Lie Group Doc

Vassili Korotkine
Department of Mechanical Engineering, McGill University
817 Sherbrooke Street West, Montreal QC H3A 0C3

March 19, 2025

Contents

1 Introduction to Lie groups **1**
1.1 Plus and Minus Operations 3

2 Notation and Preliminaries **3**

3 Plus and Minus Jacobians **4**
3.1 Left and Right Group Jacobian 5
3.2 Left Perturbation 5
3.3 Right Perturbation 6
3.4 Summary 6

4 Batch **7**

5 The Adjoint Operator **8**

1 Introduction to Lie groups

States involving rotations are not a vectorspace, which means that the notions of addition and subtraction must be reconsidered to be able to do statistics and calculus that are required for state estimation. They form a *manifold*, which is is a lower dimensional space subject to a constraint. For instance, a rotation matrix in 2D, $\mathbf{C}_{ab}$, has four elements but only one degree of freedom, the angle of rotation. The manifold is a line in four-dimensional space. The constraint is $\mathbf{C}^{\mathsf{T}}\mathbf{C} = \mathbf{1}$. Naively adding Lie group elements does not work. To do addition, velocity is considered which lives in the tangent space. The structure of the velocity is obtained by differentiating the group constraint. For the $SO(2)$ example,

$$\mathbf{C}^{\mathsf{T}}\mathbf{C} = \mathbf{1}, \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{C}^{\mathsf{T}}\mathbf{C}) = \mathbf{0}, \tag{2}$$

$$\dot{\mathbf{C}}^{\mathsf{T}}\mathbf{C} + \mathbf{C}^{\mathsf{T}}\dot{\mathbf{C}} = \mathbf{0}, \tag{3}$$

$$\boldsymbol{\omega}^{\wedge} = \dot{\mathbf{C}}^{\mathsf{T}}\mathbf{C} = -\mathbf{C}^{\mathsf{T}}\dot{\mathbf{C}} \tag{4}$$

$$\dot{\mathbf{C}} = \mathbf{C}\boldsymbol{\omega}^{\wedge}. \tag{5}$$

where $\boldsymbol{\omega}^{\wedge}$ is the angular velocity and is skew-symmetric. Importantly,the skew-symmetric matrix space is *linear*, and addition and subtraction are allowed. This applies to more complicated Lie groups that all have similar tangent spaces that are linear. For a more general Lie group,

$$\dot{\mathcal{X}} = \mathcal{X}\mathbf{v}^{\wedge}, \tag{6}$$

where $\mathbf{v}$ has a more complicated structure. The *linearity* of $\mathbf{v}^{\wedge}$, which is a matrix, means that that the Lie algebra space may be identified with $\mathbb{R}^n$ and each element described by a column vector $\mathbf{v} = (\mathbf{v}^{\wedge})^{\vee}$.

To relate the tangent space back to the Lie group, the resulting differential equation is integrated. For the general case the *exact* solution to (6) is given by

$$\mathcal{X}(t) = \mathcal{X}(0)\exp(t\mathbf{v}^{\wedge}). \tag{7}$$

For matrices, the exponential map is given by,

$$\exp(\mathbf{A}) = \sum_{i=0}^{\infty}\frac{1}{i!}\mathbf{A}^{i}. \tag{8}$$

For the $SO(2)$ example,

$$\dot{\mathbf{C}}_{ab} = \mathbf{C}_{ab}\omega_b^{ba^{\times}}, \tag{9}$$

where the angular velocity ω is a scalar and the subscripts and superscripts denote reference frames, the *exact* solution to (9) is given by

$$\mathbf{C}_{ab}(t_k) = \mathbf{C}_{ab}(t_{k-1})\exp(\omega_b^{ba^{\times}}(t_k - t_{k-1})) \tag{10}$$

$$= \mathbf{C}_{ab}(t_{k-1})\exp(\omega_b^{ba^{\times}}\Delta t), \tag{11}$$

In the $SO(2)$ example, the angular velocity is simply the rate of rotation of the angle,

$$\omega = \dot{\theta}, \tag{12}$$

and the $\wedge$ operator is the cross operator,

$$\omega^{\wedge} = \omega^{\times} = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}. \tag{13}$$

1.1 Plus and Minus Operations

Generally, define the Lie algebra increment $\boldsymbol{\tau}$ that corresponds to $t\mathbf{v}$. The analog of addition and subtraction for a Lie group is then given through the "Plus" and "minus" operations. "Plus" and "minus" operations allow us to define increments between elements of the curved manifold, and express them in the flat vector spaces. The $\oplus : G \times \mathbb{R}^m \to G$ operator defines how elements in $\mathbb{R}^m$ are used to increment elements in G. Similarly, the $\ominus : G \times G \to \mathbb{R}^m$ operator defines a difference between two group elements. These can both be defined in a left-or right manner, where the right $\ominus$ and $\oplus$ operators are defined as

$$\text{right} - \oplus : \quad \mathcal{Y} = \mathcal{X} \oplus \boldsymbol{\tau} \triangleq \mathcal{X} \circ \operatorname{Exp}(\boldsymbol{\tau}), \tag{14}$$

$$\text{right} - \ominus : \quad \boldsymbol{\tau} = \mathcal{Y} \ominus \mathcal{X} \triangleq \operatorname{Log}\left(\mathcal{X}^{-1} \circ \mathcal{Y}\right). \tag{15}$$

Similarly, the left operators are defined as

$$\text{left} - \oplus : \quad \mathcal{Y} = \mathcal{X} \oplus \boldsymbol{\tau} \triangleq \operatorname{Exp}(\boldsymbol{\tau}) \circ \mathcal{X}, \tag{16}$$

$$\text{left} - \ominus : \quad \boldsymbol{\tau} = \mathcal{Y} \ominus \mathcal{X} \triangleq \operatorname{Log}\left(\mathcal{Y} \circ \mathcal{X}^{-1}\right). \tag{17}$$

2 Notation and Preliminaries

A *Lie group* G is a smooth manifold whose elements satisfy the group axioms.. For any G, there exists an associated Lie algebra $\mathfrak{g}$, and a vector space identifiable with elements of $\mathbb{R}^m$, where m is the degrees of freedom (DoF) of G. The *exponential and logarithmic maps* are denoted $\exp : \mathfrak{g} \to G$ and $\log : G \to \mathfrak{g}$. The *vee and wedge operators* are denoted $(\cdot)^\vee : \mathfrak{g} \to \mathbb{R}^m$ and $(\cdot)^\wedge : \mathbb{R}^m \to \mathfrak{g}$. For convenience, the functions $\operatorname{Exp} : \mathbb{R}^m \to G$ and $\operatorname{Log} : G \to \mathbb{R}^m$ are additionally defined. For $\mathcal{X} \in G$ and $\boldsymbol{\tau} \in \mathbb{R}^m$,

$$\mathcal{X} \triangleq \exp(\boldsymbol{\tau}^\wedge) \triangleq \operatorname{Exp}(\boldsymbol{\tau}), \quad \boldsymbol{\tau} \triangleq \log(\mathcal{X})^\vee \triangleq \operatorname{Log}(\mathcal{X}). \tag{18}$$

"Plus" and "minus" operations allow us to define increments between elements of the curved manifold, and express them in the flat vector spaces. The $\oplus : G \times \mathbb{R}^m \to G$ operator defines how elements in $\mathbb{R}^m$ are used to increment elements in G. Similarly, the $\ominus : G \times G \to \mathbb{R}^m$ operator defines a difference between two group elements. These can both be defined in a left-or right manner, where the right $\ominus$ and $\oplus$ operators are defined as

$$\text{right} - \oplus : \quad \mathcal{Y} = \mathcal{X} \oplus \boldsymbol{\tau} \triangleq \mathcal{X} \circ \operatorname{Exp}(\boldsymbol{\tau}), \tag{19}$$

$$\text{right} - \ominus : \quad \boldsymbol{\tau} = \mathcal{Y} \ominus \mathcal{X} \triangleq \operatorname{Log}\left(\mathcal{X}^{-1} \circ \mathcal{Y}\right). \tag{20}$$

Similarly, the left operators are defined as

$$\text{left} - \oplus : \quad \mathcal{Y} = \mathcal{X} \oplus \boldsymbol{\tau} \triangleq \operatorname{Exp}(\boldsymbol{\tau}) \circ \mathcal{X}, \tag{21}$$

$$\text{left} - \ominus : \quad \boldsymbol{\tau} = \mathcal{Y} \ominus \mathcal{X} \triangleq \operatorname{Log}\left(\mathcal{Y} \circ \mathcal{X}^{-1}\right). \tag{22}$$

$$\frac{Df(\mathcal{X})}{D\mathcal{X}} = \lim_{\boldsymbol{\tau} \to \mathbf{0}} \frac{f(\mathcal{X}) \oplus \boldsymbol{\tau} \ominus f(\mathcal{X})}{\boldsymbol{\tau}}, \tag{23}$$

which for right and left definitions of $\oplus$ yield right and left Jacobians of f respectively. The right and left *group Jacobians* on a manifold $\mathcal{M}$ are defined as the derivatives of the exponential map,

$$\mathbf{J}(\boldsymbol{\tau}) = \frac{D\,\mathrm{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}}, \tag{24}$$

where the right Jacobian is obtained by using the right derivative and left Jacobian is obtained by using the left derivative. To obtain the Baker-Campbell-Hausdorff approximations, the Lie group derivative definition (23) is written for small $\boldsymbol{\tau}$ as

$$\frac{Df(\mathcal{X})}{D\mathcal{X}} \approx \frac{f(\mathcal{X}) \oplus \boldsymbol{\tau} \ominus f(\mathcal{X})}{\boldsymbol{\tau}} \tag{25}$$

$$f(\mathcal{X}) \oplus \frac{Df(\mathcal{X})}{D\mathcal{X}}\boldsymbol{\tau} \approx f(\mathcal{X} \oplus \boldsymbol{\tau}), \tag{26}$$

whereupon $f(\mathcal{X})$ is set to be the exponential map $f(\boldsymbol{\tau}) = \mathrm{Exp}(\boldsymbol{\tau}) : \mathbb{R}^m \to G$, in other words the argument $\mathcal{X}$ is the vector $\boldsymbol{\tau}$ mapped to the manifold. This yields

$$f(\mathcal{X}) \oplus \frac{Df(\mathcal{X})}{D\mathcal{X}}\boldsymbol{\tau} \approx f(\mathcal{X} \oplus \boldsymbol{\tau}), \tag{27}$$

$$\mathrm{Exp}(\boldsymbol{\tau}) \oplus \frac{D\,\mathrm{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}}\delta\boldsymbol{\tau} \approx \mathrm{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}), \tag{28}$$

$$\mathrm{Exp}(\boldsymbol{\tau}) \oplus \mathbf{J}(\boldsymbol{\tau})\delta\boldsymbol{\tau} \approx \mathrm{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}), \tag{29}$$

which is direction-agnostic as long as the perturbations used for the $\oplus, \ominus$ and the group Jacobians are consistent. Manipulating (29) yields the Baker-Campbell-Hausdorff (BCH) linear approximation,

$$\mathrm{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}) \approx \mathrm{Exp}(\boldsymbol{\tau}) \oplus \mathbf{J}(\boldsymbol{\tau})\delta\boldsymbol{\tau}, \tag{30}$$

$$\mathrm{Exp}(\boldsymbol{\tau}) \oplus \delta\boldsymbol{\tau} \approx \mathrm{Exp}(\boldsymbol{\tau} + \mathbf{J}^{-1}\delta\boldsymbol{\tau}), \tag{31}$$

$$\mathrm{Log}(\mathrm{Exp}(\boldsymbol{\tau}) \oplus \delta\boldsymbol{\tau}) \approx \boldsymbol{\tau} + \mathbf{J}^{-1}\delta\boldsymbol{\tau}. \tag{32}$$

The BCH linear approximation is illustrated in Fig. 1.

3 Plus and Minus Jacobians

Jacobians of the following operations are considered. The plus Jacobian,

$$\mathcal{Y} \oplus \delta\mathbf{y} = \mathcal{X} \oplus (\boldsymbol{\tau} + \delta\boldsymbol{\tau}) \tag{33}$$

$$\approx (\mathcal{X} \oplus \boldsymbol{\tau}) \oplus \frac{D(\mathcal{X} \oplus \boldsymbol{\tau})}{D\boldsymbol{\tau}}\delta\boldsymbol{\tau}, \tag{34}$$

as well as the minus Jacobian with respect to both inputs,

$$(\mathcal{Y} \oplus \delta\boldsymbol{\tau}) \ominus \mathcal{X} \approx (\mathcal{Y} \ominus \mathcal{X}) + \frac{D\mathcal{Y} \ominus \mathcal{X}}{D\mathcal{Y}}\delta\boldsymbol{\tau}, \tag{35}$$

$$\mathcal{Y} \ominus (\mathcal{X} \oplus \delta\boldsymbol{\tau}) \approx (\mathcal{Y} \ominus \mathcal{X}) + \frac{D\mathcal{Y} \ominus \mathcal{X}}{D\mathcal{X}}\delta\boldsymbol{\tau}, \tag{36}$$

where $\frac{D(\mathcal{X} \oplus \boldsymbol{\tau})}{D\boldsymbol{\tau}}$, $\frac{D\mathcal{Y} \ominus \mathcal{X}}{D\mathcal{Y}}$ and $\frac{D\mathcal{Y} \ominus \mathcal{X}}{D\mathcal{X}}$ are sought.

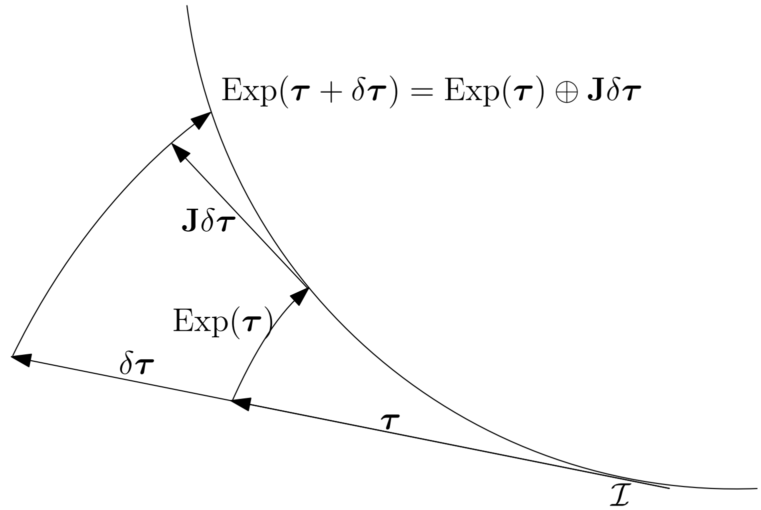

Figure 1: Illustration of the BCH linear approximation. The difference is between travelling the whole way in the tangent space versus travelling most of the way, projecting to the Lie group, and travelling the rest of the way in the tangent space expanded at that intermediate point. The Jacobian in some sense compensates for the curvature of the manifold.

3.1 Left and Right Group Jacobian

From Sola, the left and right group Jacobians are related by

$$\mathbf{J}_{R,-\boldsymbol{\tau}} = \mathbf{J}_{L,\boldsymbol{\tau}} \tag{37}$$

3.2 Left Perturbation

Perturbing both $\mathcal{Y}$ and $\mathcal{X}$ yields

$$\begin{aligned}
(\mathcal{Y} \oplus \delta\mathbf{y}) \ominus (\mathcal{X} \oplus \delta\mathbf{x}) &= \mathrm{Log}((\mathcal{Y} \oplus \delta\mathbf{y})(\mathcal{X} \oplus \delta\mathbf{x})^{-1}) && (38)\\
&= \mathrm{Log}(\mathrm{Exp}(\delta\mathbf{y})\mathcal{Y}\mathcal{X}^{-1}\,\mathrm{Exp}(-\delta\mathbf{x})) && (39)\\
&= \mathrm{Log}((\mathcal{Y}\mathcal{X}^{-1} \oplus_L \delta\mathbf{y})\ \oplus_R (-\delta\mathbf{x})). && (40)
\end{aligned}$$

Applying BCH (32) twice with the corresponding Jacobian directions yields the following expressions, where $\mathbf{J}_{L,\boldsymbol{\tau}}$ and $\mathbf{J}_{R,\boldsymbol{\tau}}$ are respectively the left and right Jacobians evaluated at $\boldsymbol{\tau}$.

$$\begin{aligned}
(\mathcal{Y} \oplus \delta\mathbf{y}) \ominus (\mathcal{X} \oplus \delta\mathbf{x}) &= \mathrm{Log}(\mathcal{Y}\mathcal{X}^{-1}) - \mathbf{J}^{-1}_{R,(\mathcal{Y}\mathcal{X}^{-1}\oplus_L\delta\mathbf{y})}\delta\mathbf{x} + \mathbf{J}^{-1}_{L,(\mathcal{Y}\mathcal{X}^{-1})}\delta\mathbf{y} && (41)\\
&\approx \mathrm{Log}(\mathcal{Y}\mathcal{X}^{-1}) - \mathbf{J}^{-1}_{R,(\mathcal{Y}\mathcal{X}^{-1})}\delta\mathbf{x} + \mathbf{J}^{-1}_{L,(\mathcal{Y}\mathcal{X}^{-1})}\delta\mathbf{y} && (42)\\
&= \mathrm{Log}(\mathcal{Y}\mathcal{X}^{-1}) - \mathbf{J}^{-1}_{L,(-\mathcal{Y}\mathcal{X}^{-1})}\delta\mathbf{x} + \mathbf{J}^{-1}_{L,(\mathcal{Y}\mathcal{X}^{-1})}\delta\mathbf{y} && (43)\\
&= \mathcal{Y} \ominus \mathcal{X} - \mathbf{J}^{-1}_{-\mathcal{Y}\ominus\mathcal{X}}\delta\mathbf{x} + \mathbf{J}^{-1}_{\mathcal{Y}\ominus\mathcal{X}}\delta\mathbf{y} && (44)\\
&= \mathcal{Y} \ominus \mathcal{X} + \mathbf{J}^{-1}_{\mathcal{Y}\ominus\mathcal{X}}\delta\mathbf{y} - \mathbf{J}^{-1}_{-\mathcal{Y}\ominus\mathcal{X}}\delta\mathbf{x}, && (45)
\end{aligned}$$

yielding the minus Jacobians with respect to both arguments in terms of the left group Jacobian.

3.3 Right Perturbation

A perturbation in $\boldsymbol{\tau}$ that causes a perturbation in $\mathcal{Y}$,

$$\begin{aligned}
\mathcal{Y} &= \mathcal{X} \oplus \boldsymbol{\tau}, && (46)\\
\mathcal{Y} \oplus \delta\mathbf{y} &= \mathcal{X} \oplus (\boldsymbol{\tau} + \delta\boldsymbol{\tau}) && (47)\\
&= \mathcal{X}\,\mathrm{Exp}(\boldsymbol{\tau} + \delta\boldsymbol{\tau}) && (48)\\
&\approx \mathcal{X}\,\mathrm{Exp}(\boldsymbol{\tau}) \oplus \mathbf{J}_{\boldsymbol{\tau}}\delta\boldsymbol{\tau} && (49)\\
&= \mathcal{Y} \oplus \mathbf{J}_{\boldsymbol{\tau}}\delta\boldsymbol{\tau}, && (50)\\
\delta\mathbf{y} &= \mathbf{J}_{\boldsymbol{\tau}}\delta\boldsymbol{\tau}, && (51)
\end{aligned}$$

which means the Jacobian of the plus operation is the same as the group Jacobian. Analogously, the Jacobian for the minus operation is obtained by considering the minus operation output,

$$\begin{aligned}
\boldsymbol{\tau} &= \mathcal{Y} \ominus \mathcal{X}, && (52)\\
\boldsymbol{\tau} + \delta\boldsymbol{\tau} &= (\mathcal{Y} \oplus \delta\mathbf{y}) \ominus \mathcal{X} && (53)\\
&= \mathrm{Log}(\mathcal{X}^{-1}\mathcal{Y}\,\mathrm{Exp}\,\delta\mathbf{y}) && (54)\\
&= \mathrm{Log}(\mathrm{Exp}\,\boldsymbol{\tau}\,\mathrm{Exp}\,\delta\mathbf{y}) && (55)\\
&\approx \boldsymbol{\tau} + \mathbf{J}|_{\boldsymbol{\tau}}^{-1}\,\delta\mathbf{y} && (56)\\
&= \boldsymbol{\tau} + \mathbf{J}|_{\mathcal{Y}\ominus\mathcal{X}}^{-1}\,\delta\mathbf{y} && (57)
\end{aligned}$$

such that the Jacobian of the minus operation w.r.t. $\mathcal{Y}$ is given by the inverse of the group Jacobian. The Jacobian w.r.t. $\mathcal{X}$ is derived as

$$\begin{aligned}
\boldsymbol{\tau} + \delta\boldsymbol{\tau} &= \mathcal{Y} \ominus (\mathcal{X} \oplus \delta\boldsymbol{\xi}) && (58)\\
&= \mathrm{Log}((\mathcal{X} \oplus \delta\boldsymbol{\xi})^{-1}\mathcal{Y}) && (59)\\
&= \mathrm{Log}(\mathrm{Exp}(-\delta\boldsymbol{\xi})\mathcal{X}^{-1}\mathcal{Y}) && (60)\\
&= \mathrm{Log}(\mathrm{Exp}(-\delta\boldsymbol{\xi})\mathcal{X}^{-1}\mathcal{Y}) && (61)\\
&\approx \boldsymbol{\tau} - \mathbf{J}_{l,\mathcal{Y}\ominus\mathcal{X}}^{-1}\delta\boldsymbol{\xi} && (62)\\
&= \boldsymbol{\tau} - \mathbf{J}_{-\boldsymbol{\tau}}^{-1}\delta\boldsymbol{\xi}. && (63)
\end{aligned}$$

3.4 Summary

The group Jacobian evaluated at $\boldsymbol{\tau}$ is denoted $\mathbf{J}|_{\boldsymbol{\tau}}$. For both the left and right perturbation,

$$\begin{aligned}
\frac{D(\mathcal{X} \oplus \boldsymbol{\tau})}{D\boldsymbol{\tau}} &= \mathbf{J}|_{\boldsymbol{\tau}} && (64)\\
\frac{D(\mathcal{Y} \ominus \mathcal{X})}{D\mathcal{Y}} &= \mathbf{J}_{\mathcal{Y}\ominus\mathcal{X}}^{-1} && (65)\\
\frac{D(\mathcal{Y} \ominus \mathcal{X})}{D\mathcal{X}} &= -\mathbf{J}_{-\mathcal{Y}\ominus\mathcal{X}}^{-1} && (66)
\end{aligned}$$

4 Batch

As the vectorspace case is a special case of the Lie group case, only the Lie group case is derived here. The process model is typically written

$$\mathcal{X}_{k+1} = f(\mathcal{X}_k, \mathbf{u}_k) \oplus \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \tag{67}$$

and the corresponding error for the process model at time index k,

$$\mathbf{e}_{p,k}(\mathcal{X}_k, \mathcal{X}_{k+1}) = \mathcal{X}_{k+1} \ominus f(\mathcal{X}_k, \mathbf{u}_k). \tag{68}$$

Generally measurements are in a vectorspace, but sometimes, such as in object-level SLAM they may be Lie group elements which gives a measurement model of the form

$$\mathcal{Y}_k = g(\mathcal{X}_k) \oplus \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k). \tag{69}$$

The corresponding error is then

$$\mathbf{e}_{y,k} = \mathcal{Y}_k \ominus g(\mathcal{X}_k). \tag{70}$$

For measurements that lie in a vectorspace, $\oplus$ and $\ominus$ reduce to standard addition and subtraction.

The overall state $\mathcal{X}$ is defined on the composite manifold [1] as the concatenation of the manifolds corresponding to each of the substates. The overall error $\mathbf{e} = \mathbf{e}(\mathcal{X})$ is defined by stacking all the suberrors into one column matrix. The corresponding nonlinear least squares problem is given by

$$J(\mathbf{e}(\mathcal{X})) = \mathbf{e}^\mathsf{T}\mathbf{W}\mathbf{e}, \tag{71}$$

where $\mathbf{e} = \mathbf{e}(\mathcal{X})$ is a nonlinear function of the state $\mathcal{X}$.

Denote the Lie Jacobian of $\mathbf{e}$ as $\mathbf{H} = \frac{D\mathbf{e}}{D\mathcal{X}}$ and $\bar{\mathbf{e}}$ as the error at a linearization point $\bar{\mathcal{X}}$, $\bar{\mathbf{e}} = \mathbf{e}(\bar{\mathcal{X}})$. The Gauss-Newton method for solving this problem may be derived through linearization of $\mathbf{e}$,

$$\begin{aligned} \mathbf{e}(\mathcal{X}) &= \mathbf{e}(\bar{\mathcal{X}} \oplus \delta\boldsymbol{\xi}) && (72)\\ &\approx \bar{\mathbf{e}} + \mathbf{H}\delta\boldsymbol{\xi} && (73) \end{aligned}$$

which gives a quadratic approximation, in $\delta\boldsymbol{\xi}$ to the loss function $J(\mathbf{e}(\mathcal{X}))$,

$$\begin{aligned} \bar{J}(\mathbf{e}(\mathcal{X}(\bar{\mathcal{X}}, \delta\boldsymbol{\xi}))) &= \mathbf{e}^\mathsf{T}\mathbf{W}\mathbf{e} && (74)\\ &\approx (\bar{\mathbf{e}} + \mathbf{H}\delta\boldsymbol{\xi})^\mathsf{T}\, \mathbf{W}\, (\bar{\mathbf{e}} + \mathbf{H}\delta\boldsymbol{\xi}) && (75)\\ &= \bar{\mathbf{e}}\mathbf{W}\bar{\mathbf{e}} + 2\delta\boldsymbol{\xi}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{W}\bar{\mathbf{e}} + \delta\boldsymbol{\xi}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{W}\mathbf{H}\delta\boldsymbol{\xi}. && (76) \end{aligned}$$

Once this approximation is made, the optimization is now over $\delta\boldsymbol{\xi}$. In other words, the perturbation $\delta\boldsymbol{\xi}$ to the current linearization state $\bar{\mathcal{X}}$ is sought such that $\bar{J}(\mathbf{e}(\mathcal{X}(\bar{\mathcal{X}}, \delta\boldsymbol{\xi})))$ is minimized. The term $\bar{\mathbf{e}}\mathbf{W}\bar{\mathbf{e}}$ does not affect the minimum of this approximation such that

$$\arg\min_{\delta\boldsymbol{\xi}} \bar{J}(\mathbf{e}(\mathcal{X})) = \arg\min_{\delta\boldsymbol{\xi}} 2\delta\boldsymbol{\xi}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{W}\bar{\mathbf{e}} + \delta\boldsymbol{\xi}^\mathsf{T}\mathbf{H}^\mathsf{T}\mathbf{W}\mathbf{H}\delta\boldsymbol{\xi}. \tag{77}$$

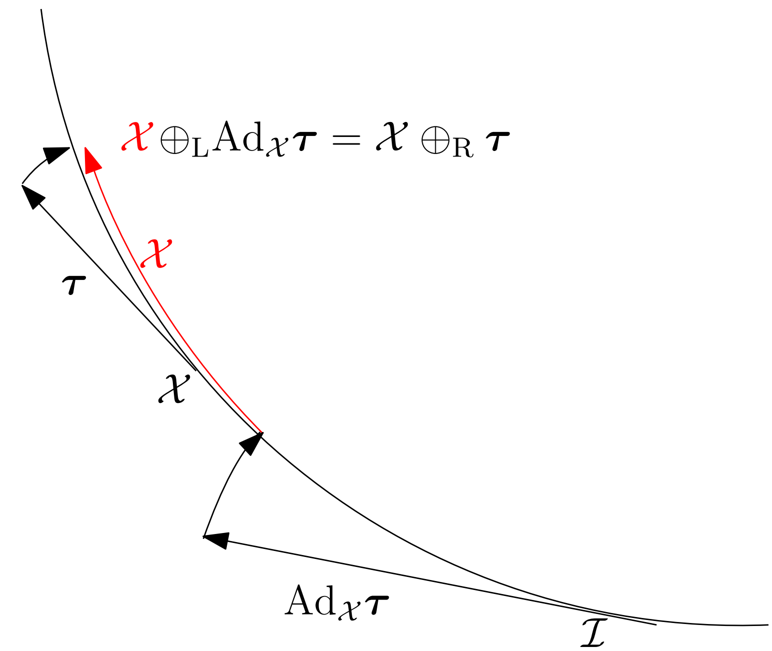

Figure 2: Illustration of the adjoint operator in how it relates the left and right plus operations on the manifold.

Taking the gradient w.r.t. $\delta \boldsymbol{\xi}$ and setting to zero yields

$$\frac{\partial}{\partial \delta \boldsymbol{\xi}} \left(2 \delta \boldsymbol{\xi}^{\mathsf{T}} \mathbf{H}^{\mathsf{T}} \mathbf{W} \bar{\mathbf{e}} + \delta \boldsymbol{\xi}^{\mathsf{T}} \mathbf{H}^{\mathsf{T}} \mathbf{W} \mathbf{H} \delta \boldsymbol{\xi} \right) = 2 \mathbf{H}^{\mathsf{T}} \mathbf{W} \bar{\mathbf{e}} + 2 \mathbf{H}^{\mathsf{T}} \mathbf{W} \mathbf{H} \delta \boldsymbol{\xi} \tag{78}$$

$$= \mathbf{0}, \tag{79}$$

$$\delta \boldsymbol{\xi} = -(\mathbf{H}^{\mathsf{T}} \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^{\mathsf{T}} \mathbf{W} \bar{\mathbf{e}}, \tag{80}$$

where in practice $\mathbf{H}^{\mathsf{T}} \mathbf{W} \mathbf{H}$ is never actually inverted, but rather the system of equations solved for $\delta \boldsymbol{\xi}$. The updated iterate $\hat{\mathcal{X}}$ is set to

$$\hat{\mathcal{X}} = \bar{\mathcal{X}} \oplus \delta \boldsymbol{\xi}, \tag{81}$$

as per the starting point of the error linearization in (72), and the process is iterated until convergence.

5 The Adjoint Operator

The adjoint operator allows to relate the left and right perturbations

$$\mathcal{X} \oplus_{\mathrm{R}} \boldsymbol{\tau} = \mathcal{X} \oplus_{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}} \boldsymbol{\tau} \tag{82}$$

depicted in in Fig. 2. The relationship between left and right Jacobians of a function $f(\mathcal{X}) : \mathcal{M} \to \mathcal{N}$ can be derived by first writing the first order Taylor series approximation for the right perturbation,

$$f(\mathcal{X} \oplus_{\mathrm{R}} \boldsymbol{\tau}) \approx f(\mathcal{X}) \oplus_{\mathrm{R}} \frac{Df}{D\mathcal{X}}^{\mathrm{R}} \boldsymbol{\tau}, \tag{83}$$

and comparing to the Taylor series approximation for the left perturbation where $\mathrm{Ad}_{\mathcal{X}}\boldsymbol{\tau}$ is used instead for $\boldsymbol{\tau}$,

$$f(\mathcal{X} \oplus_{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}}\boldsymbol{\tau}) \approx f(\mathcal{X}) \oplus_{\mathrm{L}} \frac{Df}{D\mathcal{X}}^{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}}\boldsymbol{\tau} \tag{84}$$

$$= f(\mathcal{X}) \oplus_{\mathrm{R}} \mathrm{Ad}_{f(\mathcal{X})}^{-1} \frac{Df}{D\mathcal{X}}^{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}}\boldsymbol{\tau}. \tag{85}$$

Since $\mathcal{X} \oplus_{\mathrm{R}} \boldsymbol{\tau} = \mathcal{X} \oplus_{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}}\boldsymbol{\tau}$, comparing (83) and (85) yields

$$\frac{Df}{D\mathcal{X}}^{\mathrm{R}} = \mathrm{Ad}_{f(\mathcal{X})}^{-1} \frac{Df}{D\mathcal{X}}^{\mathrm{L}} \mathrm{Ad}_{\mathcal{X}}. \tag{86}$$