指数映射的导数

Ethan Eade

November 12, 2018

1 简介

本文档计算

$$\left[\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\right]\log\left(\exp(x+\epsilon)\cdot\exp(x)^{-1}\right) \tag{1}$$

其中 exp 和 log 是李群中的指数映射及其逆映射，并且 x 和 ϵ 是相关李代数的元素。

2 定义

设 $\mathcal{G}$ 是一个李群，具有相关的李代数 $\mathfrak{g}$。则指数映射将代数元素转化为群元素：

$$\exp:\mathfrak{g}\to\mathcal{G} \tag{2}$$

$$\exp(x)=\mathbf{I}+x+\frac{1}{2!}x^2+\frac{1}{3!}x^3+\dots \tag{3}$$

群的伴随表示 Adj 通过与群元素左乘，线性地变换一个代数元素的指数映射：

$$x\in\mathfrak{g} \tag{4}$$

$$Y\in\mathcal{G} \tag{5}$$

$$Y\cdot\exp(x)=\exp\left(\mathrm{Adj}_Y\cdot x\right)\cdot Y \tag{6}$$

在代数中的伴随算子是表示李括号的线性算子：

$$x,y\in\mathfrak{g} \tag{7}$$

$$\mathrm{ad}_x\cdot y=x\cdot y-y\cdot x \tag{8}$$

伴随算子与指数映射进行交换：

$$\mathrm{Adj}_{\exp(y)}=\exp\left(\mathrm{ad}_y\right) \tag{9}$$

我们定义一个函数 f 从代数到群的微分如下：

$$f:\mathfrak{g}\to\mathcal{G} \tag{10}$$

$$\frac{\partial f(x)}{\partial x}:\mathfrak{g}\to\mathfrak{g} \tag{11}$$

$$\frac{\partial f(x)}{\partial x}\equiv\left[\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0}\right]\log\left(f(x+\epsilon)\cdot f(x)^{-1}\right) \tag{12}$$

在本文档中，我们对 exp 的导数 $D_{\exp}$ 感兴趣：

$$D_{\exp} : \mathfrak{g} \to \mathfrak{g} \tag{13}$$

$$D_{\exp}(x) = \frac{\partial \exp(x)}{\partial x} \tag{14}$$

3 $D_{\exp}(x)$ 公式的推导

这里不是一个严格的推导 (省略了两个近似步骤所需的 epsilon-delta 证明)，但我觉得它直观地令人满意。更严格的方法是使用关于连续向量场上积分流的定理。

定义 F 为 x 的 exp，由一个代数元素 ϵ 修改：

$$\epsilon \in \mathfrak{g} \tag{15}$$

$$F(x, \epsilon) = \exp(x + \epsilon) \tag{16}$$

我们也可以取同一测地线上多个较小群元素的乘积：

$$F(x, \epsilon) = \prod_{i=1}^{N} \exp\left(\frac{1}{N} \cdot (x + \epsilon)\right) \tag{17}$$

让步数 N 任意大，我们可以发送 $\frac{1}{N^2} \to 0$。则对于任意精确度，我们有

$$F(x, \epsilon) \approx \prod_{i=1}^{N} \exp\left(\frac{x}{N}\right) \cdot \exp\left(\frac{\epsilon}{N}\right) \tag{18}$$

$\exp\left(\frac{\epsilon}{N}\right)$ 的每一个因子都可以通过乘以伴随值适当的次数转移到乘积的左侧：

$$A_N \equiv \mathrm{Adj}_{\exp\left(\frac{x}{N}\right)} \tag{19}$$

$$F(x, \epsilon) \approx \left[\exp\left(\frac{1}{N} \cdot A_N \cdot \epsilon\right) \cdot \exp\left(\frac{1}{N} \cdot A_N^2 \cdot \epsilon\right) \cdot \ldots \cdot \exp\left(\frac{1}{N} \cdot A_N^N \cdot \epsilon\right)\right] \cdot \left[\prod_{i=1}^{N} \exp\left(\frac{x}{N}\right)\right] \tag{20}$$

$$= \left[\prod_{i=1}^{N} \exp\left(\frac{1}{N} \cdot A_N^i \cdot \epsilon\right)\right] \cdot \left[\prod_{i=1}^{N} \exp\left(\frac{x}{N}\right)\right] \tag{21}$$

$$= \left[\prod_{i=1}^{N} \exp\left(\frac{1}{N} \cdot A_N^i \cdot \epsilon\right)\right] \cdot \exp(x) \tag{22}$$

通过选择足够小的 ϵ，指数的乘积可以任意地很好地近似于一个总和的指数：

$$F(x, \epsilon) = \exp\left(\frac{1}{N} \cdot \sum_{i=1}^{N} A_N^i \cdot \epsilon + O\left(\|\epsilon\|^2\right)\right) \cdot \exp(x) \tag{23}$$

对于一个李群，我们可以使用伴随的性质方程 (9) 重写 A_N ：

$$A_N \equiv \mathrm{Adj}_{\exp\left(\frac{x}{N}\right)} \tag{24}$$

$$= \exp\left(\mathrm{ad}_{\frac{x}{N}}\right) \tag{25}$$

$$= \exp\left(\frac{1}{N} \cdot \mathrm{ad}_x\right) \tag{26}$$

取第 i 次方为：

$$A_N^i = \exp\left(\frac{i}{N} \cdot \mathrm{ad}_x\right) \tag{27}$$

因此当 $N \to \infty$，总和变为积分：

$$\frac{1}{N} \cdot \sum_{i=1}^{N} A_N^i = \frac{1}{N} \cdot \sum_{i=1}^{N} \exp\left(\frac{i}{N} \cdot \mathrm{ad}_x\right) \tag{28}$$

$$\to \int_0^1 \exp\left(t \cdot \mathrm{ad}_x\right) \cdot \mathrm{d}t \tag{29}$$

积分可以在矩阵指数的幂级数上进行。

$$\frac{1}{N} \cdot \sum_{i=1}^{N} A_N^i = \int_0^1 \left(\sum_{i=0}^{\infty} \frac{t^i \cdot \mathrm{ad}_x^i}{i!}\right) \cdot \mathrm{d}t \tag{30}$$

$$= \left(\sum_{i=0}^{\infty} \frac{t^{i+1} \mathrm{ad}_x^i}{(i+1)!}\right)\Bigg|_0^1 \tag{31}$$

$$= \sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!} \tag{32}$$

代入等式 (23)：

$$F(x, \epsilon) = \exp\left(\left(\sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!}\right) \cdot \epsilon + O\left(\|\epsilon\|^2\right)\right) \cdot \exp(x)$$

使用来自等式 (14) 的定义，

$$D_{\exp}(x) = \left[\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0}\right] \log\left(F(x, \epsilon) \cdot \exp(x)^{-1}\right) \tag{33}$$

$$= \left[\frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0}\right] \left(\sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!}\right) \cdot \epsilon + O\left(\|\epsilon\|^2\right) \tag{34}$$

$$= \sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!} \tag{35}$$

4 log 的导数

当 $x = \log(\exp(x))$ 时，在等式 (14) 中，我们可以倒置已微分的函数：

$$\delta \equiv f(\epsilon) = \log\left(\exp(x + \epsilon) \cdot \exp(x)^{-1}\right) \tag{36}$$

$$\epsilon = \log(\exp(\delta) \cdot \exp(x)) - x \tag{37}$$

当按 δ 微分时，第二项消失：

$$D_{\log}(x) \equiv \left[\frac{\partial}{\partial \delta}\bigg|_{\delta=0}\right] \log(\exp(\delta) \cdot \exp(x)) \tag{38}$$

在函数的双射区域中，逆映射的导数是导数的逆映射：

$$\frac{\partial \epsilon}{\partial \delta} = \left[\frac{\partial \delta}{\partial \epsilon}\right]^{-1} \tag{39}$$

$$D_{\log}(x) = D_{\exp}^{-1}(x) \tag{40}$$

5 特殊情况

等式 (35) 的无穷级数可以在某些李群中可用封闭形式表达。

5.1 SO(3)

5.1.1 exp 的导数

代数 $\mathfrak{so}(3)$ 的元素为 3×3 斜对称矩阵，且伴随表示相同：

$$\omega \in \Re^3 \tag{41}$$

$$\omega_\times = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3) \tag{42}$$

$$\mathrm{ad}_\omega = \omega_\times \tag{43}$$

$$\mathrm{ad}_\omega^3 = -\|\omega\|^2 \cdot \mathrm{ad}_\omega \tag{44}$$

由于 ad 的高阶次幂折回到低阶次幂，因此我们可以收集系列中的项：

$$D_{\exp}(\omega) = \mathbf{I} + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\omega\|^{2i}}{(2i+2)!}\right) \cdot \mathrm{ad}_\omega + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\omega\|^{2i}}{(2i+3)!}\right) \cdot \mathrm{ad}_\omega^2 \tag{45}$$

$$= \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2}\right) \cdot \omega_\times + \left(\frac{1-\frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2}\right) \cdot \omega_\times^2 \tag{46}$$

注意这个

$$\omega_\times^2 = \omega\omega^T - \|\omega\|^2\mathbf{I} \tag{47}$$

所以 $D_{\exp}(\omega)$ 可被重写为：

$$D_{\exp}(\omega) = \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2}\right) \cdot \omega_\times + \left(\frac{1-\frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2}\right) \cdot \left(\omega\omega^T - \|\omega\|^2\mathbf{I}\right) \tag{48}$$

$$= \frac{\sin\|\omega\|}{\|\omega\|} \cdot \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2}\right) \cdot \omega_\times + \left(\frac{1-\frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2}\right) \cdot \omega\omega^T \tag{49}$$

为方便起见，我们标记系数：

$$a_\theta = \frac{\sin\theta}{\theta} \tag{50}$$

$$b_\theta = \frac{1-\cos\theta}{\theta^2} \tag{51}$$

$$c_\theta = \frac{1-a_\theta}{\theta^2} \tag{52}$$

$$D_{\exp}(\omega) = a_{\|\omega\|} \cdot \mathbf{I} + b_{\|\omega\|} \cdot \omega_\times + c_{\|\omega\|} \cdot \omega\omega^T \tag{53}$$

5.1.2 log 的导数

回想一下，在 exp 和 log 的双射区域中，

$$D_{\log}(\omega) = D_{\exp}^{-1}(\omega) \tag{54}$$

对于 $\|\omega\| < 2\pi$，$D_{\exp}(\omega)$ 存在一个封闭形式的逆映射：

$$D_{\exp}^{-1}(\omega) = \mathbf{I} - \frac{1}{2}\omega_\times + e_{\|\omega\|}\omega_\times^2 \tag{55}$$

$$e_\theta = \frac{b_\theta - 2c_\theta}{2a_\theta} \tag{56}$$

$$= \frac{b_\theta - \frac{1}{2}a_\theta}{1-\cos\theta} \tag{57}$$

根据 θ 的值，应在等式 (56) 或等式 (57) 中使用一个更方便的等式来计算 e_θ。

5.2 SE(3)

5.2.1 exp 的导数

同样的，ad 的高阶次幂可表达为低阶次幂：

$$u, \omega \in \Re^3 \tag{58}$$

$$\theta \equiv \|\omega\| \tag{59}$$

$$x = \begin{pmatrix} \omega_\times & u \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(3) \tag{60}$$

$$\mathrm{ad}_x = \begin{pmatrix} \omega_\times & u_\times \\ 0 & \omega_\times \end{pmatrix} \tag{61}$$

$$\mathrm{ad}_x^2 = \begin{pmatrix} \omega_\times^2 & (\omega_\times u_\times + u_\times \omega_\times) \\ 0 & \omega_\times^2 \end{pmatrix} \tag{62}$$

$$\mathrm{ad}_x^3 = -\theta^2 \cdot \mathrm{ad}_x - 2\left(\omega^T u\right)\begin{pmatrix} 0 & \omega_\times \\ 0 & 0 \end{pmatrix} \tag{63}$$

收集各项，我们有：

$$Q(\omega) \equiv \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right) \cdot \omega_x + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right) \cdot \omega_\times^2 \tag{64}$$

$$D_{\exp}(x) = \mathbf{I} + a_\theta \cdot \mathrm{ad}_x + c_\theta \cdot \mathrm{ad}_x^2 + \left(\omega^T u\right) \cdot \begin{pmatrix} 0 & Q(\omega) \\ 0 & 0 \end{pmatrix} \tag{65}$$

$$= \begin{pmatrix} D_{\exp}(\omega) & \left(b_\theta \cdot u_\times + c_\theta \cdot (\omega_\times u_\times + u_\times \omega_\times) + \left(\omega^T u\right) \cdot Q(\omega)\right) \\ 0 & D_{\exp}(\omega) \end{pmatrix} \tag{66}$$

使用特征式

$$\omega_\times u_\times + u_\times \omega_\times = \omega u^T + u\omega^T - 2\left(\omega^T u\right)\mathbf{I} \tag{67}$$

我们可以重写 $D_{\exp}(x)$：

$$W(\omega) \equiv -2c_\theta \cdot \mathbf{I} + Q(\omega) \tag{68}$$

$$= -2c_\theta \cdot \mathbf{I} + \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right) \cdot \omega_\times + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right) \cdot \left(\omega\omega^T - \theta^2 \mathbf{I}\right) \tag{69}$$

$$= (c_\theta - b_\theta) \cdot \mathbf{I} + \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right) \cdot \omega_\times + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right) \cdot \omega\omega^T \tag{70}$$

$$D_{\exp}(x) = \begin{pmatrix} D_{\exp}(\omega) & \left(b_\theta \cdot u_\times + c_\theta \cdot \left(\omega u^T + u\omega^T\right) + \left(\omega^T u\right) \cdot W(\omega)\right) \\ 0 & D_{\exp}(\omega) \end{pmatrix} \tag{71}$$

5.2.2 log 的导数

一个平方分块矩阵 M 具有形式 -

$$M = \begin{pmatrix} A & B \\ 0 & A \end{pmatrix} \tag{72}$$

并有逆矩阵

$$M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1} \cdot B \cdot A^{-1} \\ 0 & A^{-1} \end{pmatrix} \tag{73}$$

因此，当 $\|\omega\| < 2\pi$ 时，使用等式 (55) 给定的 $D_{\exp}^{-1}(\omega)$，对于 $D_{\exp}^{-1}(x)$ 存在封闭形式：

$$B \equiv b_\theta \cdot u_\times + c_\theta \cdot \left(\omega u^T + u\omega^T\right) + \left(\omega^T u\right) \cdot W(\omega) \tag{74}$$

$$D_{\exp}^{-1}(x) = \begin{pmatrix} D_{\exp}^{-1}(\omega) & -D_{\exp}^{-1}(\omega) \cdot B \cdot D_{\exp}^{-1}(\omega) \\ 0 & D_{\exp}^{-1}(\omega) \end{pmatrix} \tag{75}$$

5.3 SE(2)

5.3.1 exp 的导数

在 $\mathfrak{se}(2)$ 中，ad 的高阶次幂折回到低阶次幂：

$$\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \in \Re^3 \tag{76}$$

$$m = \begin{pmatrix} 0 & -\theta & x \\ \theta & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{se}(2) \tag{77}$$

$$\mathrm{ad}_m = \begin{pmatrix} 0 & -\theta & y \\ \theta & 0 & -x \\ 0 & 0 & 0 \end{pmatrix} \tag{78}$$

$$\mathrm{ad}_m^2 = \begin{pmatrix} -\theta^2 & 0 & \theta x \\ 0 & -\theta^2 & \theta y \\ 0 & 0 & 0 \end{pmatrix} \tag{79}$$

$$\mathrm{ad}_m^3 = -\theta^3 \mathrm{ad}_m \tag{80}$$

收集各项：

$$D_{\exp}(m) = \mathbf{I} + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \theta^{2i}}{(2i+2)!}\right) \mathrm{ad}_m + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \theta^{2i}}{(2i+3)!}\right) \mathrm{ad}_m^2 \tag{81}$$

$$= \mathbf{I} + \left(\frac{1-\cos\theta}{\theta^2}\right) \cdot \mathrm{ad}_m + \left(\frac{1-\frac{\sin\theta}{\theta}}{\theta^2}\right) \cdot \mathrm{ad}_m^2 \tag{82}$$

$$= \begin{pmatrix} a_\theta & -\theta b_\theta & (c_\theta x + b_\theta y) \\ \theta b_\theta & a_\theta & (c_\theta y - b_\theta x) \\ 0 & 0 & 1 \end{pmatrix} \tag{83}$$

5.3.2 log 的导数

将来自等式 (83) 的 $D_{\exp}$ 写为块矩阵形式，给出为：

$$D_{\exp} = \begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \tag{84}$$

并具有逆矩阵

$$D_{\log} = \begin{pmatrix} A^{-1} & -A^{-1} \cdot v \\ 0 & 1 \end{pmatrix} \tag{85}$$

5.4 Sim(2)

5.4.1 exp 的导数

在 $\mathfrak{sim}(2)$ 中，ad 的高阶次幂不会折回到低阶次幂：

$$\begin{pmatrix} x \\ y \\ \theta \\ \lambda \end{pmatrix} \in \Re^4 \tag{86}$$

$$m = \begin{pmatrix} 0 & -\theta & x \\ \theta & 0 & y \\ 0 & 0 & -\lambda \end{pmatrix} \in \mathfrak{sim}(2) \tag{87}$$

$$\mathrm{ad}_m = \begin{pmatrix} \lambda & -\theta & y & -x \\ \theta & \lambda & -x & -y \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{88}$$

$$= \begin{pmatrix} Q & P \\ 0 & 0 \end{pmatrix} \tag{89}$$

$$\mathrm{ad}_m^n = \begin{pmatrix} Q^n & Q^{n-1} \cdot P \\ 0 & 0 \end{pmatrix} \tag{90}$$

为了计算 $D_{\exp}$，我们可以通过特征分解 $(i \equiv \sqrt{-1})$ 将 Q 对角化：

$$Q = V \cdot D \cdot V^* \tag{91}$$

$$V \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \tag{92}$$

$$E \equiv \begin{pmatrix} \lambda - \theta i & \\ & \lambda + \theta i \end{pmatrix} \tag{93}$$

现在我们可以依据 E 及其指数来表达 $D_{\exp}$：

$$\begin{aligned}
D_{\exp}(m) &= \sum_{j=0}^{\infty} \frac{\mathrm{ad}_m^j}{(j+1)!} && (94)\\
&= \sum_{j=0}^{\infty} \frac{1}{(j+1)!} \begin{pmatrix} Q^j & Q^{j-1} \cdot P \\ 0 & 0 \end{pmatrix} && (95)\\
&= \left(\begin{bmatrix} \left[\sum_{j=0}^{\infty} \frac{Q^j}{(j+1)!}\right] & \left[\left(\sum_{j=0}^{\infty} \frac{Q^j}{(j+2)!}\right) \cdot P\right] \\ 0 & 0 \end{bmatrix} \right) && (96)\\
&= \begin{pmatrix} \left[V \cdot \left(\sum_{j=0}^{\infty} \frac{E^j}{(j+1)!}\right) \cdot V^*\right] & \left[V \cdot \left(\sum_{j=0}^{\infty} \frac{E^j}{(j+2)!}\right) \cdot V^* \cdot P\right] \\ 0 & 0 \end{pmatrix} && (97)\\
&= \begin{pmatrix} [V \cdot E^{-1} \cdot (\exp_0(E) - \mathbf{I}) \cdot V^*] & [V \cdot E^{-2} \cdot (\exp(E) - \mathbf{I} - E) \cdot V^* \cdot P] \\ 0 & 0 \end{pmatrix} && (98)
\end{aligned}$$

当 E 的逆矩阵存在时，E^{-1} 有一个简单形式：

$$E^{-1} = \frac{1}{\lambda^2 + \theta^2} \begin{pmatrix} \lambda + \theta i & \\ & \lambda - \theta i \end{pmatrix} \tag{99}$$

乘回等式仅产生实数元素。

$$D_{\exp}(m) = \begin{pmatrix} \begin{pmatrix} p & -q \\ q & p \end{pmatrix} & \begin{pmatrix} g & -h \\ h & g \end{pmatrix} \cdot P \\ 0 & \mathbf{I} \end{pmatrix} \tag{100}$$

$$p \equiv \frac{1}{\lambda^2 + \theta^2} \left[e^{\lambda} \cdot (\lambda \cos\theta + \theta \sin\theta) - \lambda\right] \tag{101}$$

$$q \equiv \frac{1}{\lambda^2 + \theta^2} \left[e^{\lambda} \cdot (\lambda \sin\theta - \theta \cos\theta) + \theta\right] \tag{102}$$

$$g \equiv \frac{1}{\lambda^2 + \theta^2} \left[\frac{1}{\lambda^2 + \theta^2} \cdot (\lambda p + \theta q) - \lambda\right] \tag{103}$$

$$h \equiv \frac{1}{\lambda^2 + \theta^2} \left[\frac{1}{\lambda^2 + \theta^2} \cdot (\lambda q - \theta p) + \theta\right] \tag{104}$$

当 $\lambda^2 + \theta^2 \to 0$ 时，则应使用泰勒展开式替代：

$$p \equiv 1 + \frac{a}{2} \tag{105}$$

$$q \equiv \frac{b}{2} \tag{106}$$

$$g \equiv \frac{1}{2} + \frac{a}{6} \tag{107}$$

$$h \equiv \frac{b}{6} \tag{108}$$

5.4.2 log 的导数

将来自等式 (100) 的 $D_{\exp}$ 写为块矩阵形式，给出为：

$$D_{\exp} = \begin{pmatrix} A & B \\ 0 & \mathbf{I} \end{pmatrix} \tag{109}$$

并具有逆矩阵

$$D_{\log} = \begin{pmatrix} A^{-1} & -A^{-1} \cdot B \\ 0 & \mathbf{I} \end{pmatrix} \tag{110}$$

6 读书笔记

6.1 对于方程 (9) 的理解与证明

对于方程 (9)，

$$\mathrm{Adj}_{\exp(y)} = \exp(\mathrm{ad}_y)$$

其表明，对于给定的李代数元素 y，它的指数映射 $\exp(y)$ 的伴随矩阵等于对应的李代数的伴随矩阵的指数映射。这个公式的意义在于，通过指数映射和伴随矩阵的关系，我们可以在李群和李代数之间进行转换和对应。

为证明方程 (9)，我们首先需要明确伴随表示。对于李群 $\mathcal{G}$ 中的一个元素 g，我们可以定义其伴随表示 Adj_g 为一个从 $\mathcal{G}$ 到 $\mathcal{G}$ 的映射，具体为 $\mathrm{Adj}_g(h) = g \cdot h \cdot g^{-1}$，其中 "$\cdot$" 表示群的乘法操作。这个映射可以推广到李代数上，即对于李代数上的一个元素 x，我们可以定义其伴随表示 ad_x 为一个从李代数到李代数的映射，具体为 $\mathrm{ad}_x(y) = [x, y]$，其中 "$[x, y]$" 表示李括号。

然后，我们需要明确指数映射。对于李代数上的一个元素 x，我们可以通过指数映射 $\exp(x)$ 来得到李群上的一个元素。这个映射的具体形式为 $\exp(x) = \sum_0^\infty (x^n/n!)$。并且我们有 $\exp(y)^{-1} = \exp(-y)$，所以有 $\exp(y) \cdot x \cdot \exp(y)^{-1} = \exp(y) \cdot x \cdot \exp(-y)$。

在推导这个公式之前，我们需要了解一些基本的概念和性质：

1. 对于李群 $\mathcal{G}$ 中的元素 g，其伴随表示 $\mathrm{Adj}_g : \mathcal{G} \to \mathcal{G}$ 定义为 $\mathrm{Adj}_g(h) = g \cdot h \cdot g^{-1}$，其中 $h \in \mathcal{G}$。
2. 对于李代数 $\mathfrak{g}$ 中的元素 x，其伴随表示 $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ 定义为 $\mathrm{ad}_x(y) = [x, y]$，其中 $y \in \mathfrak{g}$，$[x, y]$ 是 x 和 y 的李括号。
3. 指数映射 $\exp : \mathfrak{g} \to \mathcal{G}$ 是李代数 $\mathfrak{g}$ 到李群 $\mathcal{G}$ 的一个光滑映射，且满足 $\exp(0) = e$ (这里 e 是李群的单位元)，以及 $\exp(x + y) = \exp(x)\exp(y)$ (对于足够小的 $x, y \in \mathfrak{g}$)。

利用这些定义和性质，我们可以推导 $\mathrm{Adj}_{\exp(y)} = \exp(\mathrm{ad}_y)$ 这个公式。为了简化表示，我们设 $g = \exp(y)$，则 $\mathrm{Adj}_g(h) = g \cdot h \cdot g^{-1}$。我们需要计算 h 的微分 $\frac{d}{dt}\mathrm{Adj}_{\exp(ty)}(h)\Big|_{t=0}$，其中 t 是实数。

首先，我们需要计算 $\frac{d}{dt}\mathrm{Adj}_{\exp(ty)}(h)\Big|_{t=0}$，也就是 $\frac{d}{dt}(\exp(ty) \cdot h \cdot \exp(-ty))\big|_{t=0}$。

我们可以将 $\exp(ty) \cdot h \cdot \exp(-ty)$ 看作两个函数的复合，即 $f(t) = \exp(ty) \cdot h$ 和 $g(t) = \exp(-ty)$，我们要计算的是 $(f(t)g(t))'$ 在 $t = 0$ 的值。

根据链式法则，$(f(t)g(t))' = f'(t)g(t) + f(t)g'(t)$。因此，我们有

$$\frac{d}{dt}(\exp(ty) \cdot h \cdot \exp(-ty))\Big|_{t=0} = \frac{d}{dt}(\exp(ty) \cdot h) \cdot \exp(-ty) + \exp(ty) \cdot h \cdot \frac{d}{dt}\exp(-ty)\Big|_{t=0}$$

其次，我们需要计算 $\frac{d}{dt}\exp(ty)$ 和 $\frac{d}{dt}\exp(-ty)$。根据指数映射的微分性质，我们有 $\frac{d}{dt}\exp(ty) = y\exp(ty)$ 和 $\frac{d}{dt}\exp(-ty) = -y\exp(-ty)$。因此，我们得到

$$\begin{aligned}&\frac{d}{dt}(\exp(ty) \cdot h) \cdot \exp(-ty) + \exp(ty) \cdot h \cdot \frac{d}{dt}\exp(-ty)\Big|_{t=0} \\ &= y\exp(ty) \cdot h \cdot \exp(-ty) - \exp(ty) \cdot h \cdot y\exp(-ty)\big|_{t=0}\end{aligned}$$

由于 $\exp(0) = e$，我们得到

$$yehe - ehye = yh - hy = [y, h] = \mathrm{ad}_y(h)$$

整理以上步骤，我们得到

$$\begin{aligned}\frac{d}{dt}\mathrm{Adj}_{\exp(ty)}(h)\bigg|_{t=0} &= \frac{d}{dt}(\exp(ty)\cdot h\cdot\exp(-ty))\bigg|_{t=0}\\ &= [y,h]\\ &= \mathrm{ad}_y(h)\end{aligned}$$

其中，第一步使用了链式法则，第二步使用了 $\exp(ty)$ 的微分性质和李括号的定义。

因此，我们有 $\frac{d}{dt}\mathrm{Adj}_{\exp(ty)}\Big|_{t=0} = \mathrm{ad}_y$ ，这意味着 $\mathrm{Adj}_{\exp(ty)}$ 和 $\exp(t\mathrm{ad}_y)$ 在 $t=0$ 处的切空间相同。由于它们都是光滑的，并且在 $t=0$ 处相等 (即 $\mathrm{Adj}_e = \exp(0) = e$)，我们可以得出 $\mathrm{Adj}_{\exp(ty)} = \exp(t\mathrm{ad}_y)$ 。

最后，将 t 替换为 1 ，我们得到 $\mathrm{Adj}_{\exp(y)} = \exp\left(\mathrm{ad}_y\right)$ ，这就完成了公式的推导。

在李群和李代数的理论中，Adj 和 ad 是两个核心的概念。Adj 是李群的伴随表示，而 ad 是李代数的伴随表示。Adj_g 表示在李群 $\mathcal{G}$ 中元素 g 作用下, 另一个元素 x 的伴随变换。而 ad_y 表示在李代数 $\mathfrak{g}$ 中元素 y 作用下, 另一个李代数元素 x 的伴随变换。李群和李代数之间存在指数映射的对应关系。也就是说, 对于任意李代数 $\mathfrak{g}$ 中的元素 y, 我们都可以找到对应的李群 $\mathcal{G}$ 中的元素 $\exp(y)$。其次，李代数和李群之间也存在着相似的伴随表达 (adjoint representation) 变换的概念。因此, 李代数元素 y 与李群元素 $\exp(y)$ 之间, 其伴随变换应该存在某种一一对应关系。这两者之间的联系可以通过这个公式 $\mathrm{Adj}_{\exp(y)} = \exp\left(\mathrm{ad}_y\right)$ 来表达。

综上所述，这个公式是李群和李代数之间的一个重要联系。我们知道，对于一个李群 $\mathcal{G}$ ，我们可以定义它的李代数为其单位元在切空间上的向量空间，这个向量空间上的向量可以通过李括号来定义一个李代数结构。同时，我们知道，李群上的操作可以通过指数映射来转化为李代数上的操作，这就是方程 (9) 想要表达的内容。

6.2 对于方程 (35) 的探讨

对于指数映射的导数的证明，可以从网上寻找到这些文献资料：

1. A micro Lie theory for state estimation in robotics
2. Derivative of the Exponential Map (本文)
3. The Derivation of the Exponential Map of Matrices
4. Derivative of the exponential map
5. Derivatives and Differentials

对比本文的方程 (35)

$$D_{\exp}(x) = \sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!}$$

就会发现与文献 [3,4,5] 的结论

$$D_{\exp}(x) = \sum_{i=0}^{\infty} \frac{(-1)^i\mathrm{ad}_x^i}{(i+1)!}$$

不一致。这其实是选取全局坐标和局部坐标造成的差异。

本文从定义方程 (1) 开始，对应的是文献 [1] 中表示全局坐标系的左结合 (left-) 算子 $\oplus$ 和 $\ominus$，即方程 (27,28)，以及表示全局或原点处的导数，即左 Jacobian 矩阵，方程 (44)。其结果，方程 (35)，对应于文献 [1] 的方程 (71)。验证结果，本文的 SO(3) 的 exp 的导数，方程 (46)，对应于文献 [1] 的方程 (145)。全局表示与局部表示的导数之间的关系，参见文献 [1] 的方程 (75,76)。

对于幂级数

$$\sum_{i=0}^{\infty} \frac{(-1)^i \mathrm{ad}_x^i}{(i+1)!}$$

通常习惯地被简写为

$$\frac{\mathrm{id} - e^{-\mathrm{ad}_x}}{\mathrm{ad}_x},$$

这样的写法更加简洁，也反映了这个级数与指数函数关系的本质。类似地，对于幂级数

$$\sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!}$$

也可以被简写为

$$\frac{e^{\mathrm{ad}_x} - \mathrm{id}}{\mathrm{ad}_x}.$$

本文的证明不是很严格，我们可以参考文献 [3] 的方法进行证明。在参考文献 [3] 中有这一步：将 Leibniz 规则应用于等式 $e^{X+tY} = \exp\left(\left(1/n\right)X + t\left(1/n\right)Y\right)^n$，对于任意 $n \in \mathbf{Z}$ 有效，我们得到：

$$\left.\frac{d}{dt}\right|_{t=0} e^{X+tY} = \sum_{k=0}^{n-1} \exp\left(\frac{1}{n}X\right)^{n-1-k} \cdot \left(\left.\frac{d}{dt}\right|_{t=0} \exp\left(\frac{1}{n}X + t\frac{1}{n}Y\right)\right) \cdot \exp\left(\frac{1}{n}X\right)^{k}.$$

我们可以将左右项对调，得到：

$$\left.\frac{d}{dt}\right|_{t=0} e^{X+tY} = \sum_{k=0}^{n-1} \exp\left(\frac{1}{n}X\right)^{n-1+k} \cdot \left(\left.\frac{d}{dt}\right|_{t=0} \exp\left(\frac{1}{n}X + t\frac{1}{n}Y\right)\right) \cdot \exp\left(\frac{1}{n}X\right)^{-k}.$$

往下推导即可得到本文的结果。

此外，文献 [1,2] 与文献 [3,4,5]，对于导数的定义方程还有一些微妙的差异。文献 [3,4,5] 是相对于一个实数，即相对于时间 $t = 0$ 的求导；而文献 [1,2] 是相对于一个李代数元素或向量元素的求导，即相对于一个状态量的求导。这两者都有各自的应用领域。例如，在使用卡尔曼滤波器对刚体的位姿进行最优估计的算法中就得到应用。卡尔曼滤波器的一次迭代分为两个阶段：时间更新和测量更新。在时间更新阶段，我们需要使用运动学方程或动力学方程，也就是使用时间相关的方程，因此使用的导数适用于文献 [3,4,5] 的定义。在测量更新阶段，我们认为用于校正的测量值是同步测量得到，只有状态量差异而没有时间差异，因此使用的导数适用于文献 [1,2] 的定义。