ECE276A: Sensing & Estimation in Robotics
Lecture 11: Matrix Lie Groups

Nikolay Atanasov

natanasov@ucsd.edu

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Outline

Topology

- **Topology** on set $\mathcal{X}$ is a set $\mathcal{T}$ of subsets of $\mathcal{X}$, called **open sets**, such that:
  - $\mathcal{X}$ and $\emptyset$ are open
  - finite intersection of open sets is open
  - uncountably infinite union of open sets is open

- **Topological space**: set $\mathcal{X}$ with topology $\mathcal{T}$

- **Hausdorff space**: topological space $\mathcal{X}$ such that $\forall x, y \in \mathcal{X}$ with $x \neq y$ there exists disjoint neighborhoods $\mathcal{U}$ of x and $\mathcal{V}$ of y

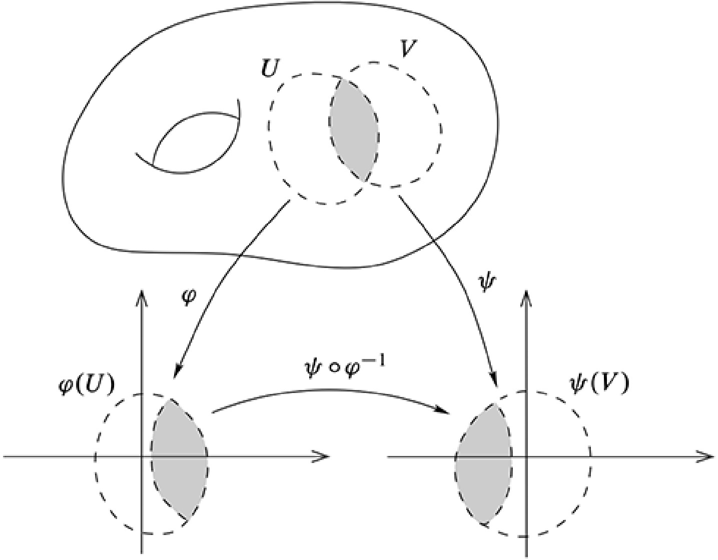
- **Separable space**: topological space $\mathcal{X}$ with a countable dense subset, i.e., there exists a sequence in $\mathcal{X}$ such that every non-empty open set contains at least one element of the sequence

- **Second-countable space**: topological space $\mathcal{X}$ with a countable base, i.e., countable collection of open sets that can express any open set as a union

Manifold

- **Homeomorphism**: continuous bijective function $f : \mathcal{X} \to \mathcal{Y}$ between two topological spaces with continuous inverse f^{-1}
- **Topological n-manifold**: Hausdorff second-countable topological space $\mathcal{M}$ such that every $p \in \mathcal{M}$ has a neighborhood $\mathcal{U}$ homeomorphic to an open subset of $\mathbb{R}^n$
- **Chart** on $\mathcal{M}$: pair $(\mathcal{U}, \phi)$ such $\phi : \mathcal{U} \subseteq \mathcal{M} \to \mathcal{V} \subseteq \mathbb{R}^n$ is a homeomorphism
- **Atlas** on $\mathcal{M}$: set of charts $\{(\mathcal{U}_\alpha, \phi_\alpha)\}_\alpha$ that cover $\mathcal{M}$
- **Coordinates of $p \in \mathcal{M}$**: elements $\phi(p) \in \mathbb{R}^n$ of a chart $(\mathcal{U}, \phi)$ containing p
- **Smooth n-manifold**: the change of coordinates function $\phi_\beta \circ \phi_\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ between any charts $(\mathcal{U}_\alpha, \phi_\alpha)$ and $(\mathcal{U}_\beta, \phi_\beta)$ with $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ is infinitely differentiable
- An open subset of a smooth n-manifold is a smooth n-manifold
- The product of smooth n_1 and n_2 manifolds is a smooth $(n_1 + n_2)$-manifold

Manifold



Embedded Submanifold

- **Directional derivative**: of $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{p} \in \mathbb{R}^n$ in direction $\mathbf{v} \in \mathbb{R}^n$:

$$Df(\mathbf{p})[\mathbf{v}] = \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{v}) - f(\mathbf{p})}{t}$$

- A nonempty subset $\mathcal{M}$ of d-dimensional Euclidean space $\mathcal{E}$ is a smooth **embedded submanifold** of dimension $n \leq d$ such that either
  1. $n = d$ and $\mathcal{M}$ is an open set in $\mathcal{E}$, called an **open submanifold**, or
  2. $n = d - k$ and, for each $p \in \mathcal{M}$, there exists a neighborhood $\mathcal{U}_p$ in $\mathcal{E}$ and a smooth function $h : \mathcal{U}_p \to \mathbb{R}^k$ such that
     2.1 if $y \in \mathcal{U}_p$, then $y \in \mathcal{M}$ iff $h(y) = 0$
     2.2 $\operatorname{rank}(Dh(p)) = k$ (rank is the range space dimension)

     The function h is called a **local defining function** for $\mathcal{M}$ at p.

- Example:
  - **unit sphere** $\mathcal{S}^{d-1} := \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{x}^\top \mathbf{x} = 1\right\}$ is an embedded submanifold of $\mathbb{R}^d$
  - $\mathcal{S}^{d-1}$ has local defining function $h(\mathbf{x}) = \mathbf{x}^\top \mathbf{x} - 1$
  - the directional derivative of h is $Dh(\mathbf{x})[\mathbf{v}] = 2\mathbf{x}^\top \mathbf{v}$ and has rank $k = 1$
  - the dimension of $\mathcal{S}^{d-1}$ is $n = d - 1$

Tangent Space

- How should directional derivative be defined for $f : \mathcal{M} \to \mathbb{R}$?
- For $p \in \mathcal{M}$, the operation $p + tv$ may not be defined. Instead, use a curve $\gamma : \mathbb{R} \to \mathcal{M}$ such that $\gamma(0) = p$.
- Let $C^\infty(\mathcal{U}_p, \mathbb{R})$ be the set of smooth real-valued functions defined on a neighborhood $\mathcal{U}_p$ of a point p on a manifold $\mathcal{M}$. A **tangent vector** v_p to $\mathcal{M}$ at p is a function from $C^\infty(\mathcal{U}_p, \mathbb{R})$ to $\mathbb{R}$ such that there exists a curve $\gamma : \mathbb{R} \to \mathcal{M}$ with $\gamma(0) = p$ and:

$$v_p[f] = \left.\frac{df(\gamma(t))}{dt}\right|_{t=0}$$

- **Tangent space to $\mathcal{M}$ at p**: set $T_p\mathcal{M}$ of all tangent vectors v_p to $\mathcal{M}$ at p

Tangent Space of Embedded Submanifold

- If $\mathcal{M}$ is an embedded submanifold, then $v \in T_p\mathcal{M}$ if and only if there exists a smooth curve γ on $\mathcal{M}$ passing through p with velocity v:

$$T_p\mathcal{M} = \left\{ \frac{d\gamma}{dt}(0) \mid \gamma : \mathcal{I} \to \mathcal{M} \quad \text{and} \quad \gamma(0) = p \right\}$$

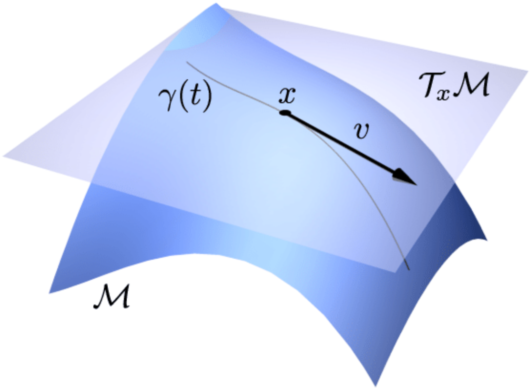
  where $\mathcal{I}$ is any open interval containing $t = 0$.

- Let $\mathcal{M}$ be an embedded submanifold of Euclidean space $\mathcal{E}$.
  - If $\mathcal{M}$ is an open submanifold of $\mathcal{E}$, then $T_p\mathcal{M} = \mathcal{E}$.
  - Otherwise, $T_p\mathcal{M} = \ker(Dh(p))$ for any local defining function h at p.

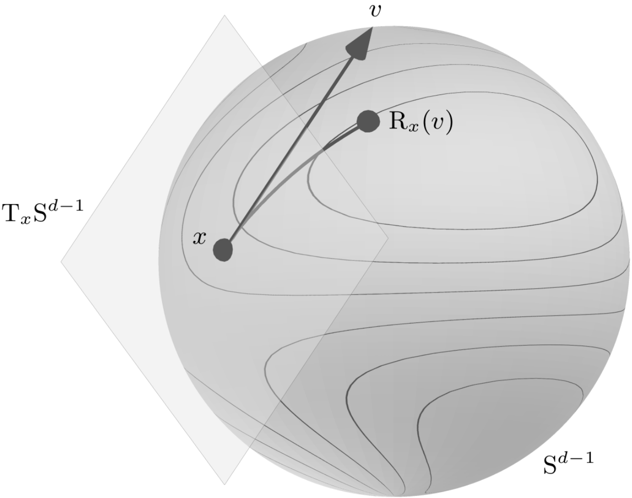
Tangent Space

- **Tangent space** $T_p\mathcal{M}$: set of all tangent vectors to $\mathcal{M}$ at p
- The tangent space $T_p\mathcal{M}$ is a **vector space** of the same dimension as $\mathcal{M}$ and can be equipped with an inner product $\langle \cdot, \cdot \rangle_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$
- **Tangent bundle of** $\mathcal{M}$: disjoint union of the tangent spaces of $\mathcal{M}$:

$$T\mathcal{M} = \{(p, v) \mid p \in \mathcal{M}, v \in T_p\mathcal{M}\}$$



Unit Sphere



- $\mathcal{S}^{d-1} := \left\{ \mathbf{x} \in \mathbb{R}^d : \mathbf{x}^\top \mathbf{x} = 1 \right\}$
- $T_{\mathbf{x}}\mathcal{S}^{d-1} = \left\{ \mathbf{v} \in \mathbb{R}^d : \mathbf{x}^\top \mathbf{v} = 0 \right\}$

Lie Group

- A **group** is a set $\mathcal{G}$ with an associated composition operator $\odot$ that satisfies:
  - **Closure**: $a \odot b \in \mathcal{G}$, $\forall a, b \in \mathcal{G}$
  - **Identity element**: $\exists e \in \mathcal{G}$ (unique) such that $e \odot a = a \odot e = a$
  - **Inverse element**: for $a \in \mathcal{G}$, $\exists b \in G$ (unique) such that $a \odot b = b \odot a = e$
  - **Associativity**: $(a \odot b) \odot c = a \odot (b \odot c)$, $\forall a, b, c, \in \mathcal{G}$
- The notion of a group is weaker than a vector space because it does not require commutativity and does not have scalar multiplication and its associated axioms (compatibility, identity, inverse, distributivity)
- **General linear group** $GL(n; \mathbb{C})$: the set of all invertible matrices in $\mathbb{C}^{n \times n}$
- A **subgroup** of group $\mathcal{G}$ is a subset that contains the identity of $\mathcal{G}$ and is closed under group composition and inverse
- **Lie group**: set $\mathcal{G}$ that is both a smooth manifold and a group with smooth composition $\odot : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ and inverse $(\cdot)^{-1} : \mathcal{G} \to \mathcal{G}$
- **Matrix Lie group**: subgroup of $GL(n; \mathbb{C})$ and embedded submanifold of $\mathbb{C}^{n \times n}$

Lie Algebra

- A **Lie algebra** is a vector space $\mathfrak{g}$ over some field $\mathcal{F}$ with a binary operation, $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called a **Lie bracket**
- For all $X, Y, Z \in \mathfrak{g}$ and $a, b \in \mathcal{F}$, the Lie bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfies:

$$\begin{aligned} \text{bilinearity :} \quad & [aX + bY, Z] = a[X, Z] + b[Y, Z] \\ & [Z, aX + bY] = a[Z, X] + b[Z, Y] \\ \text{skew-symmetry :} \quad & [X, Y] = -[Y, X] \\ \text{Jacobi identity :} \quad & [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \end{aligned}$$

- The **adjoint** $ad_X : \mathfrak{g} \to \mathfrak{g}$ of a Lie algebra at $X \in \mathfrak{g}$ is:

$$ad_X(Y) = [X, Y]$$

- Example: $\mathbb{R}^3$ with $[\mathbf{x}, \mathbf{y}] = \mathbf{x} \times \mathbf{y}$ is a Lie algebra

Lie Group and Lie Algebra

- Each matrix Lie group $\mathcal{G}$ has an associated Lie algebra $\mathfrak{g}$
- The Lie algebra $\mathfrak{g}$ of a matrix Lie group $\mathcal{G}$ is the set of all matrices X whose matrix exponential $\exp(tX)$ is in $\mathcal{G}$ for all $t \in \mathbb{R}$:

$$\mathfrak{g} = \{X \mid \exp(tX) \in \mathcal{G},\ \forall t \in \mathbb{R}\}$$

- The Lie algebra $\mathfrak{g}$ of a Lie group $\mathcal{G}$ is the tangent space at identity $T_I\mathcal{G}$
  - For $X \in \mathfrak{g}$, let $\gamma(t) = \exp(tX)$ such that $\gamma(0) = I$ and $\gamma'(0) = X$
- The **adjoint** $Ad_A : \mathfrak{g} \to \mathfrak{g}$ of a Lie group $\mathcal{G}$ at $A \in \mathcal{G}$ is:

$$Ad_A(Y) = AYA^{-1}$$

- The algebra adjoint ad_X is the derivative of the group adjoint Ad_A at $A = I$:

$$Ad_{\exp(X)} = \exp(ad_X) \qquad\qquad ad_X = \frac{d}{dt} Ad_{\exp(tX)}\bigg|_{t=0}$$

- Let $\mathcal{G}$ be a matrix Lie group with Lie algebra $\mathfrak{g}$. For $X, Y \in \mathfrak{g}$:
  - $tX \in \mathfrak{g}$ for all $t \in \mathbb{R}$
  - $X + Y \in \mathfrak{g}$
  - $ad_X(Y) = [X, Y] = XY - YX \in \mathfrak{g}$
  - $Ad_A(X) = AXA^{-1} \in \mathfrak{g}$ for all $A \in \mathcal{G}$

Lie Group and Lie Algebra

- The **exponential** and **logarithm** maps relate a matrix Lie group $\mathcal{G}$ with its Lie algebra $\mathfrak{g}$:

$$\exp(A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n \qquad \log(A) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (A - I)^n$$

- **Theorem**: Let $\mathcal{V}_\epsilon = \{X \in \mathbb{C}^{n \times n} \mid \|X\| < \epsilon\}$ and $\mathcal{U}_\epsilon = \exp(\mathcal{V}_\epsilon)$. Suppose $\mathcal{G}$ is a matrix Lie group with Lie algebra $\mathfrak{g}$. Then, there exists $\epsilon \in (0, \log 2)$ such that for all $A \in \mathcal{U}_\epsilon$, $A \in \mathcal{G}$ if and only if $\log(A) \in \mathfrak{g}$.

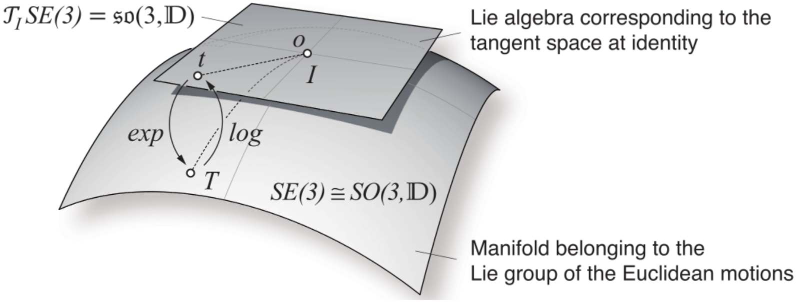


Figure: $SE(3)$ and corresponding Lie algebra $\mathfrak{se}(3)$ as tangent space at identity

Outline

Special Orthogonal Lie Group $SO(3)$

- $SO(3) := \left\{R \in \mathbb{R}^{3\times 3} \mid R^\top R = I,\ \det(R) = 1\right\}$

- $SO(3)$ is a group:
  - **Closure**: $R_1 R_2 \in SO(3)$
  - **Identity**: $I \in SO(3)$
  - **Inverse**: $R^{-1} = R^\top \in SO(3)$
  - **Associativity**: $(R_1 R_2) R_3 = R_1 (R_2 R_3)$ for all $R_1, R_2, R_3 \in SO(3)$

- $SO(3)$ is an embedded submanifold of $\mathbb{R}^{3\times 3}$ with local defining function:

$$h(R) = (R^\top R - I, \det(R) - 1)$$

- The tangent space of $SO(3)$ is:

$$T_R SO(3) = \ker(Dh(R)) = \left\{V \in \mathbb{R}^{3\times 3} \mid R^\top V + V^\top R = 0,\ \operatorname{tr}(R^\top V) = 0\right\}$$

- $SO(3)$ is a **matrix Lie group**

Special Orthogonal Lie Algebra $\mathfrak{so}(3)$

- The **Lie algebra** of $SO(3)$ is the space of skew-symmetric matrices:

$$\mathfrak{so}(3) = T_I SO(3) = \{\hat{\boldsymbol{\theta}} \in \mathbb{R}^{3\times 3} \mid \boldsymbol{\theta} \in \mathbb{R}^3\}$$

- The **Lie bracket** of $\mathfrak{so}(3)$ is:

$$[\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2] = \hat{\boldsymbol{\theta}}_1\hat{\boldsymbol{\theta}}_2 - \hat{\boldsymbol{\theta}}_2\hat{\boldsymbol{\theta}}_1 = \left(\hat{\boldsymbol{\theta}}_1\boldsymbol{\theta}_2\right)^{\wedge} \in \mathfrak{so}(3)$$

- The elements $R \in SO(3)$ are related to the elements $\hat{\boldsymbol{\theta}} \in \mathfrak{so}(3)$ through the exponential and logarithm maps:

$$R = \exp(\hat{\boldsymbol{\theta}}) = \sum_{n=0}^{\infty} \frac{1}{n!}(\hat{\boldsymbol{\theta}})^n = I + \left(\frac{\sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|}\right)\hat{\boldsymbol{\theta}} + \left(\frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right)\hat{\boldsymbol{\theta}}^2$$

$$\hat{\boldsymbol{\theta}} = \log(R) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(R-I)^n = \frac{\|\boldsymbol{\theta}\|}{2\sin\|\boldsymbol{\theta}\|}(R - R^\top)$$

$$\|\boldsymbol{\theta}\| = \arccos\left(\frac{\operatorname{tr}(R)-1}{2}\right)$$

Distance in $SO(3)$

- What is the distance between two rotations $R_1, R_2 \in SO(3)$?
- **Inner product on $\mathfrak{so}(3)$:**

$$\langle \hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2 \rangle = \frac{1}{2} \operatorname{tr}\left(\hat{\boldsymbol{\theta}}_1^\top \hat{\boldsymbol{\theta}}_2\right) = \boldsymbol{\theta}_1^\top \boldsymbol{\theta}_2$$

- **Geodesic distance on $SO(3)$:** the length of the shortest path between R_1 and R_2 on the $SO(3)$ manifold is equal to the rotation angle $\|\boldsymbol{\theta}_{12}\|_2$ of the axis-angle representation $\boldsymbol{\theta}_{12}$ of the relative rotation $R_{12} = R_1^\top R_2$:

$$\begin{aligned}\boldsymbol{\theta}_{12} &= \log\left(R_1^\top R_2\right)^\vee \\ d_\theta(R_1, R_2) &= \sqrt{\langle \hat{\boldsymbol{\theta}}_{12}, \hat{\boldsymbol{\theta}}_{12} \rangle} = \|\boldsymbol{\theta}_{12}\|_2 = \left|\arccos\left(\frac{\operatorname{tr}(R_1^\top R_2) - 1}{2}\right)\right|\end{aligned}$$

Distance in $SO(3)$

- **Chordal distance on** $SO(3)$:

$$d_c(R_1, R_2) = \|R_1 - R_2\|_F = \sqrt{\mathrm{tr}\left((R_1 - R_2)^\top (R_1 - R_2)\right)} = 2\sqrt{2}\left|\sin\left(\frac{\|\boldsymbol{\theta}_{12}\|}{2}\right)\right|$$

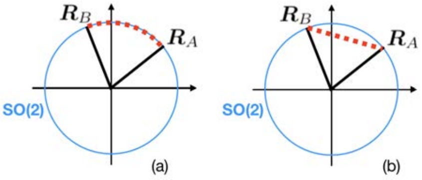


Figure: (a) Geodesic and (b) chordal distance in $SO(2)$

Baker-Campbell-Hausdorff Formulas

- The **left Jacobian** of $SO(3)$ is the matrix:

$$J_L(\boldsymbol{\theta}) := \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \left(\hat{\boldsymbol{\theta}}\right)^n \qquad\qquad R = I + \hat{\boldsymbol{\theta}} J_L(\boldsymbol{\theta})$$

- The **right Jacobian** of $SO(3)$ is the matrix:

$$J_R(\boldsymbol{\theta}) := \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \left(-\hat{\boldsymbol{\theta}}\right)^n \qquad J_R(\boldsymbol{\theta}) = J_L(-\boldsymbol{\theta}) = J_L(\boldsymbol{\theta})^\top = R^\top J_L(\boldsymbol{\theta})$$

- **Baker-Campbell-Hausdorff Formulas**: the $SO(3)$ Jacobians relate small perturbations $\delta\boldsymbol{\theta}$ in $\mathfrak{so}(3)$ to small perturbations in $SO(3)$:

$$\begin{aligned}\exp\left((\boldsymbol{\theta} + \delta\boldsymbol{\theta})^\wedge\right) &\approx \exp(\hat{\boldsymbol{\theta}}) \exp\left((J_R(\boldsymbol{\theta})\delta\boldsymbol{\theta})^\wedge\right) \\ &\approx \exp\left((J_L(\boldsymbol{\theta})\delta\boldsymbol{\theta})^\wedge\right) \exp(\hat{\boldsymbol{\theta}})\end{aligned}$$

$$\log(\exp(\hat{\boldsymbol{\theta}}_1)\exp(\hat{\boldsymbol{\theta}}_2))^\vee \approx \begin{cases} J_L(\boldsymbol{\theta}_2)^{-1}\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 & \text{if } \boldsymbol{\theta}_1 \text{ is small} \\ \boldsymbol{\theta}_1 + J_R(\boldsymbol{\theta}_1)^{-1}\boldsymbol{\theta}_2 & \text{if } \boldsymbol{\theta}_2 \text{ is small} \end{cases}$$

Closed-forms of the $SO(3)$ Jacobians

$$
\begin{aligned}
J_L(\boldsymbol{\theta}) &= I + \left(\frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right)\hat{\boldsymbol{\theta}} + \left(\frac{\|\boldsymbol{\theta}\| - \sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3}\right)\hat{\boldsymbol{\theta}}^2 \approx I + \frac{1}{2}\hat{\boldsymbol{\theta}} \\
J_L(\boldsymbol{\theta})^{-1} &= I - \frac{1}{2}\hat{\boldsymbol{\theta}} + \left(\frac{1}{\|\boldsymbol{\theta}\|^2} - \frac{1+\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|\sin\|\boldsymbol{\theta}\|}\right)\hat{\boldsymbol{\theta}}^2 \approx I - \frac{1}{2}\hat{\boldsymbol{\theta}}
\end{aligned}
$$

$$
\begin{aligned}
J_R(\boldsymbol{\theta}) &= I - \left(\frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right)\hat{\boldsymbol{\theta}} + \left(\frac{\|\boldsymbol{\theta}\| - \sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3}\right)\hat{\boldsymbol{\theta}}^2 \approx I - \frac{1}{2}\hat{\boldsymbol{\theta}} \\
J_R(\boldsymbol{\theta})^{-1} &= I + \frac{1}{2}\hat{\boldsymbol{\theta}} + \left(\frac{1}{\|\boldsymbol{\theta}\|^2} - \frac{1+\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|\sin\|\boldsymbol{\theta}\|}\right)\hat{\boldsymbol{\theta}}^2 \approx I + \frac{1}{2}\hat{\boldsymbol{\theta}}
\end{aligned}
$$

$$
\begin{aligned}
J_L(\boldsymbol{\theta})J_L(\boldsymbol{\theta})^T &= I + \left(1 - 2\frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right)\hat{\boldsymbol{\theta}}^2 \succ 0 \\
\left(J_L(\boldsymbol{\theta})J_L(\boldsymbol{\theta})^T\right)^{-1} &= I + \left(1 - 2\frac{\|\boldsymbol{\theta}\|^2}{1-\cos\|\boldsymbol{\theta}\|}\right)\hat{\boldsymbol{\theta}}^2
\end{aligned}
$$

Integration in $SO(3)$

- The geodesic distance between a rotation $R = \exp(\hat{\boldsymbol{\theta}})$ and a small perturbation $\exp((\boldsymbol{\theta} + \delta\boldsymbol{\theta})^\wedge)$ can be approximated using the BCH formulas:

$$\log\left(\exp(\hat{\boldsymbol{\theta}})^\top \exp((\boldsymbol{\theta} + \delta\boldsymbol{\theta})^\wedge)\right)^\vee \approx \log\left(R^\top R \exp\left((J_R(\boldsymbol{\theta})\delta\boldsymbol{\theta})^\wedge\right)\right)^\vee = J_R(\boldsymbol{\theta})\delta\boldsymbol{\theta}$$

- This allows to define an infinitesimal volume element:

$$dR = |\det(J_R(\boldsymbol{\theta}))| d\boldsymbol{\theta} = 2\left(\frac{1 - \cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right) d\boldsymbol{\theta} \qquad \det(J_R(\boldsymbol{\theta})) = \det(J_L(\boldsymbol{\theta}))$$

- Integrating functions of rotations can be carried out as follows:

$$\int_{SO(3)} f(R) dR = \int_{\|\boldsymbol{\theta}\| < \pi} f\left(\exp(\hat{\boldsymbol{\theta}})\right) |\det(J_R(\boldsymbol{\theta}))| d\boldsymbol{\theta}$$

Adjoint $SO(3)$ Lie Group and Lie Algebra

- The adjoint operator $Ad_A : \mathfrak{g} \to \mathfrak{g}$ represents the elements A of a Lie group $\mathcal{G}$ as linear transformations on the Lie algebra $\mathfrak{g}$

- The adjoint Ad_R at $R \in SO(3)$ transforms $\hat{\boldsymbol{\omega}} \in \mathfrak{so}(3)$ from one coordinate frame (e.g., body frame) to another (e.g., world frame):

$$Ad_R(\hat{\boldsymbol{\omega}}) = R\hat{\boldsymbol{\omega}}R^{-1} = (R\boldsymbol{\omega})^{\wedge}$$

- The adjoint operator $Ad_R(\hat{\boldsymbol{\omega}})$ is linear and can be represented as a matrix R acting on $\boldsymbol{\omega} \in \mathbb{R}^3$

- The space of adjoint operators on $SO(3)$ is a matrix Lie group $Ad(SO(3)) \cong SO(3)$ with associated Lie algebra $ad(\mathfrak{so}(3)) \cong \mathfrak{so}(3)$

Outline

Special Euclidean Lie Group $SE(3)$

- $SE(3) := \left\{ T = \begin{bmatrix} R & \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4} \,\middle|\, R \in SO(3), \mathbf{p} \in \mathbb{R}^3 \right\}$
- $SE(3)$ is a group:
  - **Closure**: $T_1 T_2 = \begin{bmatrix} R_1 & \mathbf{p}_1 \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} R_2 & \mathbf{p}_2 \\ \mathbf{0}^\top & 1 \end{bmatrix} = \begin{bmatrix} R_1 R_2 & R_1 \mathbf{p}_2 + \mathbf{p}_1 \\ \mathbf{0}^\top & 1 \end{bmatrix} \in SE(3)$
  - **Identity**: $I \in SE(3)$
  - **Inverse**: $\begin{bmatrix} R & \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^\top & -R^\top \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in SE(3)$
  - **Associativity**: $(T_1 T_2) T_3 = T_1 (T_2 T_3)$ for all T_1, T_2, $T_3 \in SE(3)$
- $SE(3)$ is an embedded submanifold of $\mathbb{R}^{4\times 4}$
- $SE(3)$ is a **matrix Lie group**

Special Euclidean Lie Algebra $\mathfrak{se}(3)$

- The **Lie algebra** of $SE(3)$ is the space of twist matrices:

$$\mathfrak{se}(3) := T_I SE(3) = \left\{ \hat{\boldsymbol{\xi}} := \begin{bmatrix} \hat{\boldsymbol{\theta}} & \boldsymbol{\rho} \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4} \;\middle|\; \boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix} \in \mathbb{R}^6 \right\}$$

- The **Lie bracket** of $\mathfrak{se}(3)$ is:

$$[\hat{\boldsymbol{\xi}}_1, \hat{\boldsymbol{\xi}}_2] = \hat{\boldsymbol{\xi}}_1 \hat{\boldsymbol{\xi}}_2 - \hat{\boldsymbol{\xi}}_2 \hat{\boldsymbol{\xi}}_1 = \left(\overset{\curlywedge}{\boldsymbol{\xi}}_1 \boldsymbol{\xi}_2 \right)^{\wedge} \in \mathfrak{se}(3) \qquad \overset{\curlywedge}{\boldsymbol{\xi}} := \begin{bmatrix} \hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\rho}} \\ 0 & \hat{\boldsymbol{\theta}} \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

- The elements $T \in SE(3)$ are related to the elements $\hat{\boldsymbol{\xi}} \in \mathfrak{se}(3)$ through the exponential and logarithm maps:

$$T = \exp(\hat{\boldsymbol{\xi}}) = \sum_{n=0}^{\infty} \frac{1}{n!} (\hat{\boldsymbol{\xi}})^n$$

$$\hat{\boldsymbol{\xi}} = \log(T) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (T - I)^n$$

Exponential Map from $\mathfrak{se}(3)$ to $SE(3)$

- **Exponential map** $\exp : \mathfrak{se}(3) \to SE(3)$: has closed-form expression obtained using $\hat{\boldsymbol{\xi}}^4 + \|\boldsymbol{\theta}\|^2 \hat{\boldsymbol{\xi}}^2 = 0$:

$$T = \exp(\hat{\boldsymbol{\xi}}) = \begin{bmatrix} \exp(\hat{\boldsymbol{\theta}}) & J_L(\boldsymbol{\theta})\boldsymbol{\rho} \\ \mathbf{0}^T & 1 \end{bmatrix} = \sum_{n=0}^{\infty} \frac{1}{n!} \hat{\boldsymbol{\xi}}^n =$$

$$= I + \hat{\boldsymbol{\xi}} + \left(\frac{1 - \cos \|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2} \right) \hat{\boldsymbol{\xi}}^2 + \left(\frac{\|\boldsymbol{\theta}\| - \sin \|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3} \right) \hat{\boldsymbol{\xi}}^3$$

- The exponential map is **surjective** but **not injective**, i.e., every element of $SE(3)$ can be generated from multiple elements of $\mathfrak{se}(3)$

- **Logarithm map** $\log : SE(3) \to \mathfrak{se}(3)$: for any $T \in SE(3)$, there exists a (non-unique) $\boldsymbol{\xi} \in \mathbb{R}^6$ such that:

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix} = \log(T)^{\vee} = \begin{cases} \boldsymbol{\theta} = \log(R)^{\vee}, \; \boldsymbol{\rho} = J_L^{-1}(\boldsymbol{\theta})\mathbf{p}, & \text{if } R \neq I, \\ \boldsymbol{\theta} = 0, \; \boldsymbol{\rho} = \mathbf{p}, & \text{if } R = I. \end{cases}$$

Distance in $SE(3)$

- **Inner product** on $\mathfrak{se}(3)$:

$$\langle \hat{\boldsymbol{\xi}}_1, \hat{\boldsymbol{\xi}}_2 \rangle = \mathrm{tr}\left(\hat{\boldsymbol{\xi}}_1 \begin{bmatrix} \frac{1}{2} I & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \hat{\boldsymbol{\xi}}_2^\top \right) = \boldsymbol{\xi}_1^\top \boldsymbol{\xi}_2$$

- **Distance on** $SE(3)$: induced by the inner product on $\mathfrak{se}(3)$ evaluated at the vector representation $\hat{\boldsymbol{\xi}}_{12}$ of the relative pose $T_{12} = T_1^{-1} T_2$:

$$\boldsymbol{\xi}_{12} = \log(T_1^{-1} T_2)^\vee$$

$$d(T_1, T_2) = \sqrt{\langle \hat{\boldsymbol{\xi}}_{12}, \hat{\boldsymbol{\xi}}_{12} \rangle} = \|\boldsymbol{\xi}_{12}\|_2$$

Baker-Campbell-Hausdorff Formulas

- **Left Jacobian of** $SE(3)$: $\mathcal{J}_L(\boldsymbol{\xi}) = \begin{bmatrix} J_L(\boldsymbol{\theta}) & Q_L(\boldsymbol{\xi}) \\ 0 & J_L(\boldsymbol{\theta}) \end{bmatrix}$
- **Right Jacobian of** $SE(3)$: $\mathcal{J}_R(\boldsymbol{\xi}) = \begin{bmatrix} J_R(\boldsymbol{\theta}) & Q_R(\boldsymbol{\xi}) \\ 0 & J_R(\boldsymbol{\theta}) \end{bmatrix}$
- **Baker-Campbell-Hausdorff Formulas**: the $SE(3)$ Jacobians relate small perturbations $\delta\boldsymbol{\xi}$ in $\mathfrak{se}(3)$ to small perturbations in $SE(3)$:

$$\begin{aligned} \exp\left((\boldsymbol{\xi} + \delta\boldsymbol{\xi})^\wedge\right) &\approx \exp(\hat{\boldsymbol{\xi}}) \exp\left((\mathcal{J}_R(\boldsymbol{\xi})\delta\boldsymbol{\xi})^\wedge\right) \\ &\approx \exp\left((\mathcal{J}_L(\boldsymbol{\xi})\delta\boldsymbol{\xi})^\wedge\right) \exp(\hat{\boldsymbol{\xi}}) \end{aligned}$$

$$\log(\exp(\hat{\boldsymbol{\xi}}_1)\exp(\hat{\boldsymbol{\xi}}_2))^\vee \approx \begin{cases} \mathcal{J}_L(\boldsymbol{\xi}_2)^{-1}\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 & \text{if } \boldsymbol{\xi}_1 \text{ is small} \\ \boldsymbol{\xi}_1 + \mathcal{J}_R(\boldsymbol{\xi}_1)^{-1}\boldsymbol{\xi}_2 & \text{if } \boldsymbol{\xi}_2 \text{ is small} \end{cases}$$

Closed-forms of the $SE(3)$ Jacobians

$$
\begin{aligned}
\mathcal{J}_L(\boldsymbol{\xi}) &= \sum_{n=0}^{\infty} \frac{1}{(n+1)!}(\boldsymbol{\xi}^{\curlywedge})^n = \begin{bmatrix} J_L(\boldsymbol{\theta}) & Q_L(\boldsymbol{\xi}) \\ 0 & J_L(\boldsymbol{\theta}) \end{bmatrix} \\
&= I + \left(\frac{4 - \|\boldsymbol{\theta}\| \sin\|\boldsymbol{\theta}\| - 4\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^2}\right)\boldsymbol{\xi}^{\curlywedge} + \left(\frac{4\|\boldsymbol{\theta}\| - 5\sin\|\boldsymbol{\theta}\| + \|\boldsymbol{\theta}\|\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^3}\right){\boldsymbol{\xi}^{\curlywedge}}^2 \\
&\quad + \left(\frac{2 - \|\boldsymbol{\theta}\| \sin\|\boldsymbol{\theta}\| - 2\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^4}\right){\boldsymbol{\xi}^{\curlywedge}}^3 + \left(\frac{2\|\boldsymbol{\theta}\| - 3\sin\|\boldsymbol{\theta}\| + \|\boldsymbol{\theta}\|\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^5}\right){\boldsymbol{\xi}^{\curlywedge}}^4 \\
&\approx I + \frac{1}{2}\boldsymbol{\xi}^{\curlywedge}
\end{aligned}
$$

$$
\mathcal{J}_L(\boldsymbol{\xi})^{-1} = \begin{bmatrix} J_L(\boldsymbol{\theta})^{-1} & -J_L(\boldsymbol{\theta})^{-1} Q_L(\boldsymbol{\xi}) J_L(\boldsymbol{\theta})^{-1} \\ \mathbf{0} & J_L(\boldsymbol{\theta})^{-1} \end{bmatrix} \approx I - \frac{1}{2}\boldsymbol{\xi}^{\curlywedge}
$$

$$
\begin{aligned}
Q_L(\boldsymbol{\xi}) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{1}{(n+m+2)!}\hat{\boldsymbol{\theta}}^n \hat{\boldsymbol{\rho}} \hat{\boldsymbol{\theta}}^m \\
&= \frac{1}{2}\hat{\boldsymbol{\rho}} + \left(\frac{\|\boldsymbol{\theta}\| - \sin\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^3}\right)\left(\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\rho}} + \hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}}\right) + \left(\frac{\|\boldsymbol{\theta}\|^2 + 2\cos\|\boldsymbol{\theta}\| - 2}{2\|\boldsymbol{\theta}\|^4}\right)\left(\hat{\boldsymbol{\theta}}^2\hat{\boldsymbol{\rho}} + \hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}}^2 - 3\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}}\right) \\
&\quad + \left(\frac{2\|\boldsymbol{\theta}\| - 3\sin\|\boldsymbol{\theta}\| + \|\boldsymbol{\theta}\|\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^5}\right)\left(\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}}^2 + \hat{\boldsymbol{\theta}}^2\hat{\boldsymbol{\rho}}\hat{\boldsymbol{\theta}}\right)
\end{aligned}
$$

$$
Q_R(\boldsymbol{\xi}) = Q_L(-\boldsymbol{\xi}) = RQ_L(\boldsymbol{\xi}) + (J_L(\boldsymbol{\theta})\boldsymbol{\rho})^{\wedge} R J_L(\boldsymbol{\theta})
$$

Integration in $SE(3)$

- The distance between a pose $T = \exp(\hat{\xi})$ and a small perturbation $\exp((\boldsymbol{\xi} + \delta\boldsymbol{\xi})^\wedge)$ can be approximated using the BCH formulas:

$$\log\left(\exp(\hat{\xi})^{-1}\exp((\boldsymbol{\xi} + \delta\boldsymbol{\xi})^\wedge)\right)^\vee \approx \mathcal{J}_R(\boldsymbol{\xi})\delta\boldsymbol{\xi}$$

- This allows to define an infinitesimal volume element:

$$dT = |\det(\mathcal{J}_R(\boldsymbol{\xi}))|d\boldsymbol{\xi} = |\det(J_R(\boldsymbol{\theta}))|^2 d\boldsymbol{\xi} = 4\left(\frac{1-\cos\|\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|^2}\right)^2 d\boldsymbol{\xi}$$

- Integrating functions of poses can then be carried out as follows:

$$\int_{SE(3)} f(T)dT = \int_{\mathbb{R}^3, \|\boldsymbol{\theta}\| < \pi} f\left(\exp(\hat{\xi})\right)|det(\mathcal{J}_R(\boldsymbol{\xi}))|d\boldsymbol{\xi}$$

Adjoint $SE(3)$ Lie Group and Lie Algebra

- The adjoint Ad_T at $T \in SE(3)$ transforms $\hat{\boldsymbol{\zeta}} \in \mathfrak{se}(3)$ from one coordinate frame to another:
$$Ad_T(\hat{\boldsymbol{\zeta}}) = T\hat{\boldsymbol{\zeta}}T^{-1} = (\mathcal{T}\boldsymbol{\zeta})^{\wedge}$$
- The adjoint operator Ad_T is linear and can be represented as a matrix $\mathcal{T}$ acting on $\boldsymbol{\zeta} \in \mathbb{R}^6$:
$$\mathcal{T} = \begin{bmatrix} R & \hat{\mathbf{p}}R \\ \mathbf{0} & R \end{bmatrix} \in \mathbb{R}^{6\times 6}$$
- The space of adjoint operators on $SE(3)$ is a matrix Lie group:
$$Ad(SE(3)) = \left\{ \mathcal{T} = \begin{bmatrix} R & \hat{\mathbf{p}}R \\ \mathbf{0} & R \end{bmatrix} \in \mathbb{R}^{6\times 6} \;\middle|\; T = \begin{bmatrix} R & \mathbf{p} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in SE(3) \right\}$$
- The Lie algebra associated with $Ad(SE(3))$ is:
$$ad(\mathfrak{se}(3)) = \left\{ \overset{\curlywedge}{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\rho}} \\ \mathbf{0} & \hat{\boldsymbol{\theta}} \end{bmatrix} \in \mathbb{R}^{6\times 6} \;\middle|\; \boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix} \in \mathbb{R}^6 \right\}$$

Rodrigues Formula for the Adjoint of $SE(3)$

- **Rodrigues Formula**: using $(\overset{\curlywedge}{\boldsymbol{\xi}})^5 + 2\|\boldsymbol{\theta}\|^2(\overset{\curlywedge}{\boldsymbol{\xi}})^3 + \|\boldsymbol{\theta}\|^4\overset{\curlywedge}{\boldsymbol{\xi}} = 0$ we can obtain a direct expression of $\mathcal{T} \in Ad(SE(3))$ in terms of $\boldsymbol{\xi} = \begin{bmatrix}\boldsymbol{\rho}\\ \boldsymbol{\theta}\end{bmatrix} \in \mathbb{R}^6$:

$$
\begin{aligned}
\mathcal{T} = Ad(T) = \exp\left(\overset{\curlywedge}{\boldsymbol{\xi}}\right) &= \begin{bmatrix}\exp(\hat{\boldsymbol{\theta}}) & (J_L(\boldsymbol{\theta})\boldsymbol{\rho})^{\wedge}\exp(\hat{\boldsymbol{\theta}})\\ \mathbf{0} & \exp(\hat{\boldsymbol{\theta}})\end{bmatrix} = \sum_{n=0}^{\infty}\frac{1}{n!}(\overset{\curlywedge}{\boldsymbol{\xi}})^n\\
&= I + \left(\frac{3\sin\|\boldsymbol{\theta}\| - \|\boldsymbol{\theta}\|\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|}\right)\overset{\curlywedge}{\boldsymbol{\xi}} + \left(\frac{4 - \|\boldsymbol{\theta}\|\sin\|\boldsymbol{\theta}\| - 4\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^2}\right)(\overset{\curlywedge}{\boldsymbol{\xi}})^2\\
&\quad + \left(\frac{\sin\|\boldsymbol{\theta}\| - \|\boldsymbol{\theta}\|\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^3}\right)(\overset{\curlywedge}{\boldsymbol{\xi}})^3 + \left(\frac{2 - \|\boldsymbol{\theta}\|\sin\|\boldsymbol{\theta}\| - 2\cos\|\boldsymbol{\theta}\|}{2\|\boldsymbol{\theta}\|^4}\right)(\overset{\curlywedge}{\boldsymbol{\xi}})^4
\end{aligned}
$$

- The exponential map is **surjective** but **not injective**, i.e., every element of $Ad(SE(3))$ can be generated from multiple elements of $ad(\mathfrak{se}(3))$

Distance in $Ad(SE(3))$

- **Inner product** on $ad(\mathfrak{se}(3))$:

$$\langle \overset{\curlywedge}{\boldsymbol{\xi}}_1, \overset{\curlywedge}{\boldsymbol{\xi}}_2 \rangle = \operatorname{tr}\left(\overset{\curlywedge}{\boldsymbol{\xi}}_1 \begin{bmatrix} \frac{1}{4} I & \mathbf{0} \\ \mathbf{0} & \frac{1}{2} I \end{bmatrix} \overset{\curlywedge}{\boldsymbol{\xi}}_2^{\top} \right) = \boldsymbol{\xi}_1^{\top} \boldsymbol{\xi}_2$$

- **Distance on** $Ad(SE(3))$: induced by the inner product on $ad(\mathfrak{se}(3))$ evaluated at the vector representation $\overset{\curlywedge}{\boldsymbol{\xi}}_{12}$ of $\mathcal{T}_{12} = \mathcal{T}_1^{-1} \mathcal{T}_2$:

$$\boldsymbol{\xi}_{12} = \log \left(\mathcal{T}_1^{-1} \mathcal{T}_2 \right)^{\curlyvee}$$

$$d(\mathcal{T}_1, \mathcal{T}_2) = \sqrt{\langle \overset{\curlywedge}{\boldsymbol{\xi}}_{12}, \overset{\curlywedge}{\boldsymbol{\xi}}_{12} \rangle} = \|\boldsymbol{\xi}_{12}\|_2$$

Pose Lie Groups and Lie Algebras

$$\begin{array}{cccc} & \text{Lie algebra} & & \text{Lie group} \\ 4\times 4 & \boldsymbol{\xi}^\wedge \in \mathfrak{se}(3) & \xrightarrow{\exp} & \mathbf{T} \in SE(3) \\ & \big\downarrow \text{ad} & & \big\downarrow \text{Ad} \\ 6\times 6 & \boldsymbol{\xi}^\curlywedge \in \text{ad}(\mathfrak{se}(3)) & \xrightarrow{\exp} & \boldsymbol{\mathcal{T}} \in \text{Ad}(SE(3)) \end{array}$$

$$\mathcal{T} = Ad\underbrace{\left(\exp(\hat{\xi})\right)}_{T} = \exp\underbrace{\left(ad(\hat{\xi})\right)}_{\overset{\curlywedge}{\xi}} \qquad \xi = \begin{bmatrix}\boldsymbol{\rho}\\ \boldsymbol{\theta}\end{bmatrix} \in \mathbb{R}^6$$

$$= Ad\left(\exp\left(\begin{bmatrix}\hat{\boldsymbol{\theta}} & \boldsymbol{\rho}\\ \mathbf{0}^T & 0\end{bmatrix}\right)\right) = \exp\left(ad\left(\begin{bmatrix}\hat{\boldsymbol{\theta}} & \boldsymbol{\rho}\\ \mathbf{0}^T & 0\end{bmatrix}\right)\right)$$

$$= Ad\left(\begin{bmatrix}\exp(\hat{\boldsymbol{\theta}}) & J_L(\boldsymbol{\theta})\boldsymbol{\rho}\\ \mathbf{0}^T & 1\end{bmatrix}\right) = \exp\left(\begin{bmatrix}\hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\rho}}\\ \mathbf{0} & \hat{\boldsymbol{\theta}}\end{bmatrix}\right)$$

$$= \begin{bmatrix}\exp(\hat{\boldsymbol{\theta}}) & (J_L(\boldsymbol{\theta})\boldsymbol{\rho})^\wedge \exp(\hat{\boldsymbol{\theta}})\\ \mathbf{0} & \exp(\hat{\boldsymbol{\theta}})\end{bmatrix}$$

$\mathfrak{se}(3)$ Identities

$$\hat{\boldsymbol{\xi}} = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}^{\wedge} = \begin{bmatrix} \hat{\boldsymbol{\theta}} & \boldsymbol{\rho} \\ \mathbf{0}^{\top} & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4} \qquad \overset{\curlywedge}{\boldsymbol{\xi}} = ad(\hat{\boldsymbol{\xi}}) = \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}^{\curlywedge} = \begin{bmatrix} \hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\rho}} \\ \mathbf{0} & \hat{\boldsymbol{\theta}} \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

$$\overset{\curlywedge}{\boldsymbol{\zeta}}\boldsymbol{\xi} = -\overset{\curlywedge}{\boldsymbol{\xi}}\boldsymbol{\zeta} \qquad \boldsymbol{\zeta} \in \mathbb{R}^{6}$$

$$\overset{\curlywedge}{\boldsymbol{\xi}}\boldsymbol{\xi} = 0$$

$$\hat{\boldsymbol{\xi}}^{4} + \left(\mathbf{s}^{\top}\mathbf{s}\right)\hat{\boldsymbol{\xi}}^{2} = 0 \qquad \mathbf{s} \in \mathbb{R}^{3}$$

$$\left(\overset{\curlywedge}{\boldsymbol{\xi}}\right)^{5} + 2\left(\mathbf{s}^{\top}\mathbf{s}\right)\left(\overset{\curlywedge}{\boldsymbol{\xi}}\right)^{3} + \left(\mathbf{s}^{\top}\mathbf{s}\right)^{2}\overset{\curlywedge}{\boldsymbol{\xi}} = 0$$

$$\mathbf{m}^{\odot} := \begin{bmatrix} \mathbf{s} \\ \lambda \end{bmatrix}^{\odot} = \begin{bmatrix} \lambda I & -\hat{\mathbf{s}} \\ \mathbf{0}^{\top} & \mathbf{0}^{\top} \end{bmatrix} \in \mathbb{R}^{4\times 6} \qquad \mathbf{m}^{\circledcirc} := \begin{bmatrix} \mathbf{s} \\ \lambda \end{bmatrix}^{\circledcirc} = \begin{bmatrix} \mathbf{0} & \mathbf{s} \\ -\hat{\mathbf{s}} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{6\times 4}$$

$$\hat{\boldsymbol{\xi}}\mathbf{m} = \mathbf{m}^{\odot}\boldsymbol{\xi} \qquad \mathbf{m}^{\top}\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}^{\top}\mathbf{m}^{\circledcirc}$$

$SE(3)$ Identities

$$T = \exp\left(\hat{\boldsymbol{\xi}}\right) = \begin{bmatrix} \exp\left(\hat{\boldsymbol{\theta}}\right) & J_L(\boldsymbol{\theta})\boldsymbol{\rho} \\ \mathbf{0}^T & 1 \end{bmatrix} \qquad \begin{aligned} \det(T) &= 1 \\ \mathrm{tr}(T) &= 2\cos\|\boldsymbol{\theta}\| + 2 \end{aligned}$$

$$\mathcal{T} = Ad(T) = \exp\left(\overset{\curlywedge}{\boldsymbol{\xi}}\right) = \begin{bmatrix} \exp\left(\hat{\boldsymbol{\theta}}\right) & (J_L(\boldsymbol{\theta})\boldsymbol{\rho})^\wedge \exp\left(\hat{\boldsymbol{\theta}}\right) \\ \mathbf{0} & \exp\left(\hat{\boldsymbol{\theta}}\right) \end{bmatrix}$$

$$T\hat{\boldsymbol{\xi}} = \hat{\boldsymbol{\xi}}T$$

$$\mathcal{T}\boldsymbol{\xi} = \boldsymbol{\xi} \qquad \mathcal{T}\overset{\curlywedge}{\boldsymbol{\xi}} = \overset{\curlywedge}{\boldsymbol{\xi}}\mathcal{T}$$

$$(\mathcal{T}\boldsymbol{\zeta})^\wedge = T\hat{\boldsymbol{\zeta}}T^{-1} \qquad \overset{\curlywedge}{(\mathcal{T}\boldsymbol{\zeta})} = \mathcal{T}\overset{\curlywedge}{\boldsymbol{\zeta}}\mathcal{T}^{-1} \qquad \boldsymbol{\zeta} \in \mathbb{R}^6$$

$$\exp\left((\mathcal{T}\boldsymbol{\zeta})^\wedge\right) = T\exp\left(\hat{\boldsymbol{\zeta}}\right)T^{-1} \qquad \exp\left(\overset{\curlywedge}{(\mathcal{T}\boldsymbol{\zeta})}\right) = \mathcal{T}\exp\left(\overset{\curlywedge}{\boldsymbol{\zeta}}\right)\mathcal{T}^{-1}$$

$$(T\mathbf{m})^\odot = T\mathbf{m}^\odot\mathcal{T}^{-1} \qquad \left((T\mathbf{m})^\odot\right)^T(T\mathbf{m})^\odot = \mathcal{T}^{-T}\left(\mathbf{m}^\odot\right)^T\mathbf{m}^\odot\mathcal{T}^{-1}$$

Outline

Manifolds and Matrix Lie Groups

$SO(3)$ Geometry

$SE(3)$ Geometry

Manifold Optimization

Riemannian Manifold

- **Riemannian manifold**: a smooth manifold $\mathcal{M}$ equipped with a (Riemannian) metric $\langle \cdot, \cdot \rangle_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$ that varies smoothly with p
- Riemannian manifolds allow generalizing the notion of Euclidean distance to curved surfaces
- The shortest path between two points in Euclidean space is a straight line
- The shortest path between two points on a Riemannian manifold $\mathcal{M}$ is a **geodesic**, i.e., the shortest continuous curve on $\mathcal{M}$ connecting the two points
- **Smooth manifold function**: Let $\mathcal{N}$ be a smooth n-manifold and $\mathcal{M}$ be a smooth m-manifold. A function $f : \mathcal{N} \to \mathcal{M}$ is smooth at $p \in \mathcal{N}$ if, for any charts $(\mathcal{U}, \phi)$ around p and $(\mathcal{V}, \psi)$ around $f(p)$ with $f(\mathcal{U}) \subseteq \mathcal{V}$, its coordinate representation $\psi \circ f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ is smooth at $\phi(p)$

Riemannian Gradient

- A **vector field** on a manifold $\mathcal{M}$ is a map $V : \mathcal{M} \to T\mathcal{M}$ such that $V(p) \in T_p\mathcal{M}$ for all $p \in \mathcal{M}$
- **Riemannian gradient**: Let $f : \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $\mathcal{M}$. The Riemannian gradient of f is a vector field $\operatorname{grad} f : \mathcal{M} \to T\mathcal{M}$ uniquely defined by:

$$Df(p)[v] = \langle \operatorname{grad} f(p), v \rangle_p, \qquad \forall (p, v) \in T\mathcal{M}$$

- A **retraction** on a manifold $\mathcal{M}$ is a smooth map $R : T\mathcal{M} \to \mathcal{M}$ such that each curve $\gamma(t) = R_p(tv)$ satisfies $\gamma(0) = p$ and $\gamma'(0) = v$ for $(p, v) \in T\mathcal{M}$
- Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function on a Riemannian manifold $\mathcal{M}$ equipped with a retraction R. Then:

$$\operatorname{grad} f(p) = \nabla_v f(R_p(v))|_{v=0}$$

Relationship Between Riemannian and Euclidean Gradient

- Let $\mathcal{M}$ be a Riemannian manifold with metric $\langle \cdot, \cdot \rangle_p$ embedded in Euclidean space $\mathcal{E}$ with metric $\langle \cdot, \cdot \rangle$

- **Orthogonal projection to** $T_p\mathcal{M}$: linear map $\Pi_p : \mathcal{E} \to T_p\mathcal{M}$ that satisfies:
  - $\Pi_p(\Pi_p(u)) = \Pi_p(u)$ for all $u \in \mathcal{E}$
  - $\langle u - \Pi_p(u), v \rangle = 0$ for all $v \in T_p\mathcal{M}$ and $u \in \mathcal{E}$

- Let $f : \mathcal{E} \to \mathbb{R}$ be a smooth function. Since its Euclidean gradient $\nabla f(p)$ is a vector in $\mathcal{E}$ and $T_p\mathcal{M}$ is a subspace of $\mathcal{E}$, there is a unique decomposition:

$$\nabla f(p) = \nabla f(p)_{\parallel} + \nabla f(p)_{\perp}$$

  where $\nabla f(p)_{\parallel} = \Pi_p(\nabla f(p)) \in T_p\mathcal{M}$ and $\langle \nabla f(p)_{\perp}, v \rangle = 0$ for all $v \in T_p\mathcal{M}$

- **Relationship between Riemannian and Euclidean gradient**:

$$\langle \operatorname{grad} f(p), v \rangle_p = Df(p)[v] = \langle \nabla f(p)_{\parallel}, v \rangle = \langle \Pi_p(\nabla f(p)), v \rangle$$

Example: Riemannian Gradient in $SO(n)$ using Projection

- $SO(n)$ is a Riemannian manifold with metric $\langle X, Y\rangle_R = \frac{1}{2}\operatorname{tr}\left(X^\top Y\right)$ for $X, Y \in \mathfrak{so}(n)$
- $SO(n)$ is embedded in Euclidean space $\mathbb{R}^{n\times n}$ with metric $\langle X, Y\rangle = \operatorname{tr}(X^\top Y)$ for $X, Y \in \mathbb{R}^{n\times n}$
- The tangent space to $SO(n)$ at R is: $T_R SO(n) = \{R\hat{\boldsymbol{\omega}} \mid \hat{\boldsymbol{\omega}} \in \mathfrak{so}(n)\}$
- The orthogonal projection to $T_R SO(n)$ can be identified as:

$$\Pi_R(U) = R\frac{1}{2}\left(R^\top U - U^\top R\right)$$

- The Riemannian gradient of a smooth function $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ is related to its Euclidean gradient as:

$$\begin{aligned}\langle \operatorname{grad} f(R), \hat{\boldsymbol{\omega}}\rangle_R = \langle \Pi_R(\nabla f(R)), \hat{\boldsymbol{\omega}}\rangle \\ \Rightarrow \quad \operatorname{grad} f(R) &= 2\Pi_R(\nabla f(R)) \\ &= R(R^\top \nabla f(R) - \nabla f(R)^\top R) \\ &= \nabla f(R) - R\nabla f(R)^\top R\end{aligned}$$

Riemannian Gradient Descent

- Consider an optimization problem with smooth objective function $f : \mathcal{M} \to \mathbb{R}$ defined on a Riemannian manifold $\mathcal{M}$:

$$\min_{x \in \mathcal{M}} f(x)$$

- **Riemannian gradient descent**: given $x_0 \in \mathcal{M}$ and retraction R on $\mathcal{M}$:

$$x_{k+1} = R_{x_k}\left(-\alpha_k \operatorname{grad} f(x_k)\right)$$

where the step size α_k is obtained via line search:

$$\alpha_k \in \arg\min_{\alpha > 0} f(R_{x_k}(-\alpha \operatorname{grad} f(x_k)))$$

Riemannian Gradient Descent Convergence

Let $f : \mathcal{M} \to \mathbb{R}$ be smooth and bounded below, i.e., $f(x) \geq b$ for some $b \in \mathbb{R}$ and all $x \in \mathcal{M}$. Let the step size α_k ensure sufficient cost decrease for constant $c > 0$:

$$f(x_k) - f(x_{k+1}) \geq c \| \operatorname{grad} f(x_k) \|_2^2.$$

Then,

$$\lim_{k \to \infty} \| \operatorname{grad} f(x_k) \| = 0.$$

Lie Group Gradient Descent

- Consider $\min_{\mathbf{x}} f(\mathbf{x})$
- **Gradient descent in** $\mathbb{R}^d$: $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)$
- The gradient of f can be identified from the first-order Taylor series:

$$f(\mathbf{x} + \delta\mathbf{x}) \approx f(\mathbf{x}) + \nabla f(\mathbf{x})^\top \delta\mathbf{x}$$

- Consider $\min_{p \in \mathcal{G}} f(p)$
- On a matrix Lie group $\mathcal{G}$ with Lie algebra $\mathfrak{g}$, the exponential map $R_p(v) = p \exp(v)$ is a retraction that can be used to define $p + v$ for $p \in \mathcal{G}$ and $v \in \mathfrak{g}$
- **Gradient descent on** $\mathcal{G}$: $p_{k+1} = p_k \exp(-\alpha_k \operatorname{grad} f(p_k))$
- The Riemannian gradient of $f : \mathcal{G} \to \mathbb{R}$ can be identified from:

$$f(p \exp(v)) \approx f(p) + \langle \operatorname{grad} f(p), v \rangle_p \qquad (p, v) \in T\mathcal{G}$$

Example: Gradient Descent in $SO(3)$

- Consider $f(R,\mathbf{x}) = \mathbf{x}^\top R^\top A R \mathbf{x}$
- Euclidean gradient with respect to $\mathbf{x}$ using Taylor series:

$$\begin{aligned}
f(R,\mathbf{x}+\delta\mathbf{x}) &= (\mathbf{x}+\delta\mathbf{x})^\top R^\top A R(\mathbf{x}+\delta\mathbf{x}) \\
&= \mathbf{x}^\top R^\top A R\mathbf{x} + \mathbf{x}^\top R^\top A R\delta\mathbf{x} + \delta\mathbf{x}^\top R^\top A R\mathbf{x} + o(\|\delta\mathbf{x}\|_2^2) \\
&\approx f(R,\mathbf{x}) + \underbrace{\mathbf{x}^\top R^\top (A + A^\top) R}_{\nabla_{\mathbf{x}} f^\top} \delta\mathbf{x} \\
&\Rightarrow \nabla_{\mathbf{x}} f(R,\mathbf{x}) = R^\top (A+A^\top) R\mathbf{x}
\end{aligned}$$

- Verify using the product rule:

$$\begin{aligned}
\frac{d}{d\mathbf{x}} f(R,\mathbf{x}) &= \mathbf{x}^\top R^\top A R \frac{d\mathbf{x}}{d\mathbf{x}} + \mathbf{x}^\top R^\top A^\top R \frac{d\mathbf{x}}{d\mathbf{x}} \\
&= \mathbf{x}^\top R^\top (A+A^\top) R \\
&\Rightarrow \nabla_{\mathbf{x}} f(R,\mathbf{x}) = \left[\frac{d}{d\mathbf{x}} f(R,\mathbf{x})\right]^\top = R^\top (A+A^\top) R\mathbf{x}
\end{aligned}$$

- Gradient descent: $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k R^\top (A+A^\top) R\mathbf{x}_k$

Example: Gradient Descent in $SO(3)$

- Consider $f(R, \mathbf{x}) = \mathbf{x}^\top R^\top A R \mathbf{x}$
- Riemannian gradient with respect to R using Taylor series:

$$
\begin{aligned}
f(R\exp(\hat{\psi}), \mathbf{x}) &= \mathbf{x}^\top \left(R\exp(\hat{\psi})\right)^\top AR\exp(\hat{\psi})\mathbf{x} \\
&\approx \mathbf{x}^\top (I + \hat{\psi}^\top) R^\top AR(I + \hat{\psi})\mathbf{x} \\
&= f(R, \mathbf{x}) + \mathbf{x}^\top R^\top AR\hat{\psi}\mathbf{x} + \mathbf{x}^\top \hat{\psi}^\top R^\top AR\mathbf{x} + o(\|\psi\|_2^2) \\
&\approx f(R, \mathbf{x}) - \mathbf{x}^\top R^\top AR\hat{\mathbf{x}}\psi + (\hat{\psi}\mathbf{x})^\top R^\top AR\mathbf{x} \\
&= f(R, \mathbf{x}) - \mathbf{x}^\top R^\top AR\hat{\mathbf{x}}\psi - \psi^\top \hat{\mathbf{x}}^\top R^\top AR\mathbf{x} \\
&= f(R, \mathbf{x}) \underbrace{-\mathbf{x}^\top R^\top (A + A^\top) R\hat{\mathbf{x}}}_{(\operatorname{grad} f^\vee)^\top} \psi \\
&\Rightarrow \operatorname{grad} f(R, x) = \left(\hat{\mathbf{x}}R^\top (A + A^\top) R\mathbf{x}\right)^\wedge
\end{aligned}
$$

- Riemannian gradient descent: $R_{k+1} = R_k \exp\left(-\alpha_k \left(\hat{\mathbf{x}}R_k^\top (A + A^\top) R_k\mathbf{x}\right)^\wedge\right)$

Example: Gradient Descent in $SO(3)$

- Consider $f(R, \mathbf{x}) = \mathbf{x}^\top R^\top A R \mathbf{x}$
- Euclidean gradient with respect to R:
$$\nabla_R f(R, \mathbf{x}) = (A + A^\top) R \mathbf{x}\mathbf{x}^\top$$
- Riemannian gradient with respect to R using projection:
$$\operatorname{grad} f(R, \mathbf{x}) = 2\Pi_R(\nabla_R f(R, \mathbf{x})) = (A + A^\top) R \mathbf{x}\mathbf{x}^\top - R\mathbf{x}\mathbf{x}^\top R^\top (A + A^\top) R$$
- Using properties of the hat map:
$$(\hat{\mathbf{x}}\mathbf{y})^\wedge = \hat{\mathbf{x}}\hat{\mathbf{y}} - \hat{\mathbf{y}}\hat{\mathbf{x}} \qquad\qquad \hat{\mathbf{x}}\hat{\mathbf{y}} = \mathbf{y}\mathbf{x}^\top - \mathbf{x}^\top \mathbf{y} I$$
we can show that $\operatorname{grad} f(R, \mathbf{x})$ is consistent with our previous result:
$$\begin{aligned}\left(\hat{\mathbf{x}} R^\top (A + A^\top) R\mathbf{x}\right)^\wedge &= \hat{\mathbf{x}}\left(R^\top (A + A^\top) R\mathbf{x}\right)^\wedge - \left(R^\top (A + A^\top) R\mathbf{x}\right)^\wedge \hat{\mathbf{x}} \\ &= R^\top (A + A^\top) R\mathbf{x}\mathbf{x}^\top - \mathbf{x}\mathbf{x}^\top R^\top (A + A^\top) R \\ &= R^\top \operatorname{grad} f(R, \mathbf{x})\end{aligned}$$
- If $\dot{R} = -\operatorname{grad} f(R)$, the discrete-time rotation kinematics lead to the following update:
$$R_{k+1} = R_k \exp\left(-\alpha_k R_k^\top \operatorname{grad} f(R_k)\right)$$