Derivative of the Exponential Map

Ethan Eade

November 12, 2018

1 Introduction

This document computes

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

where exp and log are the exponential mapping and its inverse in a Lie group, and x and ϵ are elements of the associated Lie algebra.

2 Definitions

Let $\mathcal{G}$ be a Lie group, with associated Lie algebra $\mathfrak{g}$. Then the exponential map takes algebra elements to group elements:

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

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

The adjoint representation Adj of the group linearly transforms the exponential mapping of an algebra element through left multiplication by a group element:

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

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

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

The adjoint operator in the algebra is the linear operator representing the Lie bracket:

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

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

The adjoint operator commutes with the exponential map:

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

We define differentiation of a function f from algebra to group as follows:

$$\begin{aligned} f : \mathfrak{g} &\rightarrow \mathcal{G} & (10)\\ \frac{\partial f(x)}{\partial x} : \mathfrak{g} &\rightarrow \mathfrak{g} & (11)\\ \frac{\partial f(x)}{\partial x} &\equiv \left[\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0}\right] \log\left(f(x+\epsilon)\cdot f(x)^{-1}\right) & (12)\end{aligned}$$

In this document, we're interested in $D_{\exp}$, the derivative of exp:

$$\begin{aligned} D_{\exp} : \mathfrak{g} &\rightarrow \mathfrak{g} & (13)\\ D_{\exp}(x) &= \frac{\partial \exp(x)}{\partial x} & (14)\end{aligned}$$

3 Derivation of a formula for $D_{\exp}(x)$

This isn't a rigorous derivation (the epsilon-delta proofs required for the two approximation steps are omitted), but I find it intuitively pleasing. A more rigorous approach would use theorems about integrated flows on continuous vector fields.

Define F to be exp of x modified by an algebra element ϵ:

$$\begin{aligned} \epsilon &\in \mathfrak{g} & (15)\\ F(x,\epsilon) &= \exp(x+\epsilon) & (16)\end{aligned}$$

We can also take the product of multiple smaller group elements on the same geodesic:

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

Letting the number of steps N go arbitrarily large, we can send $\frac{1}{N^2} \rightarrow 0$. Then we have, to arbitrary accuracy:

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

Each factor of $\exp\left(\frac{\epsilon}{N}\right)$ can be shifted to the left side of the product by multiplying by the adjoint an appropriate number of times:

$$\begin{aligned} A_N &\equiv \text{Adj}_{\exp\left(\frac{x}{N}\right)} & (19)\\ F(x,\epsilon) &\approx \left[\exp\left(\frac{1}{N}\cdot A_N\cdot\epsilon\right)\cdot\exp\left(\frac{1}{N}\cdot A_N^2\cdot\epsilon\right)\cdot\ldots\cdot\exp\left(\frac{1}{N}\cdot A_N^N\cdot\epsilon\right)\right]\cdot\left[\prod_{i=1}^{N}\exp\left(\frac{x}{N}\right)\right] & (20)\\ &= \left[\prod_{i=1}^{N}\exp\left(\frac{1}{N}\cdot A_N^i\cdot\epsilon\right)\right]\cdot\left[\prod_{i=1}^{N}\exp\left(\frac{x}{N}\right)\right] & (21)\\ &= \left[\prod_{i=1}^{N}\exp\left(\frac{1}{N}\cdot A_N^i\cdot\epsilon\right)\right]\cdot\exp(x) & (22) \end{aligned}$$

By choosing ϵ sufficiently small, the product of exponentials is arbitrarily well approximated by the exponential of a sum:

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

We can use the properties of the adjoint to rewrite A_N:

$$\begin{aligned} A_N &\equiv \text{Adj}_{\exp\left(\frac{x}{N}\right)} & (24)\\ &= \exp\left(\text{ad}_{\frac{x}{N}}\right) & (25)\\ &= \exp\left(\frac{1}{N}\cdot\text{ad}_x\right) & (26) \end{aligned}$$

for a Lie group

Taking the i^{th} power:

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

Thus as $N\to\infty$, the sum becomes an integral:

$$\begin{aligned} \frac{1}{N}\cdot\sum_{i=1}^{N}A_N^i &= \frac{1}{N}\cdot\sum_{i=1}^{N}\exp\left(\frac{i}{N}\cdot\text{ad}_x\right) & (28)\\ &\to \int_0^1\exp\left(t\cdot\text{ad}_x\right)\cdot\text{d}t & (29) \end{aligned}$$

The integration can be performed on the power series of the matrix exponential.

$$\begin{aligned} \frac{1}{N}\cdot\sum_{i=1}^{N}A_N^i &= \int_0^1\left(\sum_{i=0}^{\infty}\frac{t^i\cdot\text{ad}_x^i}{i!}\right)\cdot\text{d}t & (30)\\ &= \left(\sum_{i=0}^{\infty}\frac{t^{i+1}\text{ad}_x^i}{(i+1)!}\right)\Big|_0^1 & (31)\\ &= \sum_{i=0}^{\infty}\frac{\text{ad}_x^i}{(i+1)!} & (32) \end{aligned}$$

Substituting into Eq.23:

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

Using the definition from Eq.14,

$$\begin{aligned}
D_{\exp}(x) &= \left[\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\right]\log\left(F(x,\epsilon)\cdot\exp(x)^{-1}\right) && (33)\\
&= \left[\frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\right]\left(\sum_{i=0}^{\infty}\frac{\mathrm{ad}_x^i}{(i+1)!}\right)\cdot\epsilon + O\left(\|\epsilon\|^2\right) && (34)\\
&= \sum_{i=0}^{\infty}\frac{\mathrm{ad}_x^i}{(i+1)!} && (35)
\end{aligned}$$

4 Derivative of log

When $x = \log(\exp(x))$, we can invert the function being differentiated in Eq.14:

$$\begin{aligned}
\delta \equiv f(\epsilon) &= \log\left(\exp(x+\epsilon)\cdot\exp(x)^{-1}\right) && (36)\\
\epsilon &= \log(\exp(\delta)\cdot\exp(x)) - x && (37)
\end{aligned}$$

The second term vanishes when differentiating by δ:

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

In this bijective region of the function, the derivative of the inverse is the inverse of the derivative:

$$\begin{aligned}
\frac{\partial\epsilon}{\partial\delta} &= \left[\frac{\partial\delta}{\partial\epsilon}\right]^{-1} && (39)\\
D_{\log}(x) &= D_{\exp}^{-1}(x) && (40)
\end{aligned}$$

5 Special Cases

The infinite series of Eq. 35 can be expressed in closed form in some Lie groups.

5.1 SO(3)

5.1.1 Derivative of exp

The elements of the algebra $\mathfrak{so}(3)$ are 3×3 skew-symmetric matrices, and the adjoint representation is identical:

$$\begin{aligned}
\omega &\in \Re^3 && (41)\\
\omega_\times &= \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3) && (42)\\
\mathrm{ad}_\omega &= \omega_\times && (43)\\
\mathrm{ad}_\omega^3 &= -\|\omega\|^2 \cdot \mathrm{ad}_\omega && (44)
\end{aligned}$$

Because the higher powers of ad collapse back to lower powers, we can collect terms in the series:

$$\begin{aligned}
D_{\exp}(\omega) &= \mathbf{I} + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\omega\|^{2i}}{(2i+2)!} \right) \cdot \mathrm{ad}_\omega + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\omega\|^{2i}}{(2i+3)!} \right) \cdot \mathrm{ad}_\omega^2 && (45)\\
&= \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2} \right) \cdot \omega_\times + \left(\frac{1 - \frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2} \right) \cdot \omega_\times^2 && (46)
\end{aligned}$$

Note that

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

So $D_{\exp}(\omega)$ can be rewritten:

$$\begin{aligned}
D_{\exp}(\omega) &= \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2} \right) \cdot \omega_\times + \left(\frac{1 - \frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2} \right) \cdot \left(\omega\omega^T - \|\omega\|^2 \mathbf{I} \right) && (48)\\
&= \frac{\sin\|\omega\|}{\|\omega\|} \cdot \mathbf{I} + \left(\frac{1-\cos\|\omega\|}{\|\omega\|^2} \right) \cdot \omega_\times + \left(\frac{1 - \frac{\sin\|\omega\|}{\|\omega\|}}{\|\omega\|^2} \right) \cdot \omega\omega^T && (49)
\end{aligned}$$

We label the coefficients for convenience:

$$\begin{aligned}
a_\theta &= \frac{\sin\theta}{\theta} && (50)\\
b_\theta &= \frac{1-\cos\theta}{\theta^2} && (51)\\
c_\theta &= \frac{1-a_\theta}{\theta^2} && (52)\\
D_{\exp}(\omega) &= a_{\|\omega\|} \cdot \mathbf{I} + b_{\|\omega\|} \cdot \omega_\times + c_{\|\omega\|} \cdot \omega\omega^T && (53)
\end{aligned}$$

5.1.2 Derivative of log

Recall that in the bijective region of exp and log,

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

For $\|\omega\| < 2\pi$, a closed-form inverse exists for $D_{\exp}(\omega)$:

$$\begin{aligned} D_{\exp}^{-1}(\omega) &= \mathbf{I} - \frac{1}{2}\omega_\times + e_{\|\omega\|}\omega_\times^2 & (55)\\ e_\theta &= \frac{b_\theta - 2c_\theta}{2a_\theta} & (56)\\ &= \frac{b_\theta - \frac{1}{2}a_\theta}{1-\cos\theta} & (57) \end{aligned}$$

Depending on the value of θ, the more convenient of Eq. 56 or Eq.57 should be used to compute e_θ.

5.2 SE(3)

5.2.1 Derivative of exp

Again, the higher powers of ad can be expressed in terms of lower powers:

$$\begin{aligned} u, \omega &\in \Re^3 & (58)\\ \theta &\equiv \|\omega\| & (59)\\ x &= \begin{pmatrix} \omega_\times & u \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(3) & (60)\\ \mathrm{ad}_x &= \begin{pmatrix} \omega_\times & u_\times \\ 0 & \omega_\times \end{pmatrix} & (61)\\ \mathrm{ad}_x^2 &= \begin{pmatrix} \omega_\times^2 & (\omega_\times u_\times + u_\times\omega_\times) \\ 0 & \omega_\times^2 \end{pmatrix} & (62)\\ \mathrm{ad}_x^3 &= -\theta^2 \cdot \mathrm{ad}_x - 2\left(\omega^T u\right)\begin{pmatrix} 0 & \omega_\times \\ 0 & 0 \end{pmatrix} & (63) \end{aligned}$$

Collecting the terms, we have:

$$\begin{aligned} Q(\omega) &\equiv \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right)\cdot\omega_x + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right)\cdot\omega_\times^2 & (64)\\ D_{\exp}(x) &= \mathbf{I} + a_\theta\cdot\mathrm{ad}_x + c_\theta\cdot\mathrm{ad}_x^2 + \left(\omega^T u\right)\cdot\begin{pmatrix} 0 & Q(\omega) \\ 0 & 0 \end{pmatrix} & (65)\\ &= \begin{pmatrix} D_{\exp}(\omega) & \left(b_\theta\cdot u_\times + c_\theta\cdot(\omega_\times u_\times + u_\times\omega_\times) + \left(\omega^T u\right)\cdot Q(\omega)\right) \\ 0 & D_{\exp}(\omega) \end{pmatrix} & (66) \end{aligned}$$

Using the identity

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

...we can rewrite $D_{\exp}(x)$:

$$\begin{aligned}
W(\omega) &\equiv -2c_\theta \cdot \mathbf{I} + Q(\omega) && (68)\\
&= -2c_\theta \cdot \mathbf{I} + \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right)\cdot \omega_\times + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right)\cdot\left(\omega\omega^T - \theta^2\mathbf{I}\right) && (69)\\
&= (c_\theta - b_\theta)\cdot \mathbf{I} + \left(\frac{a_\theta - 2b_\theta}{\theta^2}\right)\cdot \omega_\times + \left(\frac{b_\theta - 3c_\theta}{\theta^2}\right)\cdot \omega\omega^T && (70)\\
D_{\exp}(x) &= \begin{pmatrix} D_{\exp}(\omega) & \left(b_\theta \cdot u_\times + c_\theta \cdot \left(\omega u^T + u\omega^T\right) + \left(\omega^T u\right)\cdot W(\omega)\right) \\ 0 & D_{\exp}(\omega) \end{pmatrix} && (71)
\end{aligned}$$

5.2.2 Derivative of log

A square block matrix M with the form -

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

...has an inverse:

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

Thus, when $\|\omega\| < 2\pi$, a closed form exists for $D^{-1}_{\exp}(x)$ using $D^{-1}_{\exp}(\omega)$ as given by Eq.55:

$$\begin{aligned}
B &\equiv b_\theta \cdot u_\times + c_\theta \cdot \left(\omega u^T + u\omega^T\right) + \left(\omega^T u\right)\cdot W(\omega) && (74)\\
D^{-1}_{\exp}(x) &= \begin{pmatrix} D^{-1}_{\exp}(\omega) & -D^{-1}_{\exp}(\omega)\cdot B \cdot D^{-1}_{\exp}(\omega) \\ 0 & D^{-1}_{\exp}(\omega) \end{pmatrix} && (75)
\end{aligned}$$

5.3 SE(2)

5.3.1 Derivative of exp

Higher powers of ad collapse onto lower ones in $\mathfrak{se}(2)$:

$$\begin{aligned}
\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} &\in \Re^3 && (76)\\
m &= \begin{pmatrix} 0 & -\theta & x \\ \theta & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{se}(2) && (77)
\end{aligned}$$

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

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

$$\mathrm{ad}_m^3 = -\theta^3 \mathrm{ad}_m \quad (80)$$

Collecting the terms:

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

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

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

5.3.2 Derivative of log

Writing $D_{\exp}$ from Eq. 83 in block form gives:

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

...with inverse:

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

5.4 Sim(2)

5.4.1 Derivative of exp

In $\mathfrak{sim}(2)$, higher powers of ad do not collapse onto lower ones:

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

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

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

$$= \begin{pmatrix} Q & P \\ 0 & 0 \end{pmatrix} \quad (89)$$

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

To compute $D_{\exp}$ we can diagonalize Q by eigendecomposition ($i \equiv \sqrt{-1}$):

$$Q = V \cdot D \cdot V^* \quad (91)$$

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

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

Now we can express $D_{\exp}$ in terms of E and its exponential:

$$D_{\exp}(m) = \sum_{j=0}^{\infty} \frac{\mathrm{ad}_m^j}{(j+1)!} \quad (94)$$

$$= \sum_{j=0}^{\infty} \frac{1}{(j+1)!} \begin{pmatrix} Q^j & Q^{j-1} \cdot P \\ 0 & 0 \end{pmatrix} \quad (95)$$

$$= \begin{pmatrix} \left[\sum_{j=0}^{\infty} \frac{Q^j}{(j+1)!}\right] & \left[\left(\sum_{j=0}^{\infty} \frac{Q^j}{(j+2)!}\right) \cdot P\right] \\ 0 & 0 \end{pmatrix} \quad (96)$$

$$= \begin{pmatrix} \left[V \cdot \left(\sum_{j=0}^{\infty} \frac{E^j}{(j+1)!}\right) \cdot V^*\right] & \left[V \cdot \left(\sum_{j=0}^{\infty} \frac{E^j}{(j+2)!}\right) \cdot V^* \cdot P\right] \\ 0 & 0 \end{pmatrix} \quad (97)$$

$$= \begin{pmatrix} \left[V \cdot E^{-1} \cdot (\exp(E) - \mathbf{I}) \cdot V^*\right] & \left[V \cdot E^{-2} \cdot (\exp(E) - \mathbf{I} - E) \cdot V^* \cdot P\right] \\ 0 & 0 \end{pmatrix} \quad (98)$$

When it exists, E^{-1} has a simple form:

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

Multiplying through yields only real elements:

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

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

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

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

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

When $\lambda^2+\theta^2 \to 0$, the Taylor expansion should be used instead:

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

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

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

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

5.4.2 Derivative of log

Writing $D_{\exp}$ from Eq. 100 in block form gives:

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

...with inverse:

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