frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \text{Ad}_{\mathcal{X}}^{-1} = \text{Ad}_{\mathcal{Y}}^{-1} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}}, \quad (48)$$

where $\mathcal{Y} = f(\mathcal{X})$. Now, the upper and lower super-scripts indicate the reference frames where the differentials are expressed. Respective small-tau approximations read,

$$f(\mathcal{X} \oplus {}^x\boldsymbol{\tau}) \xrightarrow[\mathcal{X}\boldsymbol{\tau} \rightarrow 0]{} \frac{\varepsilon Df(\mathcal{X})}{\varepsilon D\mathcal{X}} {}^x\boldsymbol{\tau} \oplus f(\mathcal{X}) \quad (49)$$

$$f({}^\varepsilon\boldsymbol{\tau} \oplus \mathcal{X}) \xrightarrow[\varepsilon\boldsymbol{\tau} \rightarrow 0]{} f(\mathcal{X}) \oplus \frac{f(\mathcal{X}) Df(\mathcal{X})}{\varepsilon D\mathcal{X}} {}^\varepsilon\boldsymbol{\tau}. \quad (50)$$

H. Uncertainty in manifolds, covariance propagation

We define local perturbations $\boldsymbol{\tau}$ around a point $\bar{\mathcal{X}} \in \mathcal{M}$ in the tangent vector space $T_{\bar{\mathcal{X}}}\mathcal{M}$, using right- \oplus and \ominus ,

$$\mathcal{X} = \bar{\mathcal{X}} \oplus \boldsymbol{\tau}, \quad \boldsymbol{\tau} = \mathcal{X} \ominus \bar{\mathcal{X}} \in T_{\bar{\mathcal{X}}}\mathcal{M}. \quad (51)$$

Covariance matrices can be properly defined on this tangent space at $\bar{\mathcal{X}}$ through the standard expectation operator $\mathbb{E}[\cdot]$,

$$\Sigma_{\mathcal{X}} \triangleq \mathbb{E}[\boldsymbol{\tau}\boldsymbol{\tau}^T] = \mathbb{E}[(\mathcal{X} \ominus \bar{\mathcal{X}})(\mathcal{X} \ominus \bar{\mathcal{X}})^T] \in \mathbb{R}^{m \times m}, \quad (52)$$

allowing us to define Gaussian variables on manifolds, $\mathcal{X} \sim \mathcal{N}(\bar{\mathcal{X}}, \Sigma_{\mathcal{X}})$, see Fig. 10. Notice that although we write $\Sigma_{\mathcal{X}}$, the covariance is rather that of the tangent perturbation $\boldsymbol{\tau}$. Since the dimension m of $T\mathcal{M}$ matches the degrees of freedom of \mathcal{M} , these covariances are well defined.⁵

Perturbations can also be expressed in the global reference, that is, in the tangent space at the origin $T_{\mathcal{E}}\mathcal{M}$, using left- \oplus and \ominus ,

$$\mathcal{X} = \boldsymbol{\tau} \oplus \bar{\mathcal{X}}, \quad \boldsymbol{\tau} = \mathcal{X} \ominus \bar{\mathcal{X}} \in T_{\mathcal{E}}\mathcal{M}. \quad (53)$$

This allows global specification of covariance matrices using left-minus in (52). For example, a 3D orientation that is known up to rotations in the horizontal plane can be associated to

⁵A naive definition $\Sigma_{\mathcal{X}} \triangleq \mathbb{E}[(\mathcal{X} - \bar{\mathcal{X}})(\mathcal{X} - \bar{\mathcal{X}})^T]$ is always ill-defined if $\text{size}(\mathcal{X}) > \dim(\mathcal{M})$, which is the case for most non-trivial manifolds.

$$\begin{bmatrix} [k+1] \\ SE(3) \end{bmatrix} = \begin{bmatrix} k \\ SE(3) \end{bmatrix} \bullet \begin{bmatrix} d\mathcal{X} \\ SE(3) \end{bmatrix}$$

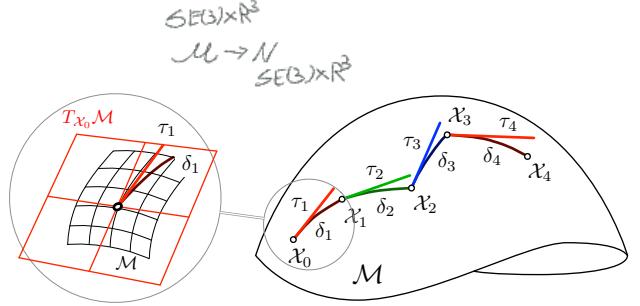


Figure 11. Motion integration on a manifold. Each motion data produces a step $\tau_k \in T_{X_{k-1}}\mathcal{M}$, which is wrapped to a local motion increment or ‘delta’ $\delta_k = \text{Exp}(\boldsymbol{\tau}_k) \in \mathcal{M}$, and then composed with X_{k-1} to yield $X_k = X_{k-1} \circ \delta_k = X_{k-1} \circ \text{Exp}(\boldsymbol{\tau}_k) = X_{k-1} \oplus \boldsymbol{\tau}_k \in \mathcal{M}$.

a covariance ${}^\varepsilon\Sigma = \text{diag}(\sigma_\phi^2, \sigma_\theta^2, \infty)$. Since “horizontal” is a global specification, ${}^\varepsilon\Sigma$ must be specified in the global reference.

Since global and local perturbations are related by the adjoint (30), their covariances can be transformed with

$$\varepsilon\Sigma_{\mathcal{X}} = \text{Ad}_{\mathcal{X}} {}^\varepsilon\Sigma_{\mathcal{X}} \text{Ad}_{\mathcal{X}}^T. \quad (54)$$

Covariance propagation through a function $f : \mathcal{M} \rightarrow \mathcal{N}; \mathcal{X} \mapsto \mathcal{Y} = f(\mathcal{X})$ just requires the linearization (43) with Jacobian matrices (41a) to yield the familiar formula,

$$\Sigma_{\mathcal{Y}} \approx \frac{Df}{D\mathcal{X}} \Sigma_{\mathcal{X}} \frac{Df^T}{D\mathcal{X}} \in \mathbb{R}^{n \times n}. \quad (55)$$

I. Discrete integration on manifolds

The exponential map $\mathcal{X}(t) = \mathcal{X}_0 \circ \text{Exp}(\mathbf{v}t)$ performs the continuous-time integral of constant velocities $\mathbf{v} \in T_{\mathcal{X}_0}\mathcal{M}$ onto the manifold. Non-constant velocities $\mathbf{v}(t)$ are typically handled by segmenting them into piecewise constant bits $\mathbf{v}_k \in T_{\mathcal{X}_{k-1}}\mathcal{M}$, of (short) duration δt_k , and writing the discrete integral

$$\begin{aligned} \mathcal{X}_k &= \mathcal{X}_0 \circ \text{Exp}(\mathbf{v}_1\delta t_1) \circ \text{Exp}(\mathbf{v}_1\delta t_2) \circ \cdots \circ \text{Exp}(\mathbf{v}_k\delta t_k) \\ &= \mathcal{X}_0 \oplus \mathbf{v}_1\delta t_1 \oplus \mathbf{v}_1\delta t_2 \oplus \cdots \oplus \mathbf{v}_k\delta t_k. \end{aligned}$$

Equivalently (Fig. 11), we can define $\tau_k = \mathbf{v}_k\delta t_k$ and construct the integral as a “sum” of (small) discrete tangent steps $\tau_k \in T_{\mathcal{X}_{k-1}}\mathcal{M}$, i.e., $\mathcal{X}_k \triangleq \mathcal{X}_0 \oplus \tau_1 \oplus \tau_2 \oplus \cdots \oplus \tau_k$. We write all these variants in recursive form,

$$\mathcal{X}_k = \mathcal{X}_{k-1} \oplus \tau_k = \mathcal{X}_{k-1} \circ \text{Exp}(\boldsymbol{\tau}_k) = \mathcal{X}_{k-1} \circ \text{Exp}(\mathbf{v}_k\delta t_k). \quad (56)$$

Common examples are the integration of 3D angular rates $\boldsymbol{\omega}$ into the rotation matrix, $\mathbf{R}_k = \mathbf{R}_{k-1} \text{Exp}(\boldsymbol{\omega}_k\delta t)$, or into the quaternion, $\mathbf{q}_k = \mathbf{q}_{k-1} \text{Exp}(\boldsymbol{\omega}_k\delta t)$.

III. DIFFERENTIATION RULES ON MANIFOLDS

For all the typical manifolds \mathcal{M} that we use, we can determine closed forms for the elementary Jacobians of *inversion*, *composition*, *exponentiation* and *action*. Moreover, some of these forms can be related to the adjoint $\text{Ad}_{\mathcal{X}}$, which becomes a central block of the differentiation process. Other forms for Log, \oplus and \ominus can be easily derived from them. Once these forms or ‘blocks’ are found, all other Jacobians follow by the chain rule. Except for the so-called *left Jacobian*, which

we also present below, all Jacobians developed here are right-Jacobians, *i.e.*, defined by (41a). By following the hints here, the interested reader should find no particular difficulties in developing the left-Jacobians. For the reader not willing to do this effort, equation (46) can be used to this end, since

$$\frac{\varepsilon Df(\mathcal{X})}{D\mathcal{X}} = \mathbf{Ad}_{f(\mathcal{X})} \frac{\mathcal{X} Df(\mathcal{X})}{D\mathcal{X}} \mathbf{Ad}_{\mathcal{X}}^{-1}. \quad (57)$$

We use the notations $\mathbf{J}_{\mathcal{X}}^{f(\mathcal{X})} \triangleq \frac{Df(\mathcal{X})}{D\mathcal{X}}$ and $\mathbf{J}_{\mathcal{X}}^{\mathcal{Y}} \triangleq \frac{D\mathcal{Y}}{D\mathcal{X}}$. We notice also that $\mathbf{Ad}_{\mathcal{X}}^{-1}$ should rather be implemented by $\mathbf{Ad}_{\mathcal{X}^{-1}}$ —see (33, 34) and the comment below them.

A. The chain rule

For $\mathcal{Y} = f(\mathcal{X})$ and $\mathcal{Z} = g(\mathcal{Y})$ we have $\mathcal{Z} = g(f(\mathcal{X}))$. The chain rule simply states,

$$\frac{D\mathcal{Z}}{D\mathcal{X}} = \frac{D\mathcal{Z}}{D\mathcal{Y}} \frac{D\mathcal{Y}}{D\mathcal{X}} \quad \text{or} \quad \mathbf{J}_{\mathcal{X}}^{\mathcal{Z}} = \mathbf{J}_{\mathcal{Y}}^{\mathcal{Z}} \mathbf{J}_{\mathcal{X}}^{\mathcal{Y}}. \quad (58)$$

We prove it here for the right Jacobian using (43) thrice,

$$\begin{aligned} g(f(\mathcal{X})) \oplus \mathbf{J}_{\mathcal{X}}^{\mathcal{Z}} \boldsymbol{\tau} &\leftarrow g(f(\mathcal{X} \oplus \boldsymbol{\tau})) \rightarrow g(f(\mathcal{X}) \oplus \mathbf{J}_{\mathcal{X}}^{\mathcal{Y}} \boldsymbol{\tau}) \\ &\rightarrow g(f(\mathcal{X})) \oplus \mathbf{J}_{\mathcal{Y}}^{\mathcal{Z}} \mathbf{J}_{\mathcal{X}}^{\mathcal{Y}} \boldsymbol{\tau} \end{aligned}$$

with the arrows indicating limit as $\boldsymbol{\tau} \rightarrow 0$, and so $\mathbf{J}_{\mathcal{X}}^{\mathcal{Z}} = \mathbf{J}_{\mathcal{Y}}^{\mathcal{Z}} \mathbf{J}_{\mathcal{X}}^{\mathcal{Y}}$. The proof for the left and crossed Jacobians is akin, using respectively (45, 49, 50). Notice that when mixing right, left and crossed Jacobians, we need to chain also the reference frames, as in *e.g.*

$$\frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \quad (59)$$

$$\frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} = \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}}, \quad (60)$$

where the first identity of (59) is proven by writing,

$$\begin{aligned} g(f(\varepsilon_{\boldsymbol{\tau}} \oplus \mathcal{X})) &\xrightarrow[\varepsilon_{\boldsymbol{\tau}} \rightarrow 0]{(50)} g(f(\mathcal{X})) \oplus \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{X}} \varepsilon_{\boldsymbol{\tau}}; \\ g(f(\varepsilon_{\boldsymbol{\tau}} \oplus \mathcal{X})) &\xrightarrow[\varepsilon_{\boldsymbol{\tau}} \rightarrow 0]{(50)} g\left(f(\mathcal{X}) \oplus \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \varepsilon_{\boldsymbol{\tau}}\right) \rightarrow \\ &\xrightarrow{(43)} g(f(\mathcal{X})) \oplus \frac{\varepsilon D\mathcal{Z}}{\varepsilon D\mathcal{Y}} \frac{\varepsilon D\mathcal{Y}}{\varepsilon D\mathcal{X}} \varepsilon_{\boldsymbol{\tau}}, \end{aligned}$$

and identifying (59) in the first and third rows.

B. Elementary Jacobian blocks

1) Inverse:

We define with (41a)

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} \triangleq \frac{\mathcal{X} D\mathcal{X}^{-1}}{D\mathcal{X}} \in \mathbb{R}^{m \times m}. \quad (61)$$

This can be determined from the adjoint using (20) and (31),

$$\begin{aligned} \mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{\text{Log}((\mathcal{X}^{-1})^{-1}(\mathcal{X} \text{Exp}(\boldsymbol{\tau}))^{-1})}{\boldsymbol{\tau}} \\ &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{\text{Log}(\mathcal{X} \text{Exp}(-\boldsymbol{\tau}) \mathcal{X}^{-1})}{\boldsymbol{\tau}} \\ &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{(\mathcal{X}(-\boldsymbol{\tau})^\wedge \mathcal{X}^{-1})^\vee}{\boldsymbol{\tau}} = -\mathbf{Ad}_{\mathcal{X}}. \quad (62) \end{aligned}$$

2) *Composition:* We define with (41a)

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X} \circ \mathcal{Y}} \triangleq \frac{\mathcal{X} D\mathcal{X} \circ \mathcal{Y}}{D\mathcal{X}} \in \mathbb{R}^{m \times m} \quad (63)$$

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{X} \circ \mathcal{Y}} \triangleq \frac{\mathcal{Y} D\mathcal{X} \circ \mathcal{Y}}{D\mathcal{Y}} \in \mathbb{R}^{m \times m}, \quad (64)$$

and using (20, 31) as above and (33),

$$\begin{aligned} \mathbf{J}_{\mathcal{X}}^{\mathcal{X} \circ \mathcal{Y}} &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{\text{Log}((\mathcal{X}\mathcal{Y})^{-1}(\mathcal{X} \text{Exp}(\boldsymbol{\tau})\mathcal{Y}))}{\boldsymbol{\tau}} \\ &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{\text{Log}(\mathcal{Y}^{-1} \text{Exp}(\boldsymbol{\tau})\mathcal{Y})}{\boldsymbol{\tau}} \\ &= \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{(\mathcal{Y}^{-1}\boldsymbol{\tau}^\wedge \mathcal{Y})^\vee}{\boldsymbol{\tau}} = \mathbf{Ad}_{\mathcal{Y}}^{-1} \quad (65) \end{aligned}$$

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{X} \circ \mathcal{Y}} = \dots = \mathbf{I} \quad (66)$$

3) *Jacobians of \mathcal{M} :* We define the *right Jacobian of \mathcal{M}* as the right Jacobian of $\mathcal{X} = \text{Exp}(\boldsymbol{\tau})$, *i.e.*, for $\boldsymbol{\tau} \in \mathbb{R}^m$,

$$\mathbf{J}_r(\boldsymbol{\tau}) \triangleq \frac{\boldsymbol{\tau} D \text{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}} \in \mathbb{R}^{m \times m}, \quad (67)$$

which is defined with (41a). The right Jacobian maps variations of the argument $\boldsymbol{\tau}$ into variations in the *local* tangent space at $\text{Exp}(\boldsymbol{\tau})$. From (41a) it is easy to prove that, for small $\delta\boldsymbol{\tau}$, the following approximations hold,

$$\text{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}) \approx \text{Exp}(\boldsymbol{\tau}) \text{Exp}(\mathbf{J}_r(\boldsymbol{\tau})\delta\boldsymbol{\tau}) \quad (68)$$

$$\text{Exp}(\boldsymbol{\tau}) \text{Exp}(\delta\boldsymbol{\tau}) \approx \text{Exp}(\boldsymbol{\tau} + \mathbf{J}_r^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}) \quad (69)$$

$$\text{Log}(\text{Exp}(\boldsymbol{\tau}) \text{Exp}(\delta\boldsymbol{\tau})) \approx \boldsymbol{\tau} + \mathbf{J}_r^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}. \quad (70)$$

Complementarily, the *left Jacobian of \mathcal{M}* is defined by,

$$\mathbf{J}_l(\boldsymbol{\tau}) \triangleq \frac{\varepsilon D \text{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}} \in \mathbb{R}^{m \times m}, \quad (71)$$

using the left Jacobian (44), leading to the approximations

$$\text{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}) \approx \text{Exp}(\mathbf{J}_l(\boldsymbol{\tau})\delta\boldsymbol{\tau}) \text{Exp}(\boldsymbol{\tau}) \quad (72)$$

$$\text{Exp}(\delta\boldsymbol{\tau}) \text{Exp}(\boldsymbol{\tau}) \approx \text{Exp}(\boldsymbol{\tau} + \mathbf{J}_l^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}) \quad (73)$$

$$\text{Log}(\text{Exp}(\delta\boldsymbol{\tau}) \text{Exp}(\boldsymbol{\tau})) \approx \boldsymbol{\tau} + \mathbf{J}_l^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}. \quad (74)$$

The left Jacobian maps variations of the argument $\boldsymbol{\tau}$ into variations in the *global* tangent space or Lie algebra. From (68, 72) we can relate left- and right- Jacobians with the adjoint,

$$\mathbf{Ad}_{\text{Exp}(\boldsymbol{\tau})} = \mathbf{J}_l(\boldsymbol{\tau}) \mathbf{J}_r^{-1}(\boldsymbol{\tau}). \quad (75)$$

Also, the chain rule allows us to relate \mathbf{J}_r and \mathbf{J}_l ,

$$\begin{aligned} \mathbf{J}_r(-\boldsymbol{\tau}) &\triangleq \mathbf{J}_{-\boldsymbol{\tau}}^{\text{Exp}(-\boldsymbol{\tau})} = \mathbf{J}_{\boldsymbol{\tau}}^{\text{Exp}(-\boldsymbol{\tau})} \mathbf{J}_{-\boldsymbol{\tau}}^{\boldsymbol{\tau}} = \mathbf{J}_{\boldsymbol{\tau}}^{\text{Exp}(\boldsymbol{\tau})^{-1}}(-\mathbf{I}) \\ &= -\mathbf{J}_{\text{Exp}(\boldsymbol{\tau})}^{\text{Exp}(\boldsymbol{\tau})^{-1}} \mathbf{J}_{\boldsymbol{\tau}}^{\text{Exp}(\boldsymbol{\tau})} = \mathbf{Ad}_{\text{Exp}(\boldsymbol{\tau})} \mathbf{J}_r(\boldsymbol{\tau}) \\ &= \mathbf{J}_l(\boldsymbol{\tau}). \quad (76) \end{aligned}$$

Closed forms of \mathbf{J}_r , \mathbf{J}_r^{-1} , \mathbf{J}_l and \mathbf{J}_l^{-1} exist for the typical manifolds in use. See the appendices for reference.

4) *Group action:* For $\mathcal{X} \in \mathcal{M}$ and $v \in \mathcal{V}$, we define with (41a)

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X} \cdot v} \triangleq \frac{\mathcal{X} D\mathcal{X} \cdot v}{D\mathcal{X}} \in \mathbb{R}^{m \times m} \quad (77)$$

$$\mathbf{J}_v^{\mathcal{X} \cdot v} \triangleq \frac{v D\mathcal{X} \cdot v}{Dv} . \quad (78)$$

Since group actions depend on the set \mathcal{V} , these expressions cannot be generalized. See the appendices for reference.

C. Useful, but deduced, Jacobian blocks

1) Log map: For $\tau = \text{Log}(\mathcal{X})$, and from (70),

$$\mathbf{J}_{\mathcal{X}}^{\text{Log}(\mathcal{X})} = \mathbf{J}_r^{-1}(\tau). \quad (79)$$

2) Plus and minus: We have

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X} \oplus \tau} = \mathbf{J}_{\mathcal{X}}^{\mathcal{X} \circ (\text{Exp}(\tau))} = \mathbf{Ad}_{\text{Exp}(\tau)}^{-1} \quad (80)$$

$$\mathbf{J}_{\tau}^{\mathcal{X} \oplus \tau} = \mathbf{J}_{\text{Exp}(\tau)}^{\mathcal{X} \circ (\text{Exp}(\tau))} \mathbf{J}_{\tau}^{\text{Exp}(\tau)} = \mathbf{J}_r(\tau) \quad (81)$$

and given $\mathcal{Z} = \mathcal{X}^{-1} \circ \mathcal{Y}$ and $\tau = \mathcal{Y} \ominus \mathcal{X} = \text{Log}(\mathcal{Z})$,

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{Y} \ominus \mathcal{X}} = \mathbf{J}_{\mathcal{Z}}^{\text{Log}(\mathcal{Z})} \mathbf{J}_{\mathcal{X}^{-1}}^{\mathcal{Z}} \mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} = -\mathbf{J}_l^{-1}(\tau) \quad (82)$$

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{Y} \ominus \mathcal{X}} = \mathbf{J}_{\mathcal{Z}}^{\text{Log}(\mathcal{Z})} \mathbf{J}_{\mathcal{Y}}^{\mathcal{Z}} = \mathbf{J}_r^{-1}(\tau). \quad (83)$$

where the former is proven here

$$\begin{aligned} \mathbf{J}_{\mathcal{X}}^{\mathcal{Y} \ominus \mathcal{X}} &= \mathbf{J}_{(\mathcal{X}^{-1} \circ \mathcal{Y})}^{\text{Log}(\mathcal{X}^{-1} \circ \mathcal{Y})} \mathbf{J}_{\mathcal{X}^{-1}}^{(\mathcal{X}^{-1} \circ \mathcal{Y})} \mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} \\ (79, 65, 62) &= \mathbf{J}_r^{-1}(\tau) \mathbf{Ad}_{\mathcal{Y}}^{-1} (-\mathbf{Ad}_{\mathcal{X}}) \\ (33, 34) &= -\mathbf{J}_r^{-1}(\tau) \mathbf{Ad}_{\mathcal{Y}^{-1} \circ \mathcal{X}} \\ &= -\mathbf{J}_r^{-1}(\tau) \mathbf{Ad}_{\text{Exp}(\tau)}^{-1} \\ (75) &= -\mathbf{J}_l^{-1}(\tau). \end{aligned}$$

IV. COMPOSITE MANIFOLDS

At the price of losing some consistency with the Lie theory, but at the benefit of obtaining some advantages in notation and manipulation, one can consider large and heterogeneous states as manifold composites (or bundles).

A *composite manifold* $\mathcal{M} = \langle \mathcal{M}_1, \dots, \mathcal{M}_M \rangle$ is no less than the concatenation of M non-interacting manifolds. This stems from defining identity, inverse and composition acting on each block of the composite separately,

$$\mathcal{E}_{\diamond} \triangleq \begin{bmatrix} \mathcal{E}_1 \\ \vdots \\ \mathcal{E}_M \end{bmatrix}, \quad \mathcal{X}^{\diamond} \triangleq \begin{bmatrix} \mathcal{X}^{-1} \\ \vdots \\ \mathcal{X}_M^{-1} \end{bmatrix}, \quad \mathcal{X} \diamond \mathcal{Y} \triangleq \begin{bmatrix} \mathcal{X} \circ \mathcal{Y}_1 \\ \vdots \\ \mathcal{X}_M \circ \mathcal{Y}_M \end{bmatrix}, \quad (84)$$

thereby fulfilling the group axioms, as well as a non-interacting retraction map, which we will also note as “exponential map” for the sake of unifying notations (notice the angled brackets),

$$\text{Exp}(\tau) \triangleq \begin{bmatrix} \text{Exp}(\tau_1) \\ \vdots \\ \text{Exp}(\tau_M) \end{bmatrix}, \quad \text{Log}(\mathcal{X}) \triangleq \begin{bmatrix} \text{Log}(\mathcal{X}) \\ \vdots \\ \text{Log}(\mathcal{X}_M) \end{bmatrix}, \quad (85)$$

thereby ensuring smoothness. These yield the composite's right- plus and minus (notice the diamond symbols),

$$\mathcal{X} \diamond \tau \triangleq \mathcal{X} \diamond \text{Exp}(\tau) \quad (86)$$

$$\mathcal{Y} \ominus \mathcal{X} \triangleq \text{Log}(\mathcal{X}^{\diamond} \diamond \mathcal{Y}). \quad (87)$$

The key consequence of these considerations (see Ex. 7) is that new derivatives can be defined,⁶ using \diamond and \ominus ,

$$\frac{Df(\mathcal{X})}{D\mathcal{X}} \triangleq \lim_{\tau \rightarrow 0} \frac{f(\mathcal{X} \diamond \tau) \ominus f(\mathcal{X})}{\tau}. \quad (88)$$

⁶We assume here right derivatives, but the same applies to left derivatives.

Example 7: $SE(n)$ vs. $T(n) \times SO(n)$ vs. $\langle \mathbb{R}^n, SO(n) \rangle$

We consider the space of translations $\mathbf{t} \in \mathbb{R}^n$ and rotations $\mathbf{R} \in SO(n)$. We have for this the well-known $SE(n)$ manifold of rigid motions $\mathbf{M} = [\mathbf{R} \ \mathbf{t}]$ (see Apps. C and D), which can also be constructed as $T(n) \times SO(n)$ (see Apps. A, B and E). These two are very similar, but have different tangent parametrizations: while $SE(n)$ uses $\tau = (\theta, \rho)$ with $\mathbf{M} = \exp(\tau^\wedge)$, $T(n) \times SO(n)$ uses $\tau = (\theta, \mathbf{p})$ with $\mathbf{M} = \exp(\mathbf{p}^\wedge) \exp(\theta^\wedge)$. They share the rotational part θ , but clearly $\rho \neq \mathbf{p}$ (see [11, pag. 35] for further details). In short, $SE(n)$ performs translation and rotation simultaneously as a continuum, while $T(n) \times SO(n)$ performs chained translation+rotation. In radical contrast, in the composite $\langle \mathbb{R}^n, SO(n) \rangle$ rotations and translations do not interact at all. By combining composition with $\text{Exp}()$ we obtain the (right) plus operators,

$$\begin{aligned} SE(n) : \quad \mathbf{M} \oplus \tau &= \begin{bmatrix} \mathbf{R} \text{Exp}(\theta) & \mathbf{t} + \mathbf{R}\mathbf{V}(\theta)\rho \\ 0 & 1 \end{bmatrix} \\ T(n) \times SO(n) : \quad \mathbf{M} \oplus \tau &= \begin{bmatrix} \mathbf{R} \text{Exp}(\theta) & \mathbf{t} + \mathbf{R}\mathbf{p} \\ 0 & 1 \end{bmatrix} \\ \langle \mathbb{R}^n, SO(n) \rangle : \quad \mathbf{M} \diamond \tau &= \begin{bmatrix} \mathbf{t} + \mathbf{p} \\ \mathbf{R} \text{Exp}(\theta) \end{bmatrix} \end{aligned}$$

where either \oplus may be used for the system dynamics, e.g. motion integration, but usually not \diamond , which might however be used to model perturbations. Their respective minus operators read,

$$\begin{aligned} SE(n) : \quad \mathbf{M}_2 \ominus \mathbf{M}_1 &= \begin{bmatrix} \mathbf{V}_1^{-1} \mathbf{R}_1^\top (\mathbf{p}_2 - \mathbf{p}_1) \\ \text{Log}(\mathbf{R}_1^\top \mathbf{R}_2) \end{bmatrix} \\ T(n) \times SO(n) : \quad \mathbf{M}_2 \ominus \mathbf{M}_1 &= \begin{bmatrix} \mathbf{R}_1^\top (\mathbf{p}_2 - \mathbf{p}_1) \\ \text{Log}(\mathbf{R}_1^\top \mathbf{R}_2) \end{bmatrix} \\ \langle \mathbb{R}^n, SO(n) \rangle : \quad \mathbf{M}_2 \ominus \mathbf{M}_1 &= \begin{bmatrix} \mathbf{p}_2 - \mathbf{p}_1 \\ \text{Log}(\mathbf{R}_1^\top \mathbf{R}_2) \end{bmatrix}, \end{aligned}$$

where now, interestingly, \ominus can be used to evaluate errors and uncertainty. This makes \diamond , \ominus valuable operators for computing derivatives and covariances.

With this derivative, Jacobians of functions $f : \mathcal{M} \rightarrow \mathcal{N}$ acting on composite manifolds can be determined in a per-block basis, which yields simple expressions requiring only knowledge on the manifold blocks of the composite,

$$\frac{Df(\mathcal{X})}{D\mathcal{X}} = \begin{bmatrix} \frac{Df_1}{D\mathcal{X}_1} & \cdots & \frac{Df_1}{D\mathcal{X}_M} \\ \vdots & \ddots & \vdots \\ \frac{Df_N}{D\mathcal{X}_1} & \cdots & \frac{Df_N}{D\mathcal{X}_M} \end{bmatrix}, \quad (89)$$

where $\frac{Df_i}{D\mathcal{X}_j}$ are each computed with (41a). For small values of τ the following holds,

$$f(\mathcal{X} \diamond \tau) \xrightarrow{\tau \rightarrow 0} f(\mathcal{X}) \diamond \frac{Df(\mathcal{X})}{D\mathcal{X}} \tau \in \mathcal{N}. \quad (90)$$

When using these derivatives, covariances and uncertainty propagation must follow the convention. In particular, the

covariance matrix (52) becomes

$$\Sigma_{\mathcal{X}} \triangleq \mathbb{E}[(\mathcal{X} \diamond \bar{\mathcal{X}})(\mathcal{X} \diamond \bar{\mathcal{X}})^T] \in \mathbb{R}^{n \times n}, \quad (91)$$

for which the linearized propagation (55) using (88) applies.

V. LANDMARK-BASED LOCALIZATION AND MAPPING

We provide three applicative examples of the theory for robot localization and mapping. The first one is a Kalman filter for landmark-based localization. The second one is a graph-based smoothing method for simultaneous localization and mapping. The third one adds sensor self-calibration. They are based on a common setup, explained as follows.

We consider a robot in the plane (see Section V-D for the 3D case) surrounded by a small number of punctual landmarks or *beacons*. The robot receives control actions in the form of axial and angular velocities and is able to measure the location of the beacons with respect to its own reference frame.

The robot pose is in $SE(2)$ (App. C) and the beacon positions in \mathbb{R}^2 (App. E),

$$\mathcal{X} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in SE(2), \quad \mathbf{b}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix} \in \mathbb{R}^2.$$

The control signal \mathbf{u} is a twist in $\mathfrak{se}(2)$ comprising longitudinal velocity v and angular velocity ω , with no lateral velocity component, integrated over the sampling time δt . The control is corrupted by additive Gaussian noise $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{W})$. This noise accounts for possible lateral wheel slippages u_s through a value of $\sigma_s \neq 0$,

$$\mathbf{u} = \begin{bmatrix} u_v \\ u_s \\ u_\omega \end{bmatrix} = \begin{bmatrix} v \delta t \\ 0 \\ \omega \delta t \end{bmatrix} + \mathbf{w} \in \mathfrak{se}(2) \quad (92)$$

$$\mathbf{W} = \begin{bmatrix} \sigma_v^2 \delta t & 0 & 0 \\ 0 & \sigma_s^2 \delta t & 0 \\ 0 & 0 & \sigma_\omega^2 \delta t \end{bmatrix} \in \mathbb{R}^{3 \times 3}. \quad (93)$$

At the arrival of a control \mathbf{u}_j at time j , the robot pose is updated with (56),

$$\mathcal{X}_j = \mathcal{X}_i \oplus \mathbf{u}_j \triangleq \mathcal{X}_i \text{Exp}(\mathbf{u}_j). \quad (94)$$

Landmark measurements are of the range and bearing type, though they are put in Cartesian form for simplicity. Their noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{N})$ is zero mean Gaussian,

$$\mathbf{y}_k = \mathcal{X}^{-1} \cdot \mathbf{b}_k + \mathbf{n} = \mathbf{R}^\top (\mathbf{b}_k - \mathbf{t}) + \mathbf{n} \in \mathbb{R}^2 \quad (95)$$

$$\mathbf{N} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \in \mathbb{R}^{2 \times 2}, \quad (96)$$

where we notice the rigid motion action $\mathcal{X}^{-1} \cdot \mathbf{b}_k$ (see App. C).

A. Localization with error-state Kalman filter on manifold

We initially consider the beacons \mathbf{b}_k situated at known positions. We define the pose to estimate as $\hat{\mathcal{X}} \in SE(2)$. The estimation error $\delta \mathbf{x}$ and its covariance \mathbf{P} are expressed in the tangent space at $\hat{\mathcal{X}}$ with (51, 52),

$$\delta \mathbf{x} \triangleq \mathcal{X} \ominus \hat{\mathcal{X}} \in \mathbb{R}^3 \quad (97)$$

$$\mathbf{P} \triangleq \mathbb{E}[(\mathcal{X} \ominus \hat{\mathcal{X}})(\mathcal{X} \ominus \hat{\mathcal{X}})^T] \in \mathbb{R}^{3 \times 3}. \quad (98)$$

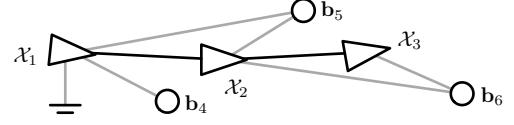


Figure 12. SAM factor graph with 3 poses and 3 beacons. Each measurement contributes a factor in the graph. There are 2 motion factors (black) and 5 beacon factors (gray). A prior factor on \mathcal{X}_1 provides global observability.

At each robot motion we apply ESKF prediction,

$$\hat{\mathcal{X}}_j = \hat{\mathcal{X}}_i \oplus \mathbf{u}_j \quad (99)$$

$$\mathbf{P}_j = \mathbf{F} \mathbf{P}_i \mathbf{F}^\top + \mathbf{G} \mathbf{W}_j \mathbf{G}^\top, \quad (100)$$

with the Jacobians computed from the blocks in App. C,

$$\mathbf{J} \triangleq \mathbf{J}_{\mathcal{X}_i}^{\mathcal{X}_j} = \mathbf{J}_{\hat{\mathcal{X}}_i}^{\hat{\mathcal{X}}_j \oplus \mathbf{u}_j} = \mathbf{Ad}_{\text{Exp}(\mathbf{u}_j)}^{-1}$$

$$\mathbf{G} \triangleq \mathbf{J}_{\mathbf{u}_j}^{\mathcal{X}_j} = \mathbf{J}_{\mathbf{u}_j}^{\hat{\mathcal{X}}_j \oplus \mathbf{u}_j} = \mathbf{J}_r(\mathbf{u}_j).$$

At each beacon measurement \mathbf{y}_k we apply ESKF correction,

$$\text{Innovation : } \mathbf{z} = \mathbf{y}_k - \hat{\mathcal{X}}^{-1} \cdot \mathbf{b}_k$$

$$\text{Innovation cov. : } \mathbf{Z} = \mathbf{H} \mathbf{P} \mathbf{H}^\top + \mathbf{N}$$

$$\text{Kalman gain : } \mathbf{K} = \mathbf{P} \mathbf{H}^\top \mathbf{Z}^{-1}$$

$$\text{Observed error : } \delta \mathbf{x} = \mathbf{K} \mathbf{z}$$

$$\text{State update : } \hat{\mathcal{X}} \leftarrow \hat{\mathcal{X}} \oplus \delta \mathbf{x} \quad (101)$$

$$\text{Cov. update : } \mathbf{P} \leftarrow \mathbf{P} - \mathbf{K} \mathbf{Z} \mathbf{K}^\top, \quad (102)$$

with the Jacobian computed from the blocks in App. C,

$$\begin{aligned} \mathbf{H} &\triangleq \mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1} \cdot \mathbf{b}_k} = \mathbf{J}_{\mathcal{X}^{-1}}^{\mathcal{X}^{-1} \cdot \mathbf{b}_k} \mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} \\ &= [\mathbf{R}^\top \quad \mathbf{R}^\top [1]_{\times} \mathbf{b}_k] \begin{bmatrix} -\mathbf{R} & [1]_{\times} \mathbf{t} \\ \mathbf{0} & -1 \end{bmatrix} \\ &= -[\mathbf{I} \quad \mathbf{R}^\top [1]_{\times} (\mathbf{b}_k - \mathbf{t})]. \end{aligned}$$

Notice that the only changes with respect to a regular EKF are in (99) and (101), where regular $+$ are substituted by \oplus . The Jacobians on the contrary are all computed using the Lie theory (see App. C). Interestingly, their usage is the same as in standard EKF — see e.g. the equation of the Kalman gain, which is the standard $\mathbf{K} = \mathbf{P} \mathbf{H}^\top (\mathbf{H} \mathbf{P} \mathbf{H}^\top + \mathbf{N})^{-1}$.

B. Smoothing and Mapping with graph-based optimization

We consider now the problem of smoothing and mapping (SAM), where the variables to estimate are the beacons' locations and the robot's trajectory. The solver of choice is a graph-based iterative least-squares optimizer. For simplicity, we assume the trajectory comprised of three robot poses $\{\mathcal{X}_1 \dots \mathcal{X}_3\}$, and a world with three beacons $\{\mathbf{b}_4 \dots \mathbf{b}_6\}$. The problem state is the composite

$$\mathcal{X} = \langle \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \mathbf{b}_4, \mathbf{b}_5, \mathbf{b}_6 \rangle, \quad \mathcal{X}_i \in SE(2), \quad \mathbf{b}_k \in \mathbb{R}^2. \quad (103)$$

The resulting factor graph [12] is shown in Fig. 12. Each prior or measurement contributes a factor in the graph. Motion measurements from pose i to j are derived from (94), while measurements of beacon k from pose i respond to (95),

$$\mathbf{u}_{ij} = \mathcal{X}_j \ominus \mathcal{X}_i + \mathbf{w}_{ij} = \text{Log}(\mathcal{X}_i^{-1} \mathcal{X}_j) + \mathbf{w}_{ij} \quad (104)$$

$$\mathbf{y}_{ik} = \mathcal{X}_i^{-1} \cdot \mathbf{b}_k + \mathbf{n}_{ik}. \quad (105)$$

Each factor comes with an information matrix, $\Omega_1 \triangleq \mathbf{W}_1^{-1}$, $\Omega_{ij} \triangleq \mathbf{W}_{ij}^{-1}$ and $\Omega_{ik} \triangleq \mathbf{N}_{ik}^{-1}$. The expectation residuals are,

$$\text{prior residual : } \mathbf{r}_1(\mathcal{X}) = \Omega_1^{\top/2} (\mathcal{X}_1 \ominus \hat{\mathcal{X}}_1)$$

$$\text{motion residual : } \mathbf{r}_{ij}(\mathcal{X}) = \Omega_{ij}^{\top/2} (\mathbf{u}_{ij} - (\hat{\mathcal{X}}_j \ominus \hat{\mathcal{X}}_i))$$

$$\text{beacon residual : } \mathbf{r}_{ik}(\mathcal{X}) = \Omega_{ik}^{\top/2} (\mathbf{y}_{ik} - \hat{\mathcal{X}}_i^{-1} \cdot \hat{\mathbf{b}}_k).$$

The optimum update step $\delta\mathbf{x}$ stems from minimizing

$$\delta\mathbf{x}^* = \arg \min_{\delta\mathbf{x}} \sum_{p \in \mathcal{P}} \mathbf{r}_p(\mathcal{X} \oplus \delta\mathbf{x})^\top \mathbf{r}_p(\mathcal{X} \oplus \delta\mathbf{x}) \quad (106)$$

with $\mathcal{P} = \{1, 12, 23, 14, 15, 25, 26, 36\}$ the set of node pairs of each measurement (see Fig. 12). The problem is solved iteratively as follows. Each residual in the sum (106) is linearized to $\mathbf{r}_p(\mathcal{X} \oplus \delta\mathbf{x}) \approx \mathbf{r}_p(\mathcal{X}) \oplus \mathbf{J}_{\mathcal{X}}^{\mathbf{r}_p} \delta\mathbf{x}$ following (90), where $\mathbf{J}_{\mathcal{X}}^{\mathbf{r}_p}$ are sparse Jacobians. The non-zero blocks of these Jacobians, that is $\mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_1}$, $\mathbf{J}_{\mathcal{X}_i}^{\mathbf{r}_{ij}}$, $\mathbf{J}_{\mathcal{X}_i}^{\mathbf{r}_{ij}}$, $\mathbf{J}_{\mathcal{X}_i}^{\mathbf{r}_{ik}}$ and $\mathbf{J}_{\mathbf{b}_k}^{\mathbf{r}_{ik}}$, can be easily computed following the methods in Section V-A, and noticing that by definition $\mathbf{J}_{\delta\mathbf{x}}^{f(\mathcal{X} \oplus \delta\mathbf{x})}|_{\delta\mathbf{x}=0} = \mathbf{J}_{\mathcal{X}}^{f(\mathcal{X} \oplus \delta\mathbf{x})}|_{\delta\mathbf{x}=0} = \mathbf{J}_{\mathcal{X}}^{f(\mathcal{X})}$. Building the total Jacobian matrix and residual vector,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{12}} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{12}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{23}} & \mathbf{J}_{\mathcal{X}_3}^{\mathbf{r}_{23}} & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{14}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_4}^{\mathbf{r}_{14}} & \mathbf{0} \\ \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{15}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_5}^{\mathbf{r}_{15}} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{25}} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_5}^{\mathbf{r}_{25}} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{26}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_6}^{\mathbf{r}_{26}} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathcal{X}_3}^{\mathbf{r}_{36}} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{r} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_{12} \\ \mathbf{r}_{23} \\ \mathbf{r}_{14} \\ \mathbf{r}_{15} \\ \mathbf{r}_{25} \\ \mathbf{r}_{26} \\ \mathbf{r}_{36} \end{bmatrix} \quad (107)$$

the linearized (106) is now transformed [12] to minimizing

$$\delta\mathbf{x}^* = \arg \min_{\delta\mathbf{x}} \|\mathbf{r} + \mathbf{J}\delta\mathbf{x}\|^2. \quad (108)$$

This is solved via least-squares using the pseudoinverse of \mathbf{J} (for large problems, QR [12], [13] or Cholesky [14], [15] factorizations are required),

$$\delta\mathbf{x}^* = -(\mathbf{J}^\top \mathbf{J})^{-1} \mathbf{J}^\top \mathbf{r}, \quad (109)$$

yielding the optimal step $\delta\mathbf{x}^*$ used to update the state,

$$\mathcal{X} \leftarrow \mathcal{X} \oplus \delta\mathbf{x}^*. \quad (110)$$

The procedure is iterated until convergence.

We highlight here the use of the composite notation in (103), which allows block-wise definitions of the Jacobian (107) and the update (110). We also remark the use of the $SE(2)$ manifold in the motion and measurement models, as we did in the ESKF case in Section V-A.

C. Smoothing and mapping with self-calibration

We consider the same problem as above but with a motion sensor affected by an unknown calibration bias $\mathbf{c} = (c_v, c_\omega)^\top$,

so that the control is now $\tilde{\mathbf{u}} = (v\delta t + c_v, 0, \omega\delta t + c_\omega)^\top + \mathbf{w}$. We define the bias correction function $c()$,

$$\mathbf{u} = c(\tilde{\mathbf{u}}, \mathbf{c}) \triangleq \begin{bmatrix} \tilde{u}_v - c_v \\ \tilde{u}_s \\ \tilde{u}_\omega - c_\omega \end{bmatrix} \in \mathbb{R}^3 \cong \mathfrak{se}(2). \quad (111)$$

The state composite is augmented with the unknowns \mathbf{c} ,

$$\mathcal{X} = \langle \mathbf{c}, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \mathbf{b}_4, \mathbf{b}_5, \mathbf{b}_6 \rangle, \quad \mathbf{c} \in \mathbb{R}^2, \quad \mathcal{X}_i \in SE(2), \quad \mathbf{b}_k \in \mathbb{R}^2,$$

and the motion residual becomes

$$\mathbf{r}_{ij}(\mathcal{X}) = \Omega_{ij}^{\top/2} (c(\tilde{\mathbf{u}}_{ij}, \mathbf{c}) - (\hat{\mathcal{X}}_j \ominus \hat{\mathcal{X}}_i)).$$

The procedure is as in Section V-B above, and just the total Jacobian is modified with an extra column on the left,

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\mathbf{c}}^{\mathbf{r}_{12}} & \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{12}} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{12}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\mathbf{c}}^{\mathbf{r}_{23}} & \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{23}} & \mathbf{J}_{\mathcal{X}_3}^{\mathbf{r}_{23}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{14}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_4}^{\mathbf{r}_{14}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\mathcal{X}_1}^{\mathbf{r}_{15}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_5}^{\mathbf{r}_{15}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{25}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_5}^{\mathbf{r}_{25}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathcal{X}_2}^{\mathbf{r}_{26}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_6}^{\mathbf{r}_{26}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathcal{X}_3}^{\mathbf{r}_{36}} & \mathbf{0} & \mathbf{0} & \mathbf{J}_{\mathbf{b}_6}^{\mathbf{r}_{36}} \end{bmatrix},$$

where $\mathbf{J}_{\mathbf{c}}^{\mathbf{r}_{ij}} = \Omega_{ij}^{\top/2} \mathbf{J}_{\mathbf{c}}^{c(\mathbf{u}_{ij}, \mathbf{c})}$, with $\mathbf{J}_{\mathbf{c}}^{c(\mathbf{u}_{ij}, \mathbf{c})}$ the 3×2 Jacobian of (111). The optimal solution is obtained with (109, 110). The resulting optimal state \mathcal{X} includes an optimal estimate of \mathbf{c} , that is, the self-calibration of the sensor bias.

D. 3D implementations

It is surprisingly easy to bring all the examples above to 3D. It suffices to define all variables in the correct spaces: $\mathcal{X} \in SE(3)$ and $\mathbf{u} \in \mathbb{R}^6 \cong \mathfrak{se}(3)$ (App. D), and $\{\mathbf{b}_k, \mathbf{y}\} \in \mathbb{R}^3$ (App. E). Jacobians and covariances matrices will follow with appropriate sizes. The interest here is in realizing that all the math in the algorithms, that is from (97) onwards, is exactly the same for 2D and 3D: the abstraction level provided by the Lie theory has made this possible.

VI. CONCLUSION

We have presented the essential of Lie theory in a form that should be useful for an audience skilled in state estimation, with a focus on robotics applications. This we have done through several initiatives:

First, a selection of materials that avoids abstract mathematical concepts as much as possible. This helps to focus Lie theory to make its tools easier to understand and to use.

Second, we chose a didactical approach, with significant redundancy. The main text is generic and covers the abstract points of Lie theory. It is accompanied by boxed examples, which ground the abstract concepts to particular Lie groups, and plenty of figures with very verbose captions.

Third, we have promoted the usage of handy operators, such as the capitalized $\text{Exp}()$ and $\text{Log}()$ maps, and the plus and minus operators $\oplus, \ominus, \oplus, \ominus$. They allow us to work on the Cartesian representation of the tangent spaces, producing

formulas for derivatives and covariance handling that greatly resemble their counterparts in standard vector spaces.

Fourth, we have made special emphasis on the definition, geometrical interpretation, and computation of Jacobians. For this, we have introduced notations for the Jacobian matrices and covariances that allow a manipulation that is visually powerful. In particular, the chain rule is clearly visible with this notation. This helps to build intuition and reducing errors.

Fifth, we present in the appendices that follow an extensive compendium of formulas for the most common groups in robotics. In 2D, we present the rotation groups of unit complex numbers S^1 and rotation matrices $SO(2)$, and the rigid motion group $SE(2)$. In 3D, we present the groups of unit quaternions S^3 and rotation matrices $SO(3)$, both used for rotations, and the rigid motion group $SE(3)$. We also present the translation groups for any dimension, which can be implemented by either the standard vector space \mathbb{R}^n under addition, or by the matrix translation group $T(n)$ under multiplication.

Sixth, we have presented some applicative examples to illustrate the capacity of Lie theory to solve robotics problems with elegance and precision. The somewhat naive concept of composite group helps to unify heterogeneous state vectors into a Lie-theoretic form.

Finally, we accompany this text with the new C++ library manif [7] implementing the tools described here. manif can be found at <https://github.com/artivis/manif>. The applications in Section V are demonstrated in manif as examples.

Though we do not introduce any new theoretical material, we believe the form in which Lie theory is here exposed will help many researchers enter the field for their future developments. We also believe this alone represents a valuable contribution.

APPENDIX A

THE 2D ROTATION GROUPS S^1 AND $SO(2)$

The Lie group S^1 is the group of unit complex numbers under the complex product. Its topology is the unit circle, or the unit 1-sphere, and therefore the name S^1 . The group, Lie algebra and vector elements have the form,

$$\mathbf{z} = \cos \theta + i \sin \theta, \quad \tau^\wedge = i\theta, \quad \tau = \theta. \quad (112)$$

Inversion and composition are achieved by conjugation $\mathbf{z}^{-1} = \mathbf{z}^*$, and product $\mathbf{z}_a \circ \mathbf{z}_b = \mathbf{z}_a \mathbf{z}_b$.

The group $SO(2)$ is the group of special orthogonal matrices in the plane, or rotation matrices, under matrix multiplication. Group, Lie algebra and vector elements have the form,

$$\mathbf{R} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \tau^\wedge = [\theta]_\times \triangleq \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix}, \quad \tau = \theta. \quad (113)$$

Inversion and composition are achieved by transposition $\mathbf{R}^{-1} = \mathbf{R}^\top$, and product $\mathbf{R}_a \circ \mathbf{R}_b = \mathbf{R}_a \mathbf{R}_b$.

Both groups rotate 2-vectors, and they have isomorphic tangent spaces. We thus study them together.

A. Exp and Log maps

Exp and Log maps may be defined for complex numbers of S^1 and rotation matrices of $SO(2)$. For S^1 we have,

$$\mathbf{z} = \text{Exp}(\theta) = \cos \theta + i \sin \theta \quad \in \mathbb{C} \quad (114)$$

$$\theta = \text{Log}(\mathbf{z}) = \arctan(\text{Im}(\mathbf{z}), \text{Re}(\mathbf{z})) \quad \in \mathbb{R}, \quad (115)$$

where (114) is the Euler formula, whereas for $SO(2)$,

$$\mathbf{R} = \text{Exp}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \in \mathbb{R}^{2 \times 2} \quad (116)$$

$$\theta = \text{Log}(\mathbf{R}) = \arctan(r_{21}, r_{11}) \quad \in \mathbb{R}. \quad (117)$$

B. Inverse, composition, exponential map

We consider generic 2D rotation elements, and note them with the sans-serif font, Q, R. We have

$$\mathbf{R}(\theta)^{-1} = \mathbf{R}(-\theta) \quad (118)$$

$$\mathbf{Q} \circ \mathbf{R} = \mathbf{R} \circ \mathbf{Q}, \quad (119)$$

i.e., planar rotations are commutative. It follows that

$$\text{Exp}(\theta_1 + \theta_2) = \text{Exp}(\theta_1) \circ \text{Exp}(\theta_2) \quad (120)$$

$$\text{Log}(\mathbf{Q} \circ \mathbf{R}) = \text{Log}(\mathbf{Q}) + \text{Log}(\mathbf{R}) \quad (121)$$

$$\mathbf{Q} \ominus \mathbf{R} = \theta_Q - \theta_R. \quad (122)$$

C. Jacobian blocks

Since our defined derivatives map tangent vector spaces, and these spaces coincide for the planar rotation manifolds of S^1 and $SO(2)$, i.e., $\theta = \text{Log}(\mathbf{z}) = \text{Log}(\mathbf{R})$, it follows that the Jacobians are independent of the representation used (\mathbf{z} or \mathbf{R}).

1) *Adjoint and other trivial Jacobians:* From (41a), Section III-B and the properties above, the following scalar derivative blocks become trivial,

$$\mathbf{Ad}_{\mathbf{R}} = 1 \quad \in \mathbb{R} \quad (123)$$

$$\mathbf{J}_{\mathbf{R}}^{\mathbf{R}^{-1}} = -1 \quad \in \mathbb{R} \quad (124)$$

$$\mathbf{J}_{\mathbf{Q}}^{\mathbf{Q} \circ \mathbf{R}} = \mathbf{J}_{\mathbf{R}}^{\mathbf{Q} \circ \mathbf{R}} = 1 \quad \in \mathbb{R} \quad (125)$$

$$\mathbf{J}_r(\theta) = \mathbf{J}_l(\theta) = 1 \quad \in \mathbb{R} \quad (126)$$

$$\mathbf{J}_{\mathbf{R}}^{\mathbf{R} \oplus \theta} = \mathbf{J}_{\theta}^{\mathbf{R} \oplus \theta} = 1 \quad \in \mathbb{R} \quad (127)$$

$$\mathbf{J}_{\mathbf{Q}}^{\mathbf{Q} \ominus \mathbf{R}} = -\mathbf{J}_{\mathbf{R}}^{\mathbf{Q} \ominus \mathbf{R}} = 1 \quad \in \mathbb{R} \quad (128)$$

2) *Rotation action:* For the action $\mathbf{R} \cdot \mathbf{v}$ we have,

$$\begin{aligned} \mathbf{J}_{\mathbf{R}}^{\mathbf{R} \cdot \mathbf{v}} &= \lim_{\theta \rightarrow 0} \frac{\mathbf{R} \text{Exp}(\theta) \mathbf{v} - \mathbf{R} \mathbf{v}}{\theta} \\ &= \lim_{\theta \rightarrow 0} \frac{\mathbf{R}(\mathbf{I} + [\theta]_\times) \mathbf{v} - \mathbf{R} \mathbf{v}}{\theta} \\ &= \lim_{\theta \rightarrow 0} \frac{\mathbf{R}[\theta]_\times \mathbf{v}}{\theta} = \mathbf{R}[1]_\times \mathbf{v} \quad \in \mathbb{R}^{2 \times 1} \end{aligned} \quad (129)$$

and

$$\mathbf{J}_{\mathbf{v}}^{\mathbf{R} \cdot \mathbf{v}} = \frac{D\mathbf{R}\mathbf{v}}{D\mathbf{v}} = \mathbf{R} \quad \in \mathbb{R}^{2 \times 2}. \quad (130)$$

APPENDIX B

THE 3D ROTATION GROUPS S^3 AND $SO(3)$

The Lie group S^3 is the group of unit quaternions under quaternion multiplication. Its topology is the unit 3-sphere in \mathbb{R}^4 , and therefore its name S^3 . Quaternions (please consult [8] for an in-depth reference) may be represented by either of these equivalent forms,

$$\begin{aligned} \mathbf{q} &= w + ix + jy + kz = w + \mathbf{v} \quad \in \mathbb{H} \\ &= [w \quad x \quad y \quad z]^\top = \begin{bmatrix} w \\ \mathbf{v} \end{bmatrix} \quad \in \mathbb{H}, \end{aligned} \quad (131)$$

where $w, x, y, z \in \mathbb{R}$, and i, j, k are three unit imaginary numbers such that $i^2 = j^2 = k^2 = ijk = -1$. The scalar w is known as the scalar or real part, and $\mathbf{v} \in \mathbb{H}_p$ as the vector or imaginary part. We note \mathbb{H}_p the set of pure quaternions, *i.e.*, of null scalar part, with dimension 3. Inversion and composition are achieved by conjugation $\mathbf{q}^{-1} = \mathbf{q}^*$, where $\mathbf{q}^* \triangleq w - \mathbf{v}$ is the conjugate, and product $\mathbf{q}_a \circ \mathbf{q}_b = \mathbf{q}_a \mathbf{q}_b$.

The group $SO(3)$ is the group of special orthogonal matrices in 3D space, or rotation matrices, under matrix multiplication. Inversion and composition are achieved with transposition and product as in all groups $SO(n)$.

Both groups rotate 3-vectors. They have isomorphic tangent spaces whose elements are identifiable with rotation vectors in \mathbb{R}^3 , so we study them together. It is in this space \mathbb{R}^3 where we define the vectors of rotation rate $\boldsymbol{\omega} \triangleq \mathbf{u}\omega$, angle-axis $\boldsymbol{\theta} \triangleq \mathbf{u}\theta$, and all perturbations and uncertainties.

The quaternion manifold S^3 is a double cover of $SO(3)$, *i.e.*, \mathbf{q} and $-\mathbf{q}$ represent the same rotation \mathbf{R} . The first cover corresponds to quaternions with positive real part $w > 0$. The two groups can be considered isomorphic up to the first cover.

A. Exp and Log maps

The Exp and Log maps may be defined for quaternions of S^3 and rotation matrices of $SO(3)$. For quaternions $\mathbf{q} = (w, \mathbf{v}) \in \mathbb{H}$ we have (see Ex. 5),

$$\mathbf{q} = \text{Exp}(\theta\mathbf{u}) \triangleq \cos(\theta/2) + \mathbf{u}\sin(\theta/2) \in \mathbb{H} \quad (132)$$

$$\theta\mathbf{u} = \text{Log}(\mathbf{q}) \triangleq 2\mathbf{v} \frac{\arctan(\|\mathbf{v}\|, w)}{\|\mathbf{v}\|} \in \mathbb{R}^3. \quad (133)$$

We can avoid eventual problems due to the double cover of \mathbf{q} by ensuring that its scalar part w is positive before doing the Log. If it is not, we can substitute \mathbf{q} by $-\mathbf{q}$ before the Log.

For rotation matrices we have (see Ex. 4),

$$\mathbf{R} = \text{Exp}(\theta\mathbf{u}) \triangleq \mathbf{I} + \sin\theta [\mathbf{u}]_\times + (1 - \cos\theta) [\mathbf{u}]_\times^2 \in \mathbb{R}^{3 \times 3} \quad (134)$$

$$\theta\mathbf{u} = \text{Log}(\mathbf{R}) \triangleq \frac{\theta(\mathbf{R} - \mathbf{R}^\top)^\vee}{2\sin\theta} \in \mathbb{R}^3, \quad (135)$$

with $\theta = \cos^{-1}\left(\frac{\text{trace}(\mathbf{R}) - 1}{2}\right)$.

B. Rotation action

Given the expressions above for the quaternion and the rotation matrix, the rotation action of quaternions on 3-vectors is performed by the double quaternion product,

$$\mathbf{x}' = \mathbf{q} \mathbf{x} \mathbf{q}^* \quad (136)$$

while rotation matrices use a single matrix product,

$$\mathbf{x}' = \mathbf{R} \mathbf{x}. \quad (137)$$

Both correspond to a right-hand rotation of θ rad around the axis \mathbf{u} . Identifying in them \mathbf{x} and \mathbf{x}' yields the identity

$$\mathbf{R}(\mathbf{q}) = \begin{bmatrix} w^2 + x^2 - y^2 - z^2 & 2(xy - wz) & 2(xz + wy) \\ 2(xy + wz) & w^2 - x^2 + y^2 - z^2 & 2(yz - wx) \\ 2(xz - wy) & 2(yz + wx) & w^2 - x^2 - y^2 + z^2 \end{bmatrix} \quad (138)$$

C. Elementary Jacobian blocks

Since our defined derivatives map tangent vector spaces, and these spaces coincide for the 3D rotation manifolds of S^3 and $SO(3)$, *i.e.*, $\boldsymbol{\theta} = \text{Log}(\mathbf{q}) = \text{Log}(\mathbf{R})$, it follows that the Jacobians are independent of the representation used (\mathbf{q} or \mathbf{R}). We thus consider generic 3D rotation elements and note them with the sans-serif font \mathbf{R} .

1) *Adjoint*: We have from (31)

$$\mathbf{Ad}_{\mathbf{R}} \boldsymbol{\theta} = (\mathbf{R} [\boldsymbol{\theta}]_\times \mathbf{R}^\top)^\vee = (([\mathbf{R}\boldsymbol{\theta}]_\times)^\vee)^\vee = \mathbf{R}\boldsymbol{\theta}$$

therefore

$$\mathbf{Ad}_{\mathbf{R}} = \mathbf{R}, \quad (139)$$

which means, just to clarify it once again, that $\mathbf{Ad}_{\mathbf{q}} = \mathbf{R}(\mathbf{q})$, see (138), and $\mathbf{Ad}_{\mathbf{R}} = \mathbf{R}$.

2) *Inversion, composition*: We have from Section III-B,

$$\mathbf{J}_{\mathbf{R}}^{\mathbf{R}^{-1}} = -\mathbf{Ad}_{\mathbf{R}} = -\mathbf{R} \quad (140)$$

$$\mathbf{J}_{\mathbf{Q}}^{\mathbf{Q}\mathbf{R}} = \mathbf{Ad}_{\mathbf{R}}^{-1} = \mathbf{R}^\top \quad (141)$$

$$\mathbf{J}_{\mathbf{R}}^{\mathbf{Q}\mathbf{R}} = \mathbf{I}. \quad (142)$$

3) *Right and left Jacobians*: They admit the closed forms [11, pag. 40],

$$\mathbf{J}_r(\boldsymbol{\theta}) = \mathbf{I} - \frac{1 - \cos\theta}{\theta^2} [\boldsymbol{\theta}]_\times + \frac{\theta - \sin\theta}{\theta^3} [\boldsymbol{\theta}]_\times^2 \quad (143)$$

$$\mathbf{J}_r^{-1}(\boldsymbol{\theta}) = \mathbf{I} + \frac{1}{2} [\boldsymbol{\theta}]_\times + \left(\frac{1}{\theta^2} - \frac{1 + \cos\theta}{2\theta \sin\theta} \right) [\boldsymbol{\theta}]_\times^2 \quad (144)$$

$$\mathbf{J}_l(\boldsymbol{\theta}) = \mathbf{I} + \frac{1 - \cos\theta}{\theta^2} [\boldsymbol{\theta}]_\times + \frac{\theta - \sin\theta}{\theta^3} [\boldsymbol{\theta}]_\times^2 \quad (145)$$

$$\mathbf{J}_l^{-1}(\boldsymbol{\theta}) = \mathbf{I} - \frac{1}{2} [\boldsymbol{\theta}]_\times + \left(\frac{1}{\theta^2} - \frac{1 + \cos\theta}{2\theta \sin\theta} \right) [\boldsymbol{\theta}]_\times^2 \quad (146)$$

where we can observe that

$$\mathbf{J}_l = \mathbf{J}_r^\top, \quad \mathbf{J}_l^{-1} = \mathbf{J}_r^{-\top}. \quad (147)$$

4) *Right- plus and minus*: We have for $\boldsymbol{\theta} = \mathbf{Q} \ominus \mathbf{R}$,

$$\mathbf{J}_{\mathbf{R}}^{\mathbf{R} \oplus \boldsymbol{\theta}} = \mathbf{R}(\boldsymbol{\theta})^\top \quad \mathbf{J}_{\boldsymbol{\theta}}^{\mathbf{R} \oplus \boldsymbol{\theta}} = \mathbf{J}_r(\boldsymbol{\theta}) \quad (148)$$

$$\mathbf{J}_{\mathbf{Q}}^{\mathbf{Q} \ominus \mathbf{R}} = \mathbf{J}_r^{-1}(\boldsymbol{\theta}) \quad \mathbf{J}_{\mathbf{R}}^{\mathbf{Q} \ominus \mathbf{R}} = -\mathbf{J}_l^{-1}(\boldsymbol{\theta}) \quad (149)$$

5) *Rotation action*: We have

$$\begin{aligned} \mathbf{J}_{\mathbf{R}}^{\mathbf{R} \cdot \mathbf{v}} &\triangleq \lim_{\theta \rightarrow 0} \frac{(\mathbf{R} \oplus \boldsymbol{\theta})\mathbf{v} - \mathbf{R}\mathbf{v}}{\theta} = \\ \lim_{\theta \rightarrow 0} \frac{\mathbf{R} \text{Exp}(\theta)\mathbf{v} - \mathbf{R}\mathbf{v}}{\theta} &= \lim_{\theta \rightarrow 0} \frac{\mathbf{R}(\mathbf{I} + [\boldsymbol{\theta}]_\times)\mathbf{v} - \mathbf{R}\mathbf{v}}{\theta} \\ &= \lim_{\theta \rightarrow 0} \frac{\mathbf{R}[\boldsymbol{\theta}]_\times \mathbf{v}}{\theta} = \lim_{\theta \rightarrow 0} \frac{-\mathbf{R}[\mathbf{v}]_\times \theta}{\theta} = -\mathbf{R}[\mathbf{v}]_\times \end{aligned} \quad (150)$$

where we used the properties $\text{Exp}(\boldsymbol{\theta}) \approx \mathbf{I} + [\boldsymbol{\theta}]_\times$ and $[\mathbf{a}]_\times \mathbf{b} = -[\mathbf{b}]_\times \mathbf{a}$. The second Jacobian yields,

$$\mathbf{J}_{\mathbf{v}}^{\mathbf{R} \cdot \mathbf{v}} \triangleq \lim_{\partial \mathbf{v} \rightarrow 0} \frac{\mathbf{R}(\mathbf{v} + \partial \mathbf{v}) - \mathbf{R}\mathbf{v}}{\partial \mathbf{v}} = \mathbf{R}. \quad (151)$$

APPENDIX C THE 2D RIGID MOTION GROUP $SE(2)$

We write elements of the rigid motion group $SE(2)$ as

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in SE(2) \subset \mathbb{R}^{3 \times 3}, \quad (152)$$

with $\mathbf{R} \in SO(2)$ a rotation and $\mathbf{t} \in \mathbb{R}^2$ a translation. The Lie algebra and vector tangents are formed by elements of the type

$$\boldsymbol{\tau}^\wedge = \begin{bmatrix} [\theta]_\times & \boldsymbol{\rho} \\ \mathbf{0} & 0 \end{bmatrix} \in \mathfrak{se}(2), \quad \boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{\rho} \\ \theta \end{bmatrix} \in \mathbb{R}^3. \quad (153)$$

A. Inverse, composition

Inversion and composition are performed respectively with matrix inversion and product,

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \quad (154)$$

$$\mathbf{M}_a \mathbf{M}_b = \begin{bmatrix} \mathbf{R}_a \mathbf{R}_b & \mathbf{t}_a + \mathbf{R}_a \mathbf{t}_b \\ \mathbf{0} & 1 \end{bmatrix}. \quad (155)$$

B. Exp and Log maps

Exp and Log are implemented via exponential maps directly from the scalar tangent space $\mathbb{R}^3 \cong \mathfrak{se}(2) = TSE(2)$ — see [5] for the derivation,

$$\mathbf{M} = \text{Exp}(\boldsymbol{\tau}) \triangleq \begin{bmatrix} \text{Exp}(\theta) & \mathbf{V}(\theta) \boldsymbol{\rho} \\ \mathbf{0} & 1 \end{bmatrix} \quad (156)$$

$$\boldsymbol{\tau} = \text{Log}(\mathbf{M}) \triangleq \begin{bmatrix} \mathbf{V}^{-1}(\theta) \mathbf{t} \\ \text{Log}(\mathbf{R}) \end{bmatrix}. \quad (157)$$

with

$$\mathbf{V}(\theta) = \frac{\sin \theta}{\theta} \mathbf{I} + \frac{1 - \cos \theta}{\theta} [1]_\times. \quad (158)$$

C. Jacobian blocks

1) *Adjoint*: The adjoint is easily found from (31) using the fact that planar rotations commute,

$$\text{Ad}_{\mathbf{M}} \boldsymbol{\tau} = (\mathbf{M} \boldsymbol{\tau}^\wedge \mathbf{M}^{-1})^\vee = \begin{bmatrix} \mathbf{R} \boldsymbol{\rho} - [\theta]_\times \mathbf{t} \\ \theta \end{bmatrix} = \text{Ad}_{\mathbf{M}} \begin{bmatrix} \boldsymbol{\rho} \\ \theta \end{bmatrix},$$

leading to

$$\text{Ad}_{\mathbf{M}} = \begin{bmatrix} \mathbf{R} & -[1]_\times \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}. \quad (159)$$

2) *Inversion, composition*: We have from Section III-B,

$$\mathbf{J}_{\mathbf{M}}^{\mathbf{M}^{-1}} = -\text{Ad}_{\mathbf{M}} = \begin{bmatrix} -\mathbf{R} & [1]_\times \mathbf{t} \\ \mathbf{0} & -1 \end{bmatrix} \quad (160)$$

$$\mathbf{J}_{\mathbf{M}_a \mathbf{M}_b}^{\mathbf{M}_a \mathbf{M}_b} = \text{Ad}_{\mathbf{M}_b}^{-1} = \begin{bmatrix} \mathbf{R}_b^\top & \mathbf{R}_b^\top [1]_\times \mathbf{t}_b \\ \mathbf{0} & 1 \end{bmatrix} \quad (161)$$

$$\mathbf{J}_{\mathbf{M}_b}^{\mathbf{M}_a \mathbf{M}_b} = \mathbf{I}. \quad (162)$$

3) *Right and left Jacobians*: We have from [11, pag. 36],

$$\mathbf{J}_r = \begin{bmatrix} \sin \theta / \theta & (1 - \cos \theta) / \theta & (\theta \rho_1 - \rho_2 + \rho_2 \cos \theta - \rho_1 \sin \theta) / \theta^2 \\ (\cos \theta - 1) / \theta & \sin \theta / \theta & (\rho_1 + \theta \rho_2 - \rho_1 \cos \theta - \rho_2 \sin \theta) / \theta^2 \\ 0 & 0 & 1 \end{bmatrix} \quad (163)$$

$$\mathbf{J}_l = \begin{bmatrix} \sin \theta / \theta & (\cos \theta - 1) / \theta & (\theta \rho_1 + \rho_2 - \rho_2 \cos \theta - \rho_1 \sin \theta) / \theta^2 \\ (1 - \cos \theta) / \theta & \sin \theta / \theta & (-\rho_1 + \theta \rho_2 + \rho_1 \cos \theta - \rho_2 \sin \theta) / \theta^2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (164)$$

4) *Rigid motion action*: We have the action on points \mathbf{p} ,

$$\mathbf{M} \cdot \mathbf{p} \triangleq \mathbf{t} + \mathbf{R} \mathbf{p}, \quad (165)$$

therefore and since for $\boldsymbol{\tau} \rightarrow 0$ we have $\text{Exp}(\boldsymbol{\tau}) \rightarrow \mathbf{I} + \boldsymbol{\tau}^\wedge$,

$$\mathbf{J}_{\mathbf{M}}^{\mathbf{M} \cdot \mathbf{p}} = \lim_{\boldsymbol{\tau} \rightarrow 0} \frac{\mathbf{M} \text{Exp}(\boldsymbol{\tau}) \cdot \mathbf{p} - \mathbf{M} \cdot \mathbf{p}}{\boldsymbol{\tau}} = \begin{bmatrix} \mathbf{R} & \mathbf{R} [1]_\times \mathbf{p} \end{bmatrix} \quad (166)$$

$$\mathbf{J}_{\mathbf{p}}^{\mathbf{M} \cdot \mathbf{p}} = \mathbf{R}. \quad (167)$$

APPENDIX D THE 3D RIGID MOTION GROUP $SE(3)$

We write elements of the 3D rigid motion group $SE(3)$ as

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in SE(3) \subset \mathbb{R}^{4 \times 4}, \quad (168)$$

with $\mathbf{R} \in SO(3)$ a rotation matrix and $\mathbf{t} \in \mathbb{R}^3$ a translation vector. The Lie algebra and vector tangents are formed by elements of the type

$$\boldsymbol{\tau}^\wedge = \begin{bmatrix} [\theta]_\times & \boldsymbol{\rho} \\ \mathbf{0} & 0 \end{bmatrix} \in \mathfrak{se}(3), \quad \boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{\rho} \\ \theta \end{bmatrix} \in \mathbb{R}^6. \quad (169)$$

A. Inverse, composition

Inversion and composition are performed respectively with matrix inversion and product,

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \quad (170)$$

$$\mathbf{M}_a \mathbf{M}_b = \begin{bmatrix} \mathbf{R}_a \mathbf{R}_b & \mathbf{t}_a + \mathbf{R}_a \mathbf{t}_b \\ \mathbf{0} & 1 \end{bmatrix}. \quad (171)$$

B. Exp and Log maps

Exp and Log are implemented via exponential maps directly from the vector tangent space $\mathbb{R}^6 \cong \mathfrak{se}(3) = TSE(3)$ — see [5] for the derivation,

$$\mathbf{M} = \text{Exp}(\boldsymbol{\tau}) \triangleq \begin{bmatrix} \text{Exp}(\theta) & \mathbf{V}(\theta) \boldsymbol{\rho} \\ \mathbf{0} & 1 \end{bmatrix} \quad (172)$$

$$\boldsymbol{\tau} = \text{Log}(\mathbf{M}) \triangleq \begin{bmatrix} \mathbf{V}^{-1}(\theta) \mathbf{t} \\ \text{Log}(\mathbf{R}) \end{bmatrix}. \quad (173)$$

with (recall for $\text{Log}(\mathbf{M})$ that $\boldsymbol{\theta} = \theta \mathbf{u} = \text{Log}(\mathbf{R})$)

$$\mathbf{V}(\theta) = \mathbf{I} + \frac{1 - \cos \theta}{\theta^2} [\theta]_\times + \frac{\theta - \sin \theta}{\theta^3} [\theta]_\times^2 \quad (174)$$

which, notice, matches (145) exactly.

C. Jacobian blocks

1) *Adjoint*: We have (see Ex. 6),

$$\text{Ad}_{\mathbf{M}} \boldsymbol{\tau} = (\mathbf{M} \boldsymbol{\tau}^\wedge \mathbf{M}^{-1})^\vee = \begin{bmatrix} \mathbf{R} \boldsymbol{\rho} + [\mathbf{t}]_\times \mathbf{R} \boldsymbol{\theta} \\ \mathbf{R} \boldsymbol{\theta} \end{bmatrix} = \text{Ad}_{\mathbf{M}} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}$$

therefore,

$$\text{Ad}_{\mathbf{M}} = \begin{bmatrix} \mathbf{R} & [\mathbf{t}]_\times \mathbf{R} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \in \mathbb{R}^{6 \times 6}. \quad (175)$$

2) *Inversion, composition:* We have from Section III-B,

$$\mathbf{J}_M^{M^{-1}} = - \begin{bmatrix} \mathbf{R} & [\mathbf{t}]_\times \mathbf{R} \\ 0 & \mathbf{R} \end{bmatrix} \quad (176)$$

$$\mathbf{J}_{M_a M_b}^{M_a M_b} = \begin{bmatrix} \mathbf{R}_b^\top & -\mathbf{R}_b^\top [\mathbf{t}_b]_\times \\ 0 & \mathbf{R}_b^\top \end{bmatrix} \quad (177)$$

$$\mathbf{J}_{M_b}^{M_a M_b} = \mathbf{I}_6. \quad (178)$$

3) *Right and left Jacobians:* Closed forms of the left Jacobian and its inverse are given by Barfoot in [10],

$$\mathbf{J}_l(\rho, \theta) = \begin{bmatrix} \mathbf{J}_l(\theta) & \mathbf{Q}(\rho, \theta) \\ 0 & \mathbf{J}_l(\theta) \end{bmatrix} \quad (179a)$$

$$\mathbf{J}_l^{-1}(\rho, \theta) = \begin{bmatrix} \mathbf{J}_l^{-1}(\theta) & -\mathbf{J}_l^{-1}(\theta) \mathbf{Q}(\rho, \theta) \mathbf{J}_l^{-1}(\theta) \\ 0 & \mathbf{J}_l^{-1}(\theta) \end{bmatrix} \quad (179b)$$

where $\mathbf{J}_l(\theta)$ is the left Jacobian of $SO(3)$, see (145), and

$$\begin{aligned} \mathbf{Q}(\rho, \theta) = & \frac{1}{2} \rho_\times + \frac{\theta - \sin \theta}{\theta^3} (\theta_\times \rho_\times + \rho_\times \theta_\times + \theta_\times \rho_\times \theta_\times) \\ & - \frac{1 - \frac{\theta^2}{2} - \cos \theta}{\theta^4} (\theta_\times^2 \rho_\times + \rho_\times \theta_\times^2 - 3 \theta_\times \rho_\times \theta_\times) \\ & - \frac{1}{2} \left(\frac{1 - \frac{\theta^2}{2} - \cos \theta}{\theta^4} - 3 \frac{\theta - \sin \theta - \frac{\theta^3}{6}}{\theta^5} \right) \\ & \times (\theta_\times \rho_\times \theta_\times^2 + \theta_\times^2 \rho_\times \theta_\times). \end{aligned} \quad (180)$$

The right Jacobian and its inverse are obtained using (76), that is, $\mathbf{J}_r(\rho, \theta) = \mathbf{J}_l(-\rho, -\theta)$ and $\mathbf{J}_r^{-1}(\rho, \theta) = \mathbf{J}_l^{-1}(-\rho, -\theta)$.

4) *Rigid motion action:* We have the action on points \mathbf{p} ,

$$\mathbf{M} \cdot \mathbf{p} \triangleq \mathbf{t} + \mathbf{R}\mathbf{p}, \quad (181)$$

therefore and since for $\tau \rightarrow 0$ we have $\text{Exp}(\tau) \rightarrow \mathbf{I} + \tau^\wedge$,

$$\mathbf{J}_M^{M \cdot \mathbf{p}} = \lim_{\tau \rightarrow 0} \frac{\mathbf{M} \text{Exp}(\tau) \cdot \mathbf{p} - \mathbf{M} \cdot \mathbf{p}}{\tau} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}[\mathbf{p}]_\times \end{bmatrix} \quad (182)$$

$$\mathbf{J}_{\mathbf{p}}^{M \cdot \mathbf{p}} = \mathbf{R}. \quad (183)$$

APPENDIX E

THE TRANSLATION GROUPS $(\mathbb{R}^n, +)$ AND $T(n)$

The group $(\mathbb{R}^n, +)$ is the group of vectors under addition and can be regarded as a translation group. We deem it *trivial* in the sense that the group elements, the Lie algebra, and the tangent spaces are all the same, so $\mathbf{t} = \mathbf{t}^\wedge = \text{Exp}(\mathbf{t})$. Its equivalent matrix group (under multiplication) is the translation group $T(n)$, whose group, Lie algebra and tangent vector elements are,

$$\mathbf{T} \triangleq \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \in T(n), \quad \mathbf{t}^\wedge \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{t} \\ \mathbf{0} & 0 \end{bmatrix} \in \mathfrak{t}(n), \quad \mathbf{t} \in \mathbb{R}^n.$$

Equivalence is easily verified by observing that $\mathbf{T}(0) = \mathbf{I}$, $\mathbf{T}(-\mathbf{t}) = \mathbf{T}(\mathbf{t})^{-1}$, and that the commutative composition

$$\mathbf{T}_1 \mathbf{T}_2 = \begin{bmatrix} \mathbf{I} & \mathbf{t}_1 + \mathbf{t}_2 \\ \mathbf{0} & 1 \end{bmatrix},$$

effectively adds the vectors \mathbf{t}_1 and \mathbf{t}_2 together. Since the sum in \mathbb{R}^n is commutative, so is the composition product in $T(n)$. Since $T(n)$ is a subgroup of $SE(n)$ with $\mathbf{R} = \mathbf{I}$, we can easily determine its exponential map by taking (156, 172) with $\mathbf{R} = \mathbf{I}$ and generalizing to any n ,

$$\text{Exp} : \mathbb{R}^n \rightarrow T(n); \quad \mathbf{T} = \text{Exp}(\mathbf{t}) = \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}. \quad (184)$$

The $T(n)$ exponential can be obtained also from the Taylor expansion of $\exp(\mathbf{t}^\wedge)$ noticing that $(\mathbf{t}^\wedge)^2 = \mathbf{0}$. This serves as immediate proof for the equivalent exponential of the $(\mathbb{R}^n, +)$ group, which is the identity,

$$\text{Exp} : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \mathbf{t} = \text{Exp}(\mathbf{t}). \quad (185)$$

This derives in trivial, commutative, right- and left- alike, plus and minus operators in \mathbb{R}^n ,

$$\mathbf{t}_1 \oplus \mathbf{t}_2 = \mathbf{t}_1 + \mathbf{t}_2 \quad (186)$$

$$\mathbf{t}_2 \ominus \mathbf{t}_1 = \mathbf{t}_2 - \mathbf{t}_1. \quad (187)$$

A. Jacobian blocks

We express translations indistinctly for $T(n)$ and \mathbb{R}^n , and note them S and T . The Jacobians are trivial (compare them with those of S^1 and $SO(2)$ in Section A-C1),

$$\mathbf{Ad}_T = \mathbf{I} \quad \in \mathbb{R}^{n \times n} \quad (188)$$

$$\mathbf{J}_T^{T^{-1}} = -\mathbf{I} \quad \in \mathbb{R}^{n \times n} \quad (189)$$

$$\mathbf{J}_T^{T \circ S} = \mathbf{J}_S^{T \circ S} = \mathbf{I} \quad \in \mathbb{R}^{n \times n} \quad (190)$$

$$\mathbf{J}_r = \mathbf{J}_l = \mathbf{I} \quad \in \mathbb{R}^{n \times n} \quad (191)$$

$$\mathbf{J}_T^{T \oplus v} = \mathbf{J}_v^{T \oplus v} = \mathbf{I} \quad \in \mathbb{R}^{n \times n} \quad (192)$$

$$\mathbf{J}_S^S \ominus T = -\mathbf{J}_T^S \ominus T = \mathbf{I} \quad \in \mathbb{R}^{n \times n}. \quad (193)$$

REFERENCES

- [1] H. Abbaspour and M. Moskowitz, *Basic Lie Theory*. WORLD SCIENTIFIC, 2007. [Online]. Available: <https://worldscientific.com/doi/abs/10.1142/6462>
- [2] R. Howe, “Very basic Lie theory,” *The American Mathematical Monthly*, vol. 90, pp. 600–623, 1983.
- [3] J. Stillwell, *Naive Lie Theory*. Springer-Verlag New York, 2008.
- [4] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2017.
- [5] E. Eade, “Lie groups for 2d and 3d transformations,” Tech. Rep.
- [6] C. Forster, L. Carbone, F. Dellaert, and D. Scaramuzza, “On-manifold preintegration for real-time visual-inertial odometry,” *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.
- [7] J. Deray and J. Solà, “Manif: A micro lie theory library for state estimation in robotics applications,” *Journal of Open Source Software*, vol. 5, no. 46, p. 1371, 2020. [Online]. Available: <https://doi.org/10.21105/joss.01371>
- [8] J. Solà, “Quaternion kinematics for the error-state Kalman filter,” *CoRR*, vol. abs/1711.02508, 2017. [Online]. Available: <http://arxiv.org/abs/1711.02508>
- [9] G. Gallego and A. Yezzi, “A compact formula for the derivative of a 3-D rotation in exponential coordinates,” Tech. Rep., 2013.
- [10] T. D. Barfoot and P. T. Furgale, “Associating uncertainty with three-dimensional poses for use in estimation problems,” *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679–693, June 2014.
- [11] G. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*, ser. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2011. [Online]. Available: <https://books.google.ch/books?id=hZ1CAAAQBAJ>
- [12] F. Dellaert and M. Kaess, “Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing,” vol. 25, no. 12, pp. 1181–1203, 2006.
- [13] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, “iSAM2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering,” in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 3281–3288.
- [14] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, “G2o: A general framework for graph optimization,” in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 3607–3613.
- [15] V. Ila, L. Polok, M. Solony, and P. Svoboda, “SLAM++ - a highly efficient and temporally scalable incremental SLAM framework,” *The International Journal of Robotics Research*, vol. 36, no. 2, pp. 210–230, 2017. [Online]. Available: <https://doi.org/10.1177/0278364917691110>