Li Group and Li Algebra in Robotics

Lin

Abstract

In the previous note, we have introduced rotaions and poses. In this note , we will look more deeply into the nature of these quantites. It's a fact that the set of the rotations is not a vector space but a non-commutative group, which is known as Li group. This note will concisely introduce Li group and Li algebra in robotics and discuss the relationship between them.

1 Introduction

Let's review some knowledge we've discussed in the previous note. In previous note, we've mentioned that the set of rotations can be represented as

$$\mathrm{SO}(n) = \left\{\mathbf{R} \in \mathbb{R}^{n\times n} | \mathbf{R}\mathbf{R}^{\mathbf{T}} = \mathbf{I}, \det(\mathbf{R}) = 1\right\}$$

which is called Special Orthogonal Group. The $\mathbf{R}\mathbf{R}^{\mathbf{T}} = \mathbf{I}$ orthogonality condition is to impose constraints on the matrix in order to reduce its degrees of freedom to n and $\det(\mathbf{R}) = 1$ ensures the rotation is a proper rotation. By the way, when $\det(\mathbf{R}) = -1$, the rotation is called improper rotation or rotary reflection.

The set of poses(rotaions and translations) can be represented as

$$\mathrm{SE}(n) = \left\{ \begin{bmatrix} \mathbf{R} & t \\ \mathbf{0}^{\mathbf{T}} & 1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} | \mathbf{R} \in \mathrm{SO}(n), \mathbf{t} \in \mathbb{R}^n \right\}$$

which is called Special Euclidean Group.

2 Li group

It's obvious that $\mathrm{SO}(n)$ is not a valid subspace of vectorspace. It's not closed under addition.

$$\mathbf{R_1} + \mathbf{R_2} \notin \mathrm{SO}(n)$$

So the same as $\mathrm{SE}(n)$. While $\mathrm{SO}(n)$ and $\mathrm{SE}(n)$ are not vectorspaces, it can be shown that they are Lie Group. Lie group is the combination of group and differential manifold. A group is a set of elements together with an operation, which can be denoted as $G = (A, \cdot)$, while satifying four conditions called group axioms:

$$(Closure) \quad \forall a_1, a_2 \in A, \quad a_1 \cdot a_2 \in A \tag{1}$$
$$(Associativity) \quad \forall a_1, a_2, a_3 \in A, \quad (a_1 \cdot a_2) \cdot a_3 = a_1 \cdot (a_2 \cdot a_3) \tag{2}$$
$$(Identity) \quad \exists e \in A, \forall a \in A, \quad a \cdot e = e \cdot a = a \tag{3}$$
$$(Invertibility) \quad \forall a \in A, \exists a^{-1} \in A, \quad a \cdot a^{-1} = e \tag{4}$$

Lie group is also a differential manifold indicates that the group operations are smooth, which means that we can use differential calculus on the group. More strictly, supposed we have a mapping f whose domain is the set X which composes the group, if we have a infinitesimal $\Delta x \in X$, and $x \in X$, then

$$\Delta y = f(x + \Delta x) - f(x) = k\Delta x + o(\Delta x)$$

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

where k is known as the derivative of f at x, denoted as $f'(x)$, and $k\Delta x$ is known as the differentiation of y, denoted as Δy.

In my perspective, in the above formula, $x + \Delta x$ indicates that there is a infinitesimal change to x which is caused by a infinitesimal Δx. However, it doesn't have any asscoiation with addition operation, it can also be a multiplication $x\Delta x$ as long as triggers only a infinitesimal change.

3 Li Algebra

3.1 Before Li Algebra

Suppose we have a rotation matrix $\mathbf{R}$,then

$$\mathbf{R}\mathbf{R}^{\mathbf{T}} = \mathbf{I}$$

Suppose rotation $\mathbf{R}$ is a mapping with repsect to time t, and taking derivative with respect to t on the both side of the equation.

$$\mathbf{R}'(t)\mathbf{R}^{\mathbf{T}}(t) + \mathbf{R}(t)(\mathbf{R}^{\mathbf{T}}(t))' = \mathbf{0}$$

Therefore,

$$\mathbf{R}'(t)\mathbf{R}^{\mathbf{T}}(t) = -(\mathbf{R}'(t)\mathbf{R}^{\mathbf{T}}(t))^{\mathbf{T}}$$

which indicates that $\mathbf{R}'(t)\mathbf{R}^{\mathbf{T}}(t)$ is a skew-symmetric matrix and can be represented by a vector $\phi(t)$.

$$\mathbf{R}'(t)\mathbf{R}^{\mathbf{T}}(t) = (\phi(t))^{\wedge}$$

Hence,

$$\mathbf{R}'(t) = (\phi(t))^{\wedge}\mathbf{R}(t) = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix} \mathbf{R}(t)$$

Suppose $\mathbf{R}(0) = \mathbf{I}$ and we implement Maclaurin expansion to $\mathbf{R}(t)$

$$\begin{aligned} \mathbf{R}(t) &= \mathbf{R}(0) + \mathbf{R}'(0)t \\ &= \mathbf{I} + \phi_0^{\wedge} t \end{aligned}$$

According to the above euqation, we can say that $\phi_0^{\wedge}$ reflects the derivative of $\mathbf{R}$ around $t = 0$.

Furthermore, if we assume the derivative of $\mathbf{R}$ is a constant around $t = 0$, then we have the following differential euqation which has initial value $\mathbf{R}(0) = \mathbf{I}$.

$$\mathbf{R}'(t) = \phi_0^{\wedge}\mathbf{R}(t)$$

The solution is

$$\mathbf{R}(t) = \mathbf{exp}(\phi_0^{\wedge} t)$$

3.2 The definiton of Li algebra

With every matrix Lie group is associated a Lie algebra, which consists of a vectorspace, $\mathbb{V}$, over a field, $\mathbb{F}$, together with a binary operation, $[\cdot,\cdot]$, called the Lie bracket and that satifies four properties:

$$\begin{aligned} (Closure): &\quad \forall \mathbf{X}, \mathbf{Y} \in \mathbb{V}, \quad [\mathbf{X}, \mathbf{Y}] \in \mathbb{V} \\ (Bilinearity): &\quad \forall \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{V}, \quad \forall a, b \in \mathbb{F}, \quad [a\mathbf{X} + b\mathbf{Y}, \mathbf{Z}] = a[\mathbf{X}, \mathbf{Z}] + b[\mathbf{Y}, \mathbf{Z}] \\ &\quad \forall \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{V}, \quad \forall a, b \in \mathbb{F}, \quad [\mathbf{Z}, a\mathbf{X} + b\mathbf{Y}] = a[\mathbf{Z}, \mathbf{X}] + b[\mathbf{Z}, \mathbf{Y}] \\ (Alternating): &\quad \forall X \in \mathbb{V}, \quad [\mathbf{X}, \mathbf{X}] = \mathbf{0} \\ (Jacobi\ identity): &\quad \forall \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{V}, \quad [\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]] + [\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]] + [\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]] = \mathbf{0} \end{aligned}$$

The vectorspace of a Lie algebra is the tangent space of the associated Lie group at the identity element of the group, and it completely captures the local structure of the group. It can be shown that $\mathfrak{g} = (\mathbb{R}^3, \mathbb{R}, \times)$ is a Lie algebra.

3.3 Li algebra $\mathfrak{so}(3)$

The Lie algebra associated with $\mathrm{SO}(3)$ is given by:

$$\begin{aligned} \text{vectorspace :} \quad & \mathfrak{so}(3) = \{\mathbf{\Phi} = \phi^\wedge \in \mathbb{R}^{3\times3} | \phi \in \mathbb{R}^3\} \\ \text{field :} \quad & \mathbb{R} \\ \text{Lie bracket :} \quad & [\mathbf{\Phi_1}, \mathbf{\Phi_2}] = \mathbf{\Phi_1}\mathbf{\Phi_2} - \mathbf{\Phi_2}\mathbf{\Phi_1} \end{aligned}$$

where $\phi^\wedge$ indicates the skew-symmetric matrix of vector ϕ.

It's obvious that $\mathfrak{so}(3)$ is a vectorspace. Let's briefly show that for Lie bracket properties hold.Suppose $\mathbf{\Phi_1} = \phi_1^\wedge$, $\mathbf{\Phi_2} = \phi_2^\wedge$, for the clousre property, we have

$$[\mathbf{\Phi_1}, \mathbf{\Phi_2}] = \phi_1^\wedge\phi_2^\wedge - \phi_2^\wedge\phi_1^\wedge = (\phi_1^\wedge\phi_2)^\wedge$$

Bilinearity follows directly from the fact that $(\cdot)^\wedge$ is a linear operator.

The alternating property can be seen easily.

$$[\mathbf{\Phi_1}, \mathbf{\Phi_1}] = \phi_1^\wedge\phi_1^\wedge - \phi_1^\wedge\phi_1^\wedge = \mathbf{0}$$

The Jacobi identity property can be shown by the following equation.

$$\begin{aligned} [\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]] + [\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]] + [\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]] &= X^\wedge(Y^\wedge Z^\wedge - Z^\wedge Y^\wedge) - (Y^\wedge Z^\wedge - Z^\wedge Y^\wedge)X^\wedge \\ &+ Y^\wedge(Z^\wedge X^\wedge - X^\wedge Z^\wedge) - (Z^\wedge X^\wedge - X^\wedge Z^\wedge)Y^\wedge \\ &+ Z^\wedge(X^\wedge Y^\wedge - Y^\wedge X^\wedge) - (X^\wedge Y^\wedge - Y^\wedge X^\wedge)Z^\wedge \\ &= \mathbf{0} \end{aligned}$$

Informally, we refer $\mathfrak{so}(3)$ as the Lie algebra associated with the $\mathrm{SO}(3)$, but technically $\mathfrak{so}(3)$ is only the associated vectorspace.

3.4 Li algebra $\mathfrak{se}(3)$

The Li algebra associated with $\mathrm{SE}(3)$ is given by:

$$\begin{aligned} \text{vectorspace :} \quad & \mathfrak{se}(3) = \{\mathbf{\Xi} = \xi^\wedge \in \mathbb{R}^{4\times4} | \xi \in \mathbb{R}^6\} \\ \text{field :} \quad & \mathbb{R} \\ \text{Lie bracket :} \quad & [\mathbf{\Xi_1}, \mathbf{\Xi_2}] = \mathbf{\Xi_1}\mathbf{\Xi_2} - \mathbf{\Xi_2}\mathbf{\Xi_1} \end{aligned}$$

where

$$\xi^\wedge = \begin{bmatrix} \rho \\ \phi \end{bmatrix}^\wedge = \begin{bmatrix} \phi^\wedge & \rho \\ \mathbf{0^T} & 0 \end{bmatrix} \in \mathbb{R}^{4\times4}, \quad \rho, \phi \in \mathbb{R}^3.$$

This is an overloading of the operator $(\cdot)^\wedge$ which used to take a vector as input and output the skew-symmetric matrix of the input vector. Now, $(\cdot)^\wedge$ takes an element of $\mathbb{R}^6$ and outputs a matrix of $\mathbb{R}^{4\times4}$ and it's still a linear operator. The proof of $\mathfrak{se}(3)$'s properties are similiar with $\mathfrak{so}(3)$'s.

Again, informally, we refer $\mathfrak{se}(3)$ as the Lie algebra associated with the $\mathrm{SE}(3)$, but technically $\mathfrak{se}(3)$ is only the associated vectorspace.

4 Exponential Mapping

4.1 Exponential Mapping from $\mathfrak{so}(3)$ to $\mathrm{SO}(3)$

In chapter 3.1, we derive the following differential equation with respect to rotation matrix $\mathbf{R}$.

$$\mathbf{R}'(t) = \phi^\wedge(t)\mathbf{R}(t)$$

The solution of the differential equation is:

$$\mathbf{R} = \exp(\phi^\wedge)$$

Therefore, we can conclude that there exists a mapping from Li algebra to Li group that is a exponential mapping. However, how to calculate the exponential of a matrix?

The matrix exponential is given by the Maclaurin expansion:

$$\exp(\mathbf{A}) = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \cdots + \frac{1}{n!}\mathbf{A}^n = \sum_{n=1}^{\infty} \frac{1}{n!}\mathbf{A}^n$$

Given the defintion of matrix exponential, we're still unable to calculate it, because it's a infinite series. We need some delicate transformations.

First, we can represent a vector ϕ by its length θ and a unit vector $\mathbf{a}$, which means that $\phi = \theta\mathbf{a}$.

To a unit vector $\mathbf{a}$, it can be shown that

$$\begin{aligned}
\mathbf{a}^\wedge\mathbf{a}^\wedge &= \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -a_2^2 - a_3^2 & a_1a_2 & a_1a_3 \\ a_1a_2 & -a_1^2 - a_3^2 & a_1a_3 \\ a_1a_3 & a_2a_3 & -a_1^2 - a_2^2 \end{bmatrix} \\
&= \mathbf{a}\mathbf{a}^\mathbf{T} - \mathbf{I}
\end{aligned}$$

Furthermore,

$$\mathbf{a}^\wedge\mathbf{a}^\wedge\mathbf{a}^\wedge = \mathbf{a}^\wedge(\mathbf{a}\mathbf{a}^\mathbf{T} - \mathbf{I}) = -\mathbf{a}^\wedge$$

Therefore,

$$\begin{aligned}
\exp(\phi^\wedge) &= \exp(\theta\mathbf{a}^\wedge) \\
&= \mathbf{I} + \theta\mathbf{a}^\wedge + \frac{1}{2!}\theta^2(\mathbf{a}^\wedge)^2 + \frac{1}{3!}\theta^3(\mathbf{a}^\wedge)^3 \cdots + \frac{1}{n!}\theta^n(\mathbf{a}^\wedge)^n \\
&= \mathbf{a}\mathbf{a}^\mathbf{T} - \mathbf{a}^\wedge\mathbf{a}^\wedge + (\theta - \frac{1}{3!}\theta^3 + \cdots + \frac{(-1)^{n-1}}{(2n-1)!}\theta^{2n-1})\mathbf{a}^\wedge \\
&+ (\frac{1}{2!}\theta^2 - \frac{1}{4!}\theta^4 + \frac{1}{6!}\theta^6 + \cdots + \frac{(-1)^{n-1}}{(2n)!}\theta^{2n})\mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{a}\mathbf{a}^\mathbf{T} + (\theta - \frac{1}{3!}\theta^3 + \cdots + \frac{(-1)^{n-1}}{(2n-1)!}\theta^{2n-1})\mathbf{a}^\wedge \\
&- (\mathbf{I} - \frac{1}{2!}\theta^2 + \frac{1}{4!}\theta^4 - \frac{1}{6!}\theta^6 + \cdots + \frac{(-1)^{n}}{(2n)!}\theta^{2n})\mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{a}\mathbf{a}^\mathbf{T} + \sin\theta\mathbf{a}^\wedge - \cos\theta\mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{a}\mathbf{a}^\mathbf{T} + \sin\theta\mathbf{a}^\wedge - \cos\theta(\mathbf{a}\mathbf{a}^\mathbf{T} - I) \\
&= \cos\theta\mathbf{I} + (1 - \cos\theta)\mathbf{a}\mathbf{a}^\mathbf{T} + \sin\theta\mathbf{a}^\wedge
\end{aligned}$$

where the result is exactly the Rodrigues's formula, which means that the vectorspace of Li algebra is the set of rotation vectors. In addition, the exponential mapping between Li group and Li algebra is surjective-only, which means that every element of $\mathrm{SO}(3)$ can be generated from multiple elements of $\mathfrak{so}(3)$. The many-to-one mapping is related to the concept of singularity (or nonuniqueness) in rotation parameterizations. But if we limit $|\phi| < \pi$, then the exponential mapping is one-to-one.

It's nature to think about that the inverse mapping of exponential mapping is logrithmic mapping, and it's true.

$$\phi = \mathrm{In}(\mathbf{R})^\vee = \left(\sum_{i=0}^{\infty} \frac{(-1)^n}{n+1}(\mathbf{R} - \mathbf{I})^{n+1}\right)^\vee$$

In fact, $\sum_{i=0}^{\infty} \frac{(-1)^n}{n+1}(x)^{n+1}$ is the Maclaurin expansion of $\mathrm{In}(x+1)$. Because $\mathrm{In}(x)$ is undefined at $x = 0$, we use $\mathrm{In}((x-1)+1)$ to represent $\mathrm{In}(x)$.

The transformation from $\mathrm{SO}(3)$ to $\mathfrak{so}(3)$ is equivalent to the transformation from rotation matrix to rotation vector (or angle/axis). Therefore, for every element $\mathbf{R} \in \mathrm{SO}(3)$, we can calculate its corresponding element $\phi \in \mathfrak{so}(3)$ as follow:

$$\arccos(\theta) = \frac{\mathrm{tr}(\mathbf{R}) - 1}{2}$$

and consider the rotation axis is invariant with the rotation opertaion, we have

$$\mathbf{R}\phi = \phi$$

Therefore, the rotation vector is an eigenvector of $\mathbf{R}$ corresponding to an eigenvalue of 1.

4.2 Exponential Mapping from $\mathfrak{se}(3)$ to $\mathrm{SE}(3)$

Compared with the exponential mapping from $\mathfrak{so}(3)$ to $\mathrm{SO}(3)$, the exponential mapping from $\mathfrak{se}(3)$ to $\mathrm{SE}(3)$ is more complicate.

Before deriving the formula of exponential mapping, it has to be known that

$$\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}^n = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \text{if n} = 0 \\ \begin{bmatrix} a^n & a^{n-1}b \\ 0 & 1 \end{bmatrix}, & \text{if n} \geq 1 \end{cases}$$

Therefore, we can derive the exponential mapping as follow

$$\begin{aligned}
\exp(\boldsymbol{\xi}^\wedge) &= \sum_{i=0}^{\infty} \frac{1}{n!} \begin{bmatrix} \boldsymbol{\phi}^\wedge & \boldsymbol{\rho} \\ \mathbf{0}^{\boldsymbol{T}} & 0 \end{bmatrix}^n \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{\infty} \frac{1}{n!}(\boldsymbol{\phi}^\wedge)^n & \sum_{i=1}^{\infty} \frac{1}{n!}(\boldsymbol{\phi}^\wedge)^{n-1}\boldsymbol{\rho} \\ \mathbf{0}^{\mathbf{T}} & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{\infty} \frac{1}{n!}(\boldsymbol{\phi}^\wedge)^n & \sum_{i=0}^{\infty} \frac{1}{(n+1)!}(\boldsymbol{\phi}^\wedge)^{n}\boldsymbol{\rho} \\ \mathbf{0}^{\mathbf{T}} & 0 \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=0}^{\infty} \frac{1}{n!}(\boldsymbol{\phi}^\wedge)^n & \sum_{i=0}^{\infty} \frac{1}{(n+1)!}(\boldsymbol{\phi}^\wedge)^{n}\boldsymbol{\rho} \\ \mathbf{0}^{\mathbf{T}} & 1 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{R} & \mathbf{J}\boldsymbol{\rho} \\ \mathbf{0}^{\mathbf{T}} & 1 \end{bmatrix}
\end{aligned}$$

where $\mathbf{J} = \sum_{i=0}^{\infty} \frac{1}{(n+1)!}(\boldsymbol{\phi}^\wedge)^n$.

Suppose we have a matrix $T = \begin{bmatrix} \mathbf{R} & \mathbf{r} \\ \mathbf{0}^{\mathbf{T}} & 1 \end{bmatrix} \in \mathrm{SE}(3)$, we've already knew how to calculate $\boldsymbol{\phi}$ from R, and if we desire to get the corresponding $\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{bmatrix} \in \mathfrak{se}(3)$, all we have to do is solve the following linear equation

$$\boldsymbol{\rho} = \mathbf{J}^{-1}\mathbf{r}$$

We can also develop a direct series expression for $\mathbf{T}$ from the exponential mapping by using the following identity

$$(\boldsymbol{\xi}^\wedge)^4 + \theta^2(\boldsymbol{\xi}^\wedge) = \mathbf{0}^{\mathbf{T}}$$

where $\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\phi} \end{bmatrix}$ and $\boldsymbol{\phi} = \theta\mathbf{a}$. Using this identity, we can represent the quartic and higher terms by the quadratic and cubic terms.

$$\begin{aligned}
\exp(\boldsymbol{\xi}^\wedge) &= \mathbf{I} + \boldsymbol{\xi}^\wedge + \frac{1}{2!}(\boldsymbol{\xi}^\wedge)^2 + \cdots + \frac{1}{n!}(\boldsymbol{\xi}^\wedge)^n \\
&= \mathbf{I} + \boldsymbol{\xi}^\wedge + (\frac{1}{2!} - \frac{1}{4!}\theta^2 + \frac{1}{6!}\theta^4 + \cdots + \frac{(-1)^{n-1}}{(2n)!}\theta^{2n-2})(\boldsymbol{\xi}^\wedge)^2 \\
&+ (\frac{1}{3!} - \frac{1}{5!}\theta^2 + \frac{1}{7!}\theta^4 + \cdots + \frac{(-1)^{n-1}}{(2n+1)!}\theta^{2n-3})(\boldsymbol{\xi}^\wedge)^3 \\
&= \mathbf{I} + \boldsymbol{\xi}^\wedge + \frac{1-\cos\theta}{\theta^2}(\boldsymbol{\xi}^\wedge)^2 + \frac{\theta - \sin\theta}{\theta^3}(\boldsymbol{\xi}^\wedge)^3.
\end{aligned}$$

Using thie approach avoids the need to deal with $\mathbf{T}$'s constituent blocks.

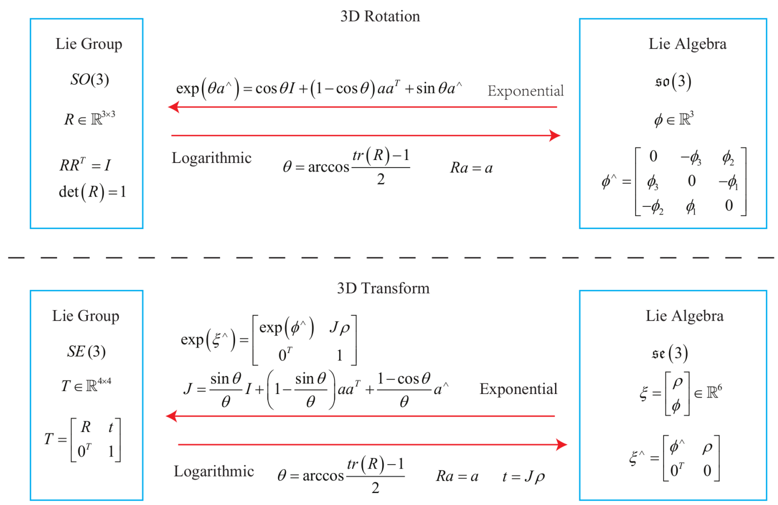

Figure 1: The relationship between SO(3),SE(3),$\mathfrak{so}(3)$,$\mathfrak{se}(3)$

5 Jacobian

The matrix $\mathbf{J}$ we've mentioned in chapter 4.2 is called (left) Jacobian of SO(3), and it not only plays an important role in the transformation between the translation component of $\mathfrak{se}(3)$ and SE(3), but also appears in the Baker-Campbell-Hausdorff Formula. The close-form expressions for $\mathbf{J}$ and its inverse enables us to calculate more easily.

$$
\begin{aligned}
\mathbf{J} &= \sum_{i=0}^{\infty} \frac{1}{(n+1)!} (\phi^\wedge)^n \\
&= \mathbf{I} + \frac{1}{2!}\theta \mathbf{a}^\wedge + \frac{1}{3!}\theta^2 (\mathbf{a}^\wedge)^2 + \cdots + \frac{1}{(n+1)!}\theta^n (\mathbf{a}^\wedge)^n \\
&= \mathbf{I} + (\frac{1}{2!}\theta - \frac{1}{4!}\theta^3 + \frac{1}{6!}\theta^5 + \cdots + \frac{(-1)^{n-1}}{(2n)!}\theta^{2n-1})\mathbf{a}^\wedge \\
&+ (\frac{1}{3!}\theta^2 - \frac{1}{5!}\theta^4 + \frac{1}{7!}\theta^6 + \cdots + \frac{(-1)^{n-1}}{(2n+1)!}\theta^{2n})\mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{I} + \frac{1}{\theta}(\frac{1}{2!}\theta^2 - \frac{1}{4!}\theta^4 + \frac{1}{6!}\theta^6 + \cdots + \frac{(-1)^{n-1}}{(2n)!}\theta^{2n})\mathbf{a}^\wedge \\
&+ \frac{1}{\theta}(\frac{1}{3!}\theta^3 - \frac{1}{5!}\theta^5 + \frac{1}{7!}\theta^7 + \cdots + \frac{(-1)^{n-1}}{(2n+1)!}\theta^{2n+1})\mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{I} - \frac{1}{\theta}(-\frac{1}{2!}\theta^2 + \frac{1}{4!}\theta^4 - \frac{1}{6!}\theta^6 + \cdots + \frac{(-1)^{n}}{(2n)!}\theta^{2n})\mathbf{a}^\wedge \\
&- \frac{1}{\theta}(\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 - \frac{1}{7!}\theta^7 + \cdots + \frac{(-1)^{n-1}}{(2n+1)!}\theta^{2n+1})\mathbf{a}^\wedge\mathbf{a}^\wedge + \mathbf{a}^\wedge\mathbf{a}^\wedge \\
&= \mathbf{I} - \frac{\cos\theta - 1}{\theta}\mathbf{a}^\wedge - \frac{\sin\theta}{\theta}(\mathbf{a}\mathbf{a}^\mathbf{T} - \mathbf{I}) + \mathbf{a}\mathbf{a}^\mathbf{T} - \mathbf{I} \\
&= \frac{\sin\theta}{\theta}\mathbf{I} - \frac{\theta - \sin\theta}{\theta}\mathbf{a}\mathbf{a}^\mathbf{T} + \frac{1-\cos\theta}{\theta}\mathbf{a}^\wedge
\end{aligned}
$$

and

$$
\mathbf{J}^{-1} = \frac{\theta}{2}\cot\frac{\theta}{2}\mathbf{I} + (1 - \frac{\theta}{2}\cot\frac{\theta}{2})\mathbf{a}\mathbf{a}^\mathbf{T} - \frac{\theta}{2}\mathbf{a}^\wedge
$$

where $\boldsymbol{\phi} = \theta\mathbf{a}$. When $\theta = 2\pi k,\ k \in \mathbb{Z}$, $\cot\frac{\theta}{2}$ is not defined. Therefore, there are singularities asscoiated with $\mathbf{J}$ (i.e., the inverse doesn't exist) at $\theta = 2\pi k$ with k a non-zero integer.

6 Adjoint

A 6×6 matrix $\boldsymbol{\mathcal{T}}$ can be constructed directly from the components of the 4×4 transformation matrix. We call this the adjoint of an element of $\mathrm{SE}(3)$.

$$\boldsymbol{\mathcal{T}} = \mathrm{Ad}(\mathbf{T}) = \mathrm{Ad}\left(\begin{bmatrix}\mathbf{R} & \mathbf{t}\\ \mathbf{0^T} & 1\end{bmatrix}\right) = \begin{bmatrix}\mathbf{R} & \mathbf{t}^\wedge\mathbf{R}\\ \mathbf{0} & \mathbf{R}\end{bmatrix}$$

We will abuse the notation a bit and say that the set of adjoints of all elements of $\mathrm{SE}(3)$ is denoted

$$\mathrm{Ad}(\mathrm{SE}(3)) = \{\boldsymbol{\mathcal{T}} = \mathrm{Ad}(\mathbf{T}) | \mathbf{T} \in \mathrm{SE}(3)\}$$

It turns out that $\mathrm{Ad}(\mathrm{SE}(3))$ is a matrix Lie group.

Lemma 1. *Suppose* $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3} \in \mathbb{R}^3$ *and* $\mathbf{e_1} \times \mathbf{e_2} = \mathbf{e_3}$*, then*

$$\forall\, \mathbf{p} \in \mathbb{R}^3, \quad \mathbf{e_2} \cdot (\mathbf{p} \times \mathbf{e_1}) = \mathbf{e_3} \cdot \mathbf{p}$$

Proof. Let's assume $\mathbf{e_1} = (e_{11}, e_{12}, e_{13})$,$\mathbf{e_2} = (e_{21}, e_{22}, e_{23})$ and $\mathbf{e_3} = (e_{31}, e_{32}, e_{33})$

$$\begin{aligned}
\mathbf{e_2} \cdot (\mathbf{p} \times \mathbf{e_1}) &= \begin{bmatrix}e_{21} & e_{22} & e_{23}\end{bmatrix} \begin{bmatrix}0 & -p_3 & p_2\\ p_3 & 0 & -p_1\\ -p_2 & p_1 & 0\end{bmatrix} \begin{bmatrix}e_{11}\\ e_{12}\\ e_{13}\end{bmatrix}\\
&= \begin{bmatrix}e_{21} & e_{22} & e_{23}\end{bmatrix} \begin{bmatrix}-p_3e_{12} + p_2e_{13}\\ p_3e_{11} - p_1e_{13}\\ -p_2e_{11} + p_1e_{12}\end{bmatrix}\\
&= (e_{22}e_{11} - e_{21}e_{12})p_3 + (e_{21}e_{13} - e_{23}e_{11})p_2 + (e_{21}e_{12} - e_{22}e_{e13})p_1\\
&= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ e_{11} & e_{12} & e_{13}\\ e_{21} & e_{22} & e_{23}\end{vmatrix} \begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix}\\
&= (\mathbf{e_1} \times \mathbf{e_2}) \cdot \mathbf{p}\\
&= \mathbf{e_3} \cdot \mathbf{p}
\end{aligned}$$

□

Lemma 2. *Suppose* $\mathbf{R} \in \mathrm{SO}(3)$*, then*

$$\forall \mathbf{v} \in \mathbb{R}^3, \quad \mathbf{R}\mathbf{v}^\wedge\mathbf{R^T} = (\mathbf{R}\mathbf{v})^\wedge$$

Proof. Let's assume $\mathbf{R}$ is a matrix of row vectors, then $\mathbf{R} = \begin{bmatrix} \mathbf{e_1} \\ \mathbf{e_2} \\ \mathbf{e_3} \end{bmatrix}$

$$\begin{aligned}
\mathbf{R}\mathbf{v}^\wedge\mathbf{R}^\mathbf{T} &= \begin{bmatrix} \mathbf{e_1} \\ \mathbf{e_2} \\ \mathbf{e_3} \end{bmatrix} \begin{bmatrix} \mathbf{v} \times \mathbf{e_1^T} & \mathbf{v} \times \mathbf{e_2^T} & \mathbf{v} \times \mathbf{e_3^T} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{e_1} \cdot (\mathbf{v} \times \mathbf{e_1^T}) & \mathbf{e_1} \cdot (\mathbf{v} \times \mathbf{e_2^T}) & \mathbf{e_1} \cdot (\mathbf{v} \times \mathbf{e_3^T}) \\ \mathbf{e_2} \cdot (\mathbf{v} \times \mathbf{e_1^T}) & \mathbf{e_2} \cdot (\mathbf{v} \times \mathbf{e_2^T}) & \mathbf{e_2} \cdot (\mathbf{v} \times \mathbf{e_3^T}) \\ \mathbf{e_3} \cdot (\mathbf{v} \times \mathbf{e_1^T}) & \mathbf{e_3} \cdot (\mathbf{v} \times \mathbf{e_2^T}) & \mathbf{e_3} \cdot (\mathbf{v} \times \mathbf{e_3^T}) \end{bmatrix} \\
&= \begin{bmatrix} (\mathbf{e_1} \times \mathbf{e_1}) \cdot \mathbf{v} & (\mathbf{e_2} \times \mathbf{e_1}) \cdot \mathbf{v} & (\mathbf{e_3} \times \mathbf{e_1}) \cdot \mathbf{v} \\ (\mathbf{e_1} \times \mathbf{e_2}) \cdot \mathbf{v} & (\mathbf{e_1} \times \mathbf{e_2}) \cdot \mathbf{v} & (\mathbf{e_3} \times \mathbf{e_2}) \cdot \mathbf{v} \\ (\mathbf{e_1} \times \mathbf{e_3}) \cdot \mathbf{v} & (\mathbf{e_2} \times \mathbf{e_3}) \cdot \mathbf{v} & (\mathbf{e_3} \times \mathbf{e_3}) \cdot \mathbf{v} \end{bmatrix} \\
&= \begin{bmatrix} 0 & (-\mathbf{e_3}) \cdot \mathbf{v} & \mathbf{e_2} \cdot \mathbf{v} \\ \mathbf{e_3} \cdot \mathbf{v} & 0 & -(\mathbf{e_1}) \cdot \mathbf{v} \\ -(\mathbf{e_2}) \cdot \mathbf{v} & \mathbf{e_1} \cdot \mathbf{v} & 0 \end{bmatrix} \\
&= \left(\begin{bmatrix} \mathbf{e_1} \cdot \mathbf{v} \\ \mathbf{e_2} \cdot \mathbf{v} \\ \mathbf{e_3} \cdot \mathbf{v} \end{bmatrix} \right)^\wedge \\
&= \left(\begin{bmatrix} \mathbf{e_1} \\ \mathbf{e_2} \\ \mathbf{e_3} \end{bmatrix} \mathbf{v} \right)^\wedge \\
&= (\mathbf{R}\mathbf{v})^\wedge
\end{aligned}$$

$\square$

The transition between the second and third line of the proof of Lemma 2 is based on Lemma 1.

Corollary 1. *Suppose* $\mathbf{R} \in \mathrm{SO}(3)$, *then*

$$\forall \mathbf{v} \in \mathbb{R}^3, \quad \mathbf{R} \exp(v^\wedge)\mathbf{R}^\mathbf{T} = \exp((\mathbf{R}\mathbf{v})^\wedge)$$

Proof. Let $\mathbf{v} = \theta\mathbf{a}$, and then

$$\begin{aligned}
\mathbf{R} \exp(v^\wedge)\mathbf{R}^\mathbf{T} &= \mathbf{R}(\cos\theta\mathbf{I} + (1 - \cos\theta)\mathbf{a}\mathbf{a}^\mathbf{T} + \sin\theta\mathbf{a}^\wedge)\mathbf{R}^\mathbf{T} \\
&= \cos\theta\mathbf{R}\mathbf{R}^\mathbf{T} + (1 - \cos\theta)\mathbf{R}\mathbf{a}\mathbf{a}^\mathbf{T}\mathbf{R}^\mathbf{T} + \sin\theta\mathbf{R}(\mathbf{a}^\wedge)\mathbf{R}^\mathbf{T} \\
&= \cos\theta\mathbf{I} + (1 - \cos\theta)\mathbf{R}\mathbf{a}(\mathbf{R}\mathbf{a})^\mathbf{T} + \sin\theta\mathbf{R}\mathbf{a}^\wedge \\
&= \exp(\theta(\mathbf{R}\mathbf{a})^\wedge) \\
&= \exp((Rv)^\wedge)
\end{aligned}$$

$\square$

Corollary 1 is called the adjoint property of $\mathrm{SO}(3)$. In addition, $\mathrm{SE}(3)$ has the same property:

$$\mathbf{T} \exp(\boldsymbol{\xi}^\wedge)\mathbf{T}^{-1} = \exp((\mathrm{Ad}(\mathbf{T})\boldsymbol{\xi})^\wedge)$$

Now, let's start proofing $\mathrm{Ad}(\mathrm{SE}(3))$ is a matrix Lie group.

For closure we let $\mathcal{T}_1 = \mathrm{Ad}(\mathbf{T}_1), \mathcal{T}_2 = \mathrm{Ad}(\mathbf{T}_2) \in \mathrm{Ad}(\mathrm{SE}(3))$, and then

$$\begin{aligned}
\mathbf{T}_1\mathbf{T}_2 &= \begin{bmatrix} \mathbf{R_1} & \mathbf{t_1}^\wedge\mathbf{R_1} \\ \mathbf{0} & \mathbf{R_1} \end{bmatrix} \begin{bmatrix} \mathbf{R_2} & \mathbf{t_2}^\wedge\mathbf{R_2} \\ \mathbf{0} & \mathbf{R_2} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{R_1}\mathbf{R_2} & \mathbf{R_1}\mathbf{t_2}^\wedge\mathbf{R_2} + \mathbf{t_1}^\wedge\mathbf{R_1}\mathbf{R_2} \\ \mathbf{0} & \mathbf{R_1}\mathbf{R_2} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{R_1}\mathbf{R_2} & (\mathbf{R_1}\mathbf{t_2} + \mathbf{t_1})^\wedge\mathbf{R_1}\mathbf{R_2} \\ \mathbf{0} & \mathbf{R_1}\mathbf{R_2} \end{bmatrix} \in \mathrm{Ad}(\mathrm{SE}(3))
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R_1}\mathbf{t_2}^\wedge\mathbf{R_2} + \mathbf{t_1}^\wedge\mathbf{R_1}\mathbf{R_2} &= \mathbf{R_1}\mathbf{t_2}^\wedge(\mathbf{R_1^T}\mathbf{R_1})\mathbf{R_2} + \mathbf{t_1}^\wedge\mathbf{R_1}\mathbf{R_2} \\
&= (\mathbf{R_1}\mathbf{t_2})^\wedge\mathbf{R_1}\mathbf{R_2} + \mathbf{t_1}^\wedge\mathbf{R_1}\mathbf{R_2} \\
&= (\mathbf{R_1}\mathbf{t_2} + \mathbf{t_1})^\wedge\mathbf{R_1}\mathbf{R_2}\ .
\end{aligned}$$

The above transition from the first line to the second line is based on Lemma 2. The transition from the second line to the third line derives directly from the linearity of $(\cdot)^\wedge$ operation.

The asscoiativity of $\mathrm{Ad}(\mathrm{SE}(3))$ follows from basic properties of matrix multiplication . The identity element of $\mathrm{Ad}(\mathrm{SE}(3))$ is the 6×6 identity matrix. With resepct to invertibility, for every element $\mathcal{T} \in \mathrm{Ad}(\mathrm{SE}(3))$,

$$\begin{aligned}
\mathcal{T}^{-1} &= \begin{bmatrix} \mathbf{R^T} & -\mathbf{R^T}\mathbf{t}\wedge \\ \mathbf{0} & \mathbf{R^T} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{R^T} & (-\mathbf{R^T}\mathbf{t}) \wedge \mathbf{R^T} \\ \mathbf{0} & \mathbf{R^T} \end{bmatrix} \\
&= \mathrm{Ad}\left(\begin{bmatrix} \mathbf{R^T} & -\mathbf{R^T}\mathbf{t} \\ \mathbf{0^T} & 1 \end{bmatrix}\right) = \mathrm{Ad}\left(\mathbf{T}^{-1}\right) \in \mathrm{Ad}(\mathrm{SE}(3))
\end{aligned}$$

Other than smoothness, these four properties (clousre, asscoiativity, identity and invertibility) show that $\mathrm{Ad}(\mathrm{SE}(3))$ is a matrix lie group.

7 Baker-Campbell-Hausdorff Formula

When $a, b \in \mathbb{R}$, we can derive the following identity:

$$\exp(a)\exp(b) = \exp(a + b)\ .$$

This is because exponential mapping is a homomorphic mapping over $\mathbb{R}$. However, this is not established in matrix case, and there exists some remainder terms. More specially, given $\mathbf{A}, \mathbf{B} \in \mathrm{SO}(3)$ and $\boldsymbol{\phi}_1 = \exp(\mathbf{A}), \boldsymbol{\phi}_2 = \exp(\mathbf{B}) \in \mathfrak{so}(3)$, then

$$\exp(\mathbf{A})\exp(\mathbf{B}) \neq \exp(\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2)$$

which is the same with respect to $\mathrm{SE}(3)$ and $\mathfrak{se}(3)$.

To compound two matrix exponentials, we use the Baker-Campbell-Hausdorff (BCH) formula:

$$\mathrm{In}(\exp(\mathbf{A})\exp(\mathbf{B})) = \sum_{i=0}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{\substack{r_i + s_i > 0 \\ 1 \le i \le n}} \frac{\left(\sum_{i=1}^{n}(r_i + s_i)\right)^{-1}}{\prod_{i=1}^{n} r_i! s_i!} [\mathbf{A^{r_1}B^{s_1}A^{r_2}B^{s_2}} \dots \mathbf{A^{r_n}B^{s_n}}]$$

where

$$\begin{aligned}
&[\mathbf{A^{r_1}B^{s_1}A^{r_2}B^{s_2}} \cdots \mathbf{A^{r_n}B^{s_n}}] \\
&= \underbrace{[\mathbf{A}, \dots [\mathbf{A},}_{r_1} \underbrace{[\mathbf{B}, \dots, [\mathbf{B},}_{s_1} \dots \underbrace{[\mathbf{A}, \dots [\mathbf{A},}_{r_n} \underbrace{[\mathbf{B}, \dots [\mathbf{B}, \mathbf{B}] \dots]}_{s_n}] \dots] \dots] \dots]] \dots],
\end{aligned}$$

which is zero if $s_n > 1$ or if $s_n = 0$ and $r_n > 1$.

The first several terms of the general terms of BCH formula is:

$$\begin{aligned}\text{In}(\exp(\mathbf{A})\exp(\mathbf{B})) &= \exp(\mathbf{A}) + \exp(\mathbf{B}) + \frac{1}{2}[\mathbf{A},\mathbf{B}] \\ &+ \frac{1}{12}[\mathbf{A},[\mathbf{A},\mathbf{B}]] - \frac{1}{12}[\mathbf{B},[\mathbf{A},\mathbf{B}]] \\ &- \frac{1}{24}[\mathbf{B},[\mathbf{A},[\mathbf{A},\mathbf{B}]]] - \frac{1}{720}[[[[\mathbf{A},\mathbf{B}],\mathbf{B},],\mathbf{B}],\mathbf{B}] + \ldots .\end{aligned}$$

If you are feeling totally confused about the BCH formula, that's fine and it's not necessary for a SLAMer to understand every detail of BCH formula, all you have to know is that **BCH forumula establish a linkage between the operations in Lie group and the operations in Lie algebra**. In the context of $\mathrm{SO}(3), \mathrm{SE}(3)$ and $\mathfrak{so}(3), \mathfrak{se}(3)$, BCH formula establish a linkage between the matrix multiplications over $\mathrm{SO}(3), \mathrm{SE}(3)$ and the operations over the vector space of $\mathfrak{so}(3), \mathfrak{se}(3)$.

Compared to the complex original form of BCH formula, it's more common to use the approximation form of BCH formula. If ϕ_1 or ϕ_2 is small, we have the following approximation:

$$\text{In}(\exp(\boldsymbol{\phi}_1^{\wedge})\exp(\boldsymbol{\phi}_2^{\wedge})) \approx \begin{cases}\mathbf{J}_l^{-1}(\boldsymbol{\phi_2}^{\wedge})\boldsymbol{\phi_1}^{\wedge} + \boldsymbol{\phi_2}^{\wedge}, & \text{if } \boldsymbol{\phi_1} \text{ is small} \\ \mathbf{J}_r^{-1}(\boldsymbol{\phi_1}^{\wedge})\boldsymbol{\phi_2}^{\wedge} + \boldsymbol{\phi_1}^{\wedge}, & \text{if } \boldsymbol{\phi_2} \text{ is small}\end{cases}$$

where $\mathbf{J}_l$ is the left Jacobian of $\mathrm{SO}(3)$ which we've mentioned in chapter 5. Suppose $\theta\mathbf{a} = \boldsymbol{\phi} \in \mathrm{SO}(3)$, we have

$$\mathbf{J}_l(\boldsymbol{\phi}) = \frac{\sin\theta}{\theta}\mathbf{I} + (1 - \frac{\sin\theta}{\theta})\mathbf{a}\mathbf{a}^{\mathbf{T}} + \frac{1-\cos\theta}{\theta}\mathbf{a}^{\wedge}$$

and its inverse

$$\mathbf{J}_l^{-1}(\boldsymbol{\phi}) = \frac{\theta}{2}\cot\frac{\theta}{2}\mathbf{I} + (1 - \frac{\theta}{2}\cot\frac{\theta}{2})\mathbf{a}\mathbf{a}^{\mathbf{T}} - \frac{\theta}{2}\mathbf{a}^{\wedge} .$$

As for the right Jacobian of $\mathrm{SO}(3)$, we have

$$\mathbf{J}_r(\boldsymbol{\phi}) = \frac{\sin\theta}{\theta}\mathbf{I} + (1 - \frac{\sin\theta}{\theta})\mathbf{a}\mathbf{a}^{\mathbf{T}} - \frac{1-\cos\theta}{\theta}\mathbf{a}^{\wedge}$$

and its inverse

$$\mathbf{J}_r^{-1}(\boldsymbol{\phi}) = \frac{\theta}{2}\cot\frac{\theta}{2}\mathbf{I} + (1 - \frac{\theta}{2}\cot\frac{\theta}{2})\mathbf{a}\mathbf{a}^{\mathbf{T}} + \frac{\theta}{2}\mathbf{a}^{\wedge} .$$

It's obvious that an relationship between left Jacobian and right Jacobian is that

$$\mathbf{J}_r(\boldsymbol{\phi}) = \mathbf{J}_l(-\boldsymbol{\phi}) \qquad \mathbf{J}_r^{-1}(\boldsymbol{\phi}) = \mathbf{J}_l^{-1}(-\boldsymbol{\phi}) .$$

There also exists another relationship between left Jacobian and right Jacobian

$$\mathbf{J}_l(\boldsymbol{\phi}) = \mathbf{R}\mathbf{J}_r(\boldsymbol{\phi}) .$$

The approximation form of BCH formula gives us an intuition that left(or right) multiplying a tiny rotation in $\mathrm{SO}(3)$ is euqivalent to adding a tiny corresponding vector which is self multiplied by left(or right) Jacobian.

More concretely, suppose we have a tiny rotation ΔR with respect to rotation R and their corresponding vector in Li algebra is $\Delta\boldsymbol{\phi}$, $\boldsymbol{\phi}$ respectively, then we can derive

$$\exp(\Delta\boldsymbol{\phi}^{\wedge})\exp(\boldsymbol{\phi}^{\wedge}) = \exp((\mathbf{J}_l^{-1}(\Delta\boldsymbol{\phi}) + \boldsymbol{\phi})^{\wedge})$$

and

$$\exp(\boldsymbol{\phi}^{\wedge})\exp(\Delta\boldsymbol{\phi}^{\wedge}) = \exp((\mathbf{J}_r^{-1}(\Delta\boldsymbol{\phi}) + \boldsymbol{\phi})) .$$

Reversely, we can derive

$$\exp((\boldsymbol{\phi} + \Delta\boldsymbol{\phi})^{\wedge}) = \exp(\mathbf{J}_l(\Delta\boldsymbol{\phi}))\exp(\boldsymbol{\phi}) = \exp(\boldsymbol{\phi})\exp(\mathbf{J}_r(\Delta\boldsymbol{\phi}))$$

Similiarly, there's an approximation form of BCH formula for $\mathrm{SE}(3)$ and $\mathfrak{se}(3)$:

$$\text{In}(\exp(\boldsymbol{\xi}_1^{\wedge})\exp(\boldsymbol{\xi}_2^{\wedge})) \approx \begin{cases}\boldsymbol{\mathcal{J}}_l^{-1}(\boldsymbol{\xi_2}^{\wedge})\boldsymbol{\xi_1}^{\wedge} + \boldsymbol{\xi_2}^{\wedge}, & \text{if } \boldsymbol{\xi_1} \text{ is small} \\ \boldsymbol{\mathcal{J}}_r^{-1}(\boldsymbol{\xi_1}^{\wedge})\boldsymbol{\xi_2}^{\wedge} + \boldsymbol{\xi_1}^{\wedge}, & \text{if } \boldsymbol{\xi_2} \text{ is small}\end{cases}$$

Here the $\boldsymbol{\mathcal{J}}_l$ and $\boldsymbol{\mathcal{J}}_r$ are 6×6 matrics, which are more complicated. Since these two matrics are less common to use in the calculation, we choose to omit them.

8 Li Algebra derivation and Perturbation Model

Since Lie group is not closed under addition, according to the definition of derivative, it can't be defined directly on Lie group. On the other hand, if we treat the rotation matrics and the transformation matrics as the common matrics, in order to keep the matrics under defined, constraints have to be introduced into the optimization. Thanks to BCH formula, we can solve the problem of derivative using Lie algebra in two ways:

1. Assume we add a infinitesimal amount on Lie algebra, then compute the change of the object function

2. Assume we left(or right) multiply an infinitesimal perturbation on the Lie group and use Lie algebra to describe the perturbation, then compute the derivative on this perturbation.

8.1 Derivative Model

Suppose we have a point $\mathbf{p}$ and a rotation $\mathbf{R}$ with corresponding Lie algebra vector $\boldsymbol{\phi}$, and then calculate the derivative of $\mathbf{Rp}$ with respect to $\boldsymbol{\phi}$

$$\frac{\partial(\mathbf{Rp})}{\partial \boldsymbol{\phi}} .$$

According to the definition of derivative, we have

$$\begin{aligned}
\frac{\partial(\mathbf{Rp})}{\partial \boldsymbol{\phi}} &= \frac{\partial(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})}{\partial \boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp((\boldsymbol{\phi}+\Delta\boldsymbol{\phi})^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp((\mathbf{J}_l(\boldsymbol{\phi})\Delta\boldsymbol{\phi})^\wedge)\exp(\boldsymbol{\phi}^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{(\mathbf{I} + (\mathbf{J}_l(\boldsymbol{\phi})\Delta\boldsymbol{\phi})^\wedge)\exp(\boldsymbol{\phi}^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{(\mathbf{J}_l(\boldsymbol{\phi})\Delta\boldsymbol{\phi})^\wedge\exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{-(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})^\wedge(\mathbf{J}_l(\boldsymbol{\phi})\Delta\boldsymbol{\phi})}{\Delta\boldsymbol{\phi}} \\
&= -(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})^\wedge\mathbf{J}_l(\boldsymbol{\phi}) \\
&= -(\mathbf{Rp})^\wedge\mathbf{J}_l(\boldsymbol{\phi})
\end{aligned}$$

However, this's not the most streamlined way to optimize $f(\mathbf{R})$. Firstly, we have to store the rotation as a rotation vector, which has the problem of singularities. Secondly, we have to compute the Jacobian matrix.

8.2 Perturbation Model

A cleaner way to optimize $f(\mathbf{R})$ is to find an update step for $\mathbf{R}$ in the form of a tiny rotation on the left rather than directly on the Lie algebra rotation vector. Suppose we have a point $\mathbf{p}$ and a rotation $\mathbf{R}$ with corresponding Lie algebra vector $\boldsymbol{\phi}$, and then calculate the derivative of $\mathbf{Rp}$ with respect to

$\boldsymbol{\phi}$ as follow:

$$
\begin{aligned}
\frac{\partial(\mathbf{R}\mathbf{p})}{\partial\boldsymbol{\phi}} &= \frac{\partial(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})}{\partial\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp((\Delta\boldsymbol{\phi})^\wedge)\exp(\phi^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\boldsymbol{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{(\mathbf{I} + (\Delta\boldsymbol{\phi})^\wedge)\exp(\boldsymbol{\phi}^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{(\Delta\boldsymbol{\phi})^\wedge \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{-(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})^\wedge \Delta\boldsymbol{\phi}}{\Delta\boldsymbol{\phi}} \\
&= -(\mathbf{R}\mathbf{p})^\wedge \, .
\end{aligned}
$$

Since the perturbation functions through left multiplication, optimizing in above way is called left perturbation model. On the other hand, there's the right perturbation model,

$$
\begin{aligned}
\frac{\partial(\mathbf{R}\mathbf{p})}{\partial\boldsymbol{\phi}} &= \frac{\partial(\exp(\boldsymbol{\phi}^\wedge)\mathbf{p})}{\partial\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp(\boldsymbol{\phi}^\wedge)\exp((\Delta\boldsymbol{\phi})^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp(\boldsymbol{\phi}^\wedge)(\mathbf{I} + (\Delta\boldsymbol{\phi})^\wedge)\mathbf{p} - \exp(\boldsymbol{\phi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{\exp(\boldsymbol{\phi}^\wedge)(\Delta\boldsymbol{\phi})^\wedge\mathbf{p}}{\Delta\boldsymbol{\phi}} \\
&= \lim_{\Delta\boldsymbol{\phi}\to 0} \frac{-\exp(\boldsymbol{\phi}^\wedge)\mathbf{p}^\wedge \Delta\boldsymbol{\phi}}{\Delta\boldsymbol{\phi}} \\
&= -\mathbf{R}(\mathbf{p}^\wedge) \, .
\end{aligned}
$$

All we've talked is the perturbation model in $\mathrm{SO}(3)$. Following the same ideology, it's easy to extend perturbation model to $\mathrm{SE}(3)$. Suppose a point $\mathbf{p}$ and a transformation matrix $\mathbf{T}$ with its corresponding Lie algebra vector $\boldsymbol{\xi}$. To optimize $\mathbf{T}\mathbf{p}$ with respect to $\boldsymbol{\xi}$ using left perturbation model, we have

$$
\begin{aligned}
\frac{\partial(\mathbf{T}\mathbf{p})}{\partial\boldsymbol{\xi}} &= \frac{\partial(\exp(\boldsymbol{\xi}^\wedge)\mathbf{p})}{\partial\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\exp((\Delta\boldsymbol{\xi})^\wedge)\exp(\boldsymbol{\xi}^\wedge)\mathbf{p} - \exp(\boldsymbol{\xi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{(\mathbf{I} + (\Delta\boldsymbol{\xi})^\wedge)\exp(\boldsymbol{\xi}^\wedge)\mathbf{p} - \exp(\boldsymbol{\xi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{(\Delta\boldsymbol{\xi})^\wedge \exp(\boldsymbol{\xi}^\wedge)\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\begin{bmatrix} (\Delta\boldsymbol{\phi})^\wedge & \Delta\boldsymbol{\rho} \\ \mathbf{0}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}\mathbf{p} + \mathbf{t} \\ 1 \end{bmatrix}}{[\Delta\boldsymbol{\rho} \quad \Delta\boldsymbol{\phi}]^T} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\begin{bmatrix} (\Delta\boldsymbol{\phi})^\wedge(\mathbf{R}\mathbf{p} + \mathbf{t}) + \Delta\boldsymbol{\rho} \\ \mathbf{0}^T \end{bmatrix}}{[\Delta\boldsymbol{\rho} \quad \Delta\boldsymbol{\phi}]^T} \\
&= \begin{bmatrix} \mathbf{I} & -(\mathbf{R}\mathbf{p} + \mathbf{t})^\wedge \\ \mathbf{0^T} & \mathbf{0^T} \end{bmatrix} \overset{def}{=} (\mathbf{T}\mathbf{p})^\odot
\end{aligned}
$$

where we define the operator $(\cdot)^{\odot}$, which transforms a spatial point of homogeneous coordinates into a matrix of 4×6.

In addition, to optimize $\mathbf{Tp}$ with respect to $\boldsymbol{\xi}$ using right perturbation model, we have

$$\begin{aligned}
\frac{\partial(\mathbf{Tp})}{\partial\boldsymbol{\xi}} &= \frac{\partial(\exp(\boldsymbol{\xi}^{\wedge})\mathbf{p})}{\partial\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\exp(\boldsymbol{\xi}^{\wedge})\exp((\Delta\boldsymbol{\xi})^{\wedge})\mathbf{p} - \exp(\boldsymbol{\xi}^{\wedge})\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\exp(\boldsymbol{\xi}^{\wedge})(\mathbf{I} + (\Delta\boldsymbol{\xi})^{\wedge})\mathbf{p} - \exp(\boldsymbol{\xi}^{\wedge})\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\exp(\boldsymbol{\xi}^{\wedge})(\Delta\boldsymbol{\xi})^{\wedge}\mathbf{p}}{\Delta\boldsymbol{\xi}} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 0 \end{bmatrix} \begin{bmatrix} (\Delta\boldsymbol{\phi})^{\wedge}\mathbf{p} + \Delta\boldsymbol{\rho} \\ 1 \end{bmatrix}}{\begin{bmatrix} \Delta\boldsymbol{\rho} & \Delta\boldsymbol{\phi} \end{bmatrix}^T} \\
&= \lim_{\Delta\boldsymbol{\xi}\to 0} \frac{\begin{bmatrix} \mathbf{R}((\Delta\boldsymbol{\phi})^{\wedge}\mathbf{p} + \Delta\boldsymbol{\rho}) + \mathbf{t} \\ \mathbf{0}^T \end{bmatrix}}{\begin{bmatrix} \Delta\boldsymbol{\rho} & \Delta\boldsymbol{\phi} \end{bmatrix}^T} \\
&= \begin{bmatrix} \mathbf{R} & -\mathbf{R}(\mathbf{p}^{\wedge}) \\ \mathbf{0^T} & \mathbf{0^T} \end{bmatrix} .
\end{aligned}$$

Above equations require further explanation about matrix differentiation. Assume that $\mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}$ are column vectors, there are following rules:

$$\frac{\mathrm{d}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}}{\mathrm{d}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}} = \left(\frac{\mathrm{d}\begin{bmatrix} \mathbf{a} & \mathbf{b} \end{bmatrix}^T}{\mathrm{d}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}} \right)^T = \begin{bmatrix} \frac{\mathrm{d}\mathbf{a}}{\mathrm{d}\mathbf{x}} & \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}\mathbf{x}} \\ \frac{\mathrm{d}\mathbf{a}}{\mathrm{d}\mathbf{y}} & \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}\mathbf{y}} \end{bmatrix}^T = \begin{bmatrix} \frac{\mathrm{d}\mathbf{a}}{\mathrm{d}\mathbf{x}} & \frac{\mathrm{d}\mathbf{a}}{\mathrm{d}\mathbf{y}} \\ \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}\mathbf{x}} & \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}\mathbf{y}} \end{bmatrix} .$$

References

[1] Gao, Xiang, et al. "14 lectures on visual SLAM: from theory to practice." Publishing House of Electronics Industry, Beijing (2017).

[2] Barfoot, Timothy D. State estimation for robotics. Cambridge University Press, 2017.