AAE 590: Lie Group Methods for Control and Estimation

Differential Operators and Jacobians

Dr. James Goppert

Purdue University
School of Aeronautics and Astronautics

Lecture Overview

Differential Operators and Jacobians

- Pushforward/pullback, left/right translation
- Left and right invariant vector fields
- Left and right Jacobians (J_l, J_r) of the exponential map
- $\mathrm{SE}(2)$ Jacobian examples

The calculus of Lie groups!

Appendix contains detailed derivations and reference material for self-study.

Greek Letters and Symbols in Lie Group Theory

	LaTeX	Say it	Write by hand	Use
ξ	`\xi`	"ksee" / "zai"	squiggle, backwards 3, tail below	twist
η	`\eta`	"AY-tah"	n with tail below	error
ζ	`\zeta`	"ZAY-tah"	C, squiggle above, tail below	alt. twist
ω	`\omega`	"oh-MAY-gah"	curly w	angular vel.
θ	`\theta`	"THAY-tah"	O with horiz. bar	angle
ϕ	`\phi`	"fie" / "fee"	O + vertical line	rotation
ψ	`\psi`	"sigh" / "psee"	trident, tail below	yaw
∂	`\partial`	"partial"	backwards 6	$\partial/\partial x$
ϵ	`\epsilon`	"EP-sih-lon"	backwards 3	perturbation
δ	`\delta`	"DEL-tah"	curly d	variation

Capitals: Ω, Σ **Lie algebras:** $\mathfrak{so}$, $\mathfrak{se}$ — handwrite as "so", "se"

What is Differential Geometry?

Definition

Differential geometry is the study of smooth shapes (manifolds) using calculus.

Core concepts:

- **Manifold:** A space that *locally* looks like $\mathbb{R}^n$ (e.g., sphere, torus, Lie groups)
- **Tangent space** T_pM: The vector space of "velocities" at a point p
- **Smooth map:** A function with continuous derivatives to all orders

Why for Robotics?

Configuration spaces are often manifolds, not vector spaces:

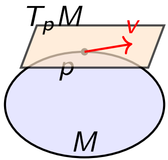
- Rotations: $\mathrm{SO}(3)$ is a 3D manifold (not $\mathbb{R}^3$!)
- Poses: $\mathrm{SE}(3)$ is a 6D manifold
- Differential geometry gives us the right tools for calculus on these spaces

Manifolds and Tangent Spaces: Quick Review

A manifold M is a space with coordinate charts $\phi_\alpha : U_\alpha \to \mathbb{R}^n$.

Tangent space T_pM at $p \in M$:

- All possible velocities $\dot{\gamma}(0)$ of curves through p
- A **vector space** of same dimension as M
- For matrix Lie groups: $T_I G = \mathfrak{g}$ (the Lie algebra)



Key Insight

Calculus on manifolds = calculus on tangent spaces (which *are* vector spaces)!

Pushforward and Pullback

Definition

Let $\phi : M \to N$ be a smooth map between manifolds. The **pushforward** (or differential) at $p \in M$ is:

$$\phi_* = d\phi_p : T_pM \to T_{\phi(p)}N$$

It maps tangent vectors at p to tangent vectors at $\phi(p)$.

Intuition: If $v \in T_pM$ is a velocity, $\phi_* v$ is the corresponding velocity after applying ϕ.

Definition

The **pullback** ϕ^* acts on covectors (differential forms) in the opposite direction:

$$\phi^* : T^*_{\phi(p)}N \to T^*_pM$$

Mnemonic: Pushforward "pushes" vectors forward; pullback "pulls" forms back.

What Are Differential Forms?

Tangent vectors $v \in T_pM$ are directions you can move at a point.
Covectors are *linear functions that eat one vector and return a number*.

Definition

A **covector** at p is a linear map $\omega_p : T_pM \to \mathbb{R}$. The space of covectors is the **cotangent space** T_p^*M.

Example: For $f : M \to \mathbb{R}$, the differential df_p is a covector:

$$df_p(v) = \text{“rate of change of } f \text{ in direction } v\text{”}$$

Vectors vs. Covectors

- **Vectors:** directions, velocities, “arrows”
- **Covectors:** measurements, “rulers” that measure vectors

In coordinates: vectors are column vectors, covectors are row vectors.

Differential Forms: The Basics

A 1-form ω is a smooth assignment of a covector to each point: $p \mapsto \omega_p \in T_p^*M$.
The "1" means it eats **one vector**. (2-forms eat two vectors, etc.)

In coordinates on $\mathbb{R}^n$: $\omega = a_1(x)\,dx_1 + a_2(x)\,dx_2 + \cdots + a_n(x)\,dx_n$

where dx_i is the "dual basis" to $\frac{\partial}{\partial x_i}$: $dx_i\left(\frac{\partial}{\partial x_j}\right) = \delta_{ij}$

Example: On $\mathbb{R}^2$, $\omega = x\,dy - y\,dx$ is a 1-form. Applied to $v = (2,3)$ at $(1,1)$:

$$\omega_{(1,1)}(v) = 1 \cdot 3 - 1 \cdot 2 = 1$$

Why Forms Matter for Lie Groups

- Invariant 1-forms pair with invariant vector fields: $\theta^i(X_j) = \delta_{ij}$
- $\mathfrak{g}$-valued forms: return Lie algebra elements, not scalars
- Integration of forms gives path-independent quantities

The Maurer-Cartan Form

Question: Given a tangent vector $v \in T_g G$, how do we map it back to the Lie algebra?

Answer: Use left translation! $(L_{g^{-1}})_* v = g^{-1} v \in T_e G = \mathfrak{g}$

Definition

The **Maurer-Cartan form** ω is a $\mathfrak{g}$-valued 1-form on G:

$$\omega_g : T_g G \to \mathfrak{g}, \quad \omega_g(v) = g^{-1} v$$

For matrix groups: $\omega = g^{-1}\, dg$ (matrix multiplication, returns a matrix in $\mathfrak{g}$).

Note: Unlike covectors ($T_p M \to \mathbb{R}$), this returns a Lie algebra element.
Key property: ω is **left-invariant**: $L_h^* \omega = \omega$ for all $h \in G$.

Physical Interpretation

If $g(t)$ is a curve in G, then $\omega(\dot{g}) = g^{-1}\dot{g} = \xi$ is the **body velocity**—the Lie algebra element generating the motion!

Maurer-Cartan: Why It Matters

The Maurer-Cartan form encodes the group structure:

Theorem (Maurer-Cartan Equation)

The Maurer-Cartan form $\omega = g^{-1}\,dg$ *satisfies:*

$$d\omega + \omega \wedge \omega = 0$$

or equivalently: $d\omega + \frac{1}{2}[\omega, \omega] = 0$

This equation:

- Encodes the Lie bracket structure
- Is the integrability condition for reconstructing g from ω
- Appears in gauge theory and differential geometry

Connection to Kinematics

For $\dot{g} = g\xi^\wedge$, the body velocity is $\omega(\dot{g}) = g^{-1}\dot{g} = \xi^\wedge$.
The Maurer-Cartan form extracts "what the system is doing" at any point!

Differential Operators: Notation by Author

Different authors use different notations! Here's a Rosetta stone:

Concept	**Math/Physics** [1, 2]	**Robotics** [3, 4]	**This Course**
Differential/pushforward	$d\phi$, ϕ_*, $T\phi$	$D\phi$, $\mathbf{J}$	df, f_*
At a point p	$d\phi_p$, $(T\phi)_p$	$D\phi(p)$	df_p
Applied to vector v	$d\phi_p(v)$, $\phi_* v$	$D\phi(p) \cdot v$	$df_p(v)$
Directional derivative	$v(f)$, $\nabla_v f$	$D_v f$	$D_v f$
Lie derivative	$\mathcal{L}_X f$	$L_X f$	$\mathcal{L}_X$
Wedge/hat map	$\widehat{\xi}$, $\xi^\wedge$	$\xi^\wedge$, $[\xi]_\times$	$\xi^\wedge$
Exp map Jacobians	—	$\mathbf{J}_l$, $\mathbf{J}_r$	J_l, J_r

Key References

Barfoot [3]: Bold **J**, detailed Jacobian derivations for SLAM.
Solà et al. [4]: Concise "micro Lie theory," our primary reference for J_l, J_r.
Hall [1]: Rigorous math; uses $d\phi$ for differential.
Bullo & Lewis [2]: $T\phi$ for tangent map; geometric control focus.

The Differential df: Coordinate-Free Definition

Definition

For $f : M \to N$ smooth, the **differential** at $p \in M$ is the linear map $df_p : T_pM \to T_{f(p)}N$ defined by: for any curve $\gamma(t)$ with $\gamma(0) = p$, $\dot{\gamma}(0) = v$:

$$df_p(v) = \frac{d}{dt}\bigg|_{t=0} f(\gamma(t))$$

Key: df_p is **independent of which curve** γ we choose—only $v = \dot{\gamma}(0)$ matters.

Why "Coordinate-Free"?

This definition doesn't mention coordinates. df_p exists intrinsically; coordinates just help *compute* it.

The Differential in Coordinates: The Jacobian

In coordinates, the differential becomes a matrix (the Jacobian):

If $f : \mathbb{R}^m \to \mathbb{R}^n$ with $f(x) = (f_1(x), \ldots, f_n(x))$, then:

$$Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

Application to tangent vector:

$$df_p(v) = Df(p) \cdot v \quad \text{(matrix-vector product)}$$

Notation Clash: df vs Df

- df: the abstract linear map (coordinate-free)
- Df: the Jacobian matrix (coordinate-dependent)

They represent the same object in different languages!

Directional Derivatives: Intrinsic to the Manifold

Definition

The **directional derivative** of $f : M \to \mathbb{R}$ along $v \in T_pM$ is:

$$D_v f = df_p(v) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t))$$

where $\gamma(t)$ is any curve with $\gamma(0) = p$, $\dot{\gamma}(0) = v$.

In coordinates: $D_v f = \nabla f \cdot v = \sum_i v_i \frac{\partial f}{\partial x_i}$
Key insight: Intrinsic—doesn't depend on coordinates, only on direction v.

Partial $\leftrightarrow$ Directional

Partial derivatives are directional derivatives along coordinate basis vectors: $\frac{\partial f}{\partial x_i} = D_{e_i} f$

The Lie Derivative: Derivative Along a Flow

Definition

The **Lie derivative** of $f : M \to \mathbb{R}$ along vector field X is:

$$\mathcal{L}_X f = \lim_{t \to 0} \frac{f(\phi_t(p)) - f(p)}{t} = \left.\frac{d}{dt}\right|_{t=0} f(\phi_t(p))$$

Key insight: $\mathcal{L}_X f = X(f) = D_X f = df(X)$ — for functions, it's the directional derivative!

But the Lie derivative is more general: it can differentiate *any* tensor field (vectors, forms, etc.) along a flow.

For Vector Fields

$$\mathcal{L}_X Y = \lim_{t \to 0} \frac{(\phi_{-t})_* Y_{\phi_t(p)} - Y_p}{t}$$

This measures how Y changes as we flow along X. Result: $\mathcal{L}_X Y = [X, Y]$ (Lie bracket)!

Lie Derivative of Vector Fields = Lie Bracket

Theorem

For vector fields X, Y on M: $\mathcal{L}_X Y = [X, Y] = XY - YX$ where $(XY)f = X(Y(f))$ means "apply Y then X to function f".

Geometric interpretation:

- Flow along X then Y for time ϵ each
- Flow along Y then X for time ϵ each
- Difference (to order ϵ^2) is flow along $[X, Y]$

$$\phi_\epsilon^X \circ \phi_\epsilon^Y \circ \phi_{-\epsilon}^X \circ \phi_{-\epsilon}^Y \approx \phi_{\epsilon^2}^{[X,Y]}$$

Non-Commutativity

$[X, Y] \neq 0$ means flows don't commute—why Lie groups have non-trivial structure!

Example: Lie Bracket on $\mathrm{SE}(2)$

Consider two vector fields on $\mathrm{SE}(2)$:

- X_v: flow that translates forward (pure forward velocity)
- X_ω: flow that rotates (pure angular velocity)

In Lie algebra terms: $\xi_v = (1, 0, 0)^T$ (translate forward), $\xi_\omega = (0, 0, 1)^T$ (rotate)
Lie bracket:

$$[\xi_v, \xi_\omega] = \begin{bmatrix} 0 \cdot 0 - 1 \cdot 0 \\ 1 \cdot 1 - 0 \cdot 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Physical meaning: Translate-then-rotate vs. rotate-then-translate differs by a **lateral translation**! The bracket captures this non-commutativity.

Intuition

Rotating changes which direction "forward" means, so order matters.

Hierarchy of Derivatives: Summary

$$\boxed{\text{Partial } \frac{\partial f}{\partial x_i}} \xrightarrow{v=e_i} \boxed{\text{Directional } D_v f} \xrightarrow{\text{all } v} \boxed{\text{Differential } df} \xrightarrow{\text{along flow}} \boxed{\text{Lie } \mathcal{L}_X}$$

- **Partial:** one coordinate direction, requires chart
- **Directional:** any tangent vector at a point
- **Differential:** packages all directional derivatives into a linear map
- **Lie:** derivative along a vector field (defined globally), works on tensors

For Lie Groups

Left/right-invariant vector fields give **global** flows, making Lie derivatives particularly powerful for studying group structure.

Worked Example: Pushforward on $\mathrm{SO}(2)$

For $L_R : \mathrm{SO}(2) \to \mathrm{SO}(2)$*,* $L_R(Q) = RQ$*, compute* $(L_R)_* \omega^\wedge$ *at identity.*

Solution: Pushforward on $\mathrm{SO}(2)$

At $Q = I$: Consider curve $\gamma(t) = \mathrm{Exp}(t\omega^\wedge)$ with $\gamma(0) = I$, $\dot{\gamma}(0) = \omega^\wedge$

$(L_R)_*\omega^\wedge = \frac{d}{dt}\big|_{t=0} L_R(\gamma(t)) = \frac{d}{dt}\big|_{t=0} R \cdot \mathrm{Exp}(t\omega^\wedge)$

$= R \cdot \frac{d}{dt}\big|_{t=0} \mathrm{Exp}(t\omega^\wedge) = R \cdot \omega^\wedge$

Result: $(L_R)_*\omega^\wedge = R\omega^\wedge \in T_R\mathrm{SO}(2)$

This is a **left-invariant vector field**: $\omega^\wedge$ at I maps to $R\omega^\wedge$ at R.

Worked Example: Pullback on $\mathrm{SO}(2)$

For $L_R : \mathrm{SO}(2) \to \mathrm{SO}(2)$, *define 1-form* $\alpha_R(V) = (R^T V)_{21}$ *at* R *(extracts angular velocity). Compute* $(L_R)^* \alpha_R$ *at identity.*

Solution: Pullback on $\mathrm{SO}(2)$

Pullback definition: $((L_R)^*\alpha_R)(v) = \alpha_R((L_R)_*v)$ for $v \in T_I\mathrm{SO}(2)$

Let $v = \omega^\wedge \in T_I\mathrm{SO}(2)$. From pushforward: $(L_R)_*\omega^\wedge = R\omega^\wedge \in T_R\mathrm{SO}(2)$

$((L_R)^*\alpha_R)(\omega^\wedge) = \alpha_R(R\omega^\wedge) = (R^T R\omega^\wedge)_{21} = (\omega^\wedge)_{21} = \omega$

Result: $(L_R)^*\alpha_R = \alpha_I$ — the pullback of α at R equals α at I!

Key insight: This 1-form is **left-invariant**. Pullback goes "backwards," pulling forms from target to source.

Differential Operators: A Reference Guide

Notation	Name	Meaning
$\frac{\partial f}{\partial x}$	Partial derivative	Derivative holding other variables fixed
$\frac{df}{dt}$ or $\dot{f}$	Total/time derivative	$\frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial x_i}\dot{x}_i$
Df or df	Differential	Linear map $T_pM \to T_{f(p)}N$ (same as pushforward)
ϕ_* or $d\phi$	Pushforward	Maps tangent vectors: $T_pM \to T_{\phi(p)}N$
ϕ^*	Pullback	Maps covectors/forms: $T^*_{\phi(p)}N \to T^*_pM$
D_vf	Directional deriv.	$\lim_{t\to 0} \frac{f(p+tv)-f(p)}{t} = \nabla f \cdot v$
$\mathcal{L}_Xf$	Lie derivative	Rate of change of f along flow of vector field X
$[X, Y]$	Lie bracket	$\mathcal{L}_XY$; measures non-commutativity of flows

Key relationships:

- For functions: $\mathcal{L}_Xf = X(f) = D_Xf = df(X)$
- For vector fields: $\mathcal{L}_XY = [X, Y] = XY - YX$
- Pushforward of composition: $(f \circ g)_* = f_* \circ g_*$

Differential Operators: Matrix Lie Groups

For matrix Lie groups, operators simplify to matrix operations:

$$
\begin{aligned}
\text{Derivative:} \quad & \frac{dF}{dt} = \lim_{h \to 0} \frac{F(t+h) - F(t)}{h} \\
\text{Pushforward:} \quad & (L_g)_* \xi = g\xi \quad \text{(matrix multiplication)} \\
\text{Lie bracket:} \quad & [A, B] = AB - BA \quad \text{(commutator)} \\
\text{Adjoint:} \quad & \mathrm{Ad}_g(\xi) = g\xi g^{-1}, \quad \mathrm{ad}_\xi(\eta) = [\xi, \eta]
\end{aligned}
$$

Good News

For matrix groups, differential geometry reduces to matrix calculus!

Left and Right Translation on Lie Groups

Definition

For $g \in G$, define **left/right translation** (diffeomorphisms $G \to G$):

$$L_g(h) = gh, \qquad R_g(h) = hg$$

Their pushforwards move tangent vectors around the group:

$$(L_g)_* : T_h G \to T_{gh} G, \quad (R_g)_* : T_h G \to T_{hg} G$$

Key Insight

Left/right translation transports vectors from identity I to any $g \in G$. This connects the Lie algebra $\mathfrak{g} = T_I G$ to the entire group!

Left-Invariant Vector Fields [1]

Definition

A vector field V on G is **left-invariant** if $(L_g)_* V_h = V_{gh}$ for all $g, h \in G$.

Construction: Given $\xi \in \mathfrak{g} = T_I G$: $V_g = (L_g)_* \xi = g\xi^\wedge$
This gives a left-invariant vector field V with $V_I = \xi$.

One-to-One Correspondence

Left-invariant vector fields on $G \longleftrightarrow$ elements of $\mathfrak{g}$

Physical meaning: Left-invariant = **body-frame** velocities—the same "body velocity" ξ at every configuration.

Right-Invariant Vector Fields

Definition

A vector field V on G is **right-invariant** if $(R_g)_* V_h = V_{hg}$ for all $g, h \in G$.

Construction: Given $\xi \in \mathfrak{g}$: $V_g = (R_g)_* \xi = \xi^\wedge g$
Physical meaning: Right-invariant = **world-frame** velocities.

Left vs Right

- **Left-invariant:** $\dot{X} = X\xi^\wedge$ (body velocity)
- **Right-invariant:** $\dot{X} = \xi^\wedge X$ (world velocity)

Adjoint Connects Left and Right

How are left and right related?

The **Adjoint map** $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ converts between them:

$$(L_g)_* \xi = (R_g)_* (\mathrm{Ad}_g \xi)$$

Equivalently:

$$g\xi^\wedge = (\mathrm{Ad}_g \xi)^\wedge g$$

Interpretation: Ad_g transforms a body-frame velocity to the equivalent world-frame velocity (and vice versa).

Example

For $\mathrm{SE}(2)$: if a robot has body velocity $\xi = (v, 0, \omega)^T$, its world velocity is $\mathrm{Ad}_X \xi$, which depends on the robot's current pose X.

Left and Right Jacobians: Motivation [4, 3]

Question: How does $\mathrm{Exp}((\xi + \delta\xi)^\wedge)$ relate to $\mathrm{Exp}(\xi^\wedge)$?

Naively, we might hope:

$$\mathrm{Exp}((\xi + \delta\xi)^\wedge) \stackrel{?}{=} \mathrm{Exp}(\xi^\wedge)\mathrm{Exp}(\delta\xi^\wedge)$$

But this is only true for abelian groups!

For general Lie groups:

$$\begin{aligned} \mathrm{Exp}((\xi + \delta\xi)^\wedge) &\approx \mathrm{Exp}(\xi^\wedge)\mathrm{Exp}((J_r^{-1}(\xi)\delta\xi)^\wedge) \\ &\approx \mathrm{Exp}((J_l^{-1}(\xi)\delta\xi)^\wedge)\mathrm{Exp}(\xi^\wedge) \end{aligned}$$

where J_r and J_l are the **right** and **left Jacobians**.

Why This Matters

Jacobians appear in optimization, uncertainty propagation, and control on Lie groups.

Jacobians and the Derivative of Exp

The derivative of the exponential map is NOT simply $\mathrm{Exp}(\xi^\wedge)$!

Theorem (Derivative of Exponential Map)

For a curve $\xi(t)$ in the Lie algebra:

$$\frac{d}{dt}\mathrm{Exp}(\xi(t)^\wedge) = \mathrm{Exp}(\xi^\wedge) \cdot (J_l(\xi)\dot{\xi})^\wedge = (J_r(\xi)\dot{\xi})^\wedge \cdot \mathrm{Exp}(\xi^\wedge)$$

Equivalently, for perturbations:

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Exp}((\xi + \epsilon\delta\xi)^\wedge) = \mathrm{Exp}(\xi^\wedge) \cdot (J_l(\xi)\delta\xi)^\wedge = (J_r(\xi)\delta\xi)^\wedge \cdot \mathrm{Exp}(\xi^\wedge)$$

Key Insight

The Jacobian "corrects" for the non-commutativity of the group. It maps algebra perturbations to group perturbations.

General Jacobian Formulas [4, 5]

The Jacobians have infinite series expansions in terms of ad_ξ:

Definition (Left Jacobian)

$$J_l(\xi) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}_\xi^k = I + \frac{1}{2!}\mathrm{ad}_\xi + \frac{1}{3!}\mathrm{ad}_\xi^2 + \frac{1}{4!}\mathrm{ad}_\xi^3 + \cdots$$

Definition (Right Jacobian)

$$J_r(\xi) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!}\mathrm{ad}_\xi^k = I - \frac{1}{2!}\mathrm{ad}_\xi + \frac{1}{3!}\mathrm{ad}_\xi^2 - \frac{1}{4!}\mathrm{ad}_\xi^3 + \cdots$$

Note: ad_ξ is the matrix such that $\mathrm{ad}_\xi \eta = [\xi, \eta]$ (Lie bracket as matrix multiplication).

Closed-Form Jacobian Expressions

Closed form using the matrix exponential:

$$J_l(\xi) = \int_0^1 e^{t\,\mathrm{ad}_\xi}\,dt = \frac{e^{\mathrm{ad}_\xi} - I}{\mathrm{ad}_\xi}, \qquad J_r(\xi) = \int_0^1 e^{-t\,\mathrm{ad}_\xi}\,dt = \frac{I - e^{-\mathrm{ad}_\xi}}{\mathrm{ad}_\xi}$$

Interpretation: The Jacobian is the "average" of $e^{t\,\mathrm{ad}_\xi}$ over $t \in [0, 1]$.

Notation Warning

The fraction $\frac{A}{B}$ means AB^{-1} and **assumes** A **and** B **commute**. This holds because e^{ad_ξ} and ad_ξ commute. When ad_ξ is singular, use L'Hôpital-style limits.

Jacobian Relationships

Key identities connecting left and right Jacobians:

1. Via Adjoint:

$$J_l(\xi) = \mathrm{Ad}_{\mathrm{Exp}(\xi^\wedge)} J_r(\xi)$$

2. Via negation:

$$J_r(\xi) = J_l(-\xi), \quad J_l(\xi) = J_r(-\xi)$$

3. Inverse formulas:

$$J_l^{-1}(\xi) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \mathrm{ad}_\xi^k, \quad J_r^{-1}(\xi) = \sum_{k=0}^{\infty} \frac{B_k}{k!} (-\mathrm{ad}_\xi)^k$$

where B_k are **Bernoulli numbers**: $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $\ldots$
The series form is standard. Our convention: $\mathrm{ad}_\xi(\eta) = [\xi, \eta]$ *with* $[A, B] = AB - BA$.

SE(2) Example: The Adjoint Matrix ad_ξ

To compute Jacobians for $\mathrm{SE}(2)$, we need the adjoint matrix ad_ξ:

For $\xi = (v_x, v_y, \omega)^T$ (translation-first ordering):

$$\mathrm{ad}_\xi = \begin{bmatrix} 0 & \omega & v_y \\ -\omega & 0 & -v_x \\ 0 & 0 & 0 \end{bmatrix}$$

Verification: Check that $\mathrm{ad}_\xi \eta = [\xi, \eta]$ for any $\eta = (u_x, u_y, \psi)^T$:

$$[\xi, \eta] = \begin{bmatrix} \omega u_y - \psi v_y \\ -\omega u_x + \psi v_x \\ 0 \end{bmatrix}$$

Block Upper Triangular

With translation-first ordering, ad_ξ has the same block structure as Ad_X and the group element X.

Left and Right Jacobians for $\mathrm{SE}(2)$ [4]

Unlike abelian groups, the Jacobians for $\mathrm{SE}(2)$ are non-trivial!

For $\xi = (v_x, v_y, \omega)^T \in \mathbb{R}^3$, the **right Jacobian** is:

$$J_r(\xi) = \begin{bmatrix} \frac{\sin\omega}{\omega} & -\frac{1-\cos\omega}{\omega} & a(\omega)v_y \\ \frac{1-\cos\omega}{\omega} & \frac{\sin\omega}{\omega} & -a(\omega)v_x \\ 0 & 0 & 1 \end{bmatrix}$$

where $a(\omega) = \frac{\omega - \sin\omega}{\omega^2}$ (with $a(0) = 0$).

Block upper triangular structure:

- Top-left 2×2: rotation-like block (the $V(\omega)$ matrix)
- Bottom-right: angular part is trivial (1D rotation)
- Top-right: rotation-translation coupling

Left Jacobian for $\mathrm{SE}(2)$

The **left Jacobian** is related by:

$$J_l(\xi) = \mathrm{Ad}_{\mathrm{Exp}(\xi^\wedge)} J_r(\xi) = J_r(-\xi)$$

Explicit form:

$$J_l(\xi) = \begin{bmatrix} \frac{\sin\omega}{\omega} & \frac{1-\cos\omega}{\omega} & -a(\omega)v_y \\ -\frac{1-\cos\omega}{\omega} & \frac{\sin\omega}{\omega} & a(\omega)v_x \\ 0 & 0 & 1 \end{bmatrix}$$

Key Difference from Abelian Groups

The coupling terms (third column, rows 1-2) depend on **both** ω and $\boldsymbol{v}$. This is the semi-direct product structure manifesting in the Jacobian!

Convention Warning: Parameterization Order

Critical Point

The **ordering** of Lie algebra components is an **author's choice**!

Our convention: $\xi = (v_x, v_y, \omega)^T$ (translation first)

Alternative convention: $\xi = (\omega, v_x, v_y)^T$ (angular first)

With angular-first ordering, the Jacobian becomes block lower triangular:

$$J_r(\xi) = \begin{bmatrix} 1 & 0 & 0 \\ a(\omega)v_y & \frac{\sin\omega}{\omega} & -\frac{1-\cos\omega}{\omega} \\ -a(\omega)v_x & \frac{1-\cos\omega}{\omega} & \frac{\sin\omega}{\omega} \end{bmatrix}$$

Same mathematical content, different matrix layout!

When Reading Papers

Always check the author's convention before using their formulas. The Jacobian structure changes with ordering—this is **not** a typo!

Jacobian Inverse and Small Angle Limits

Small angle approximation (as $\omega \to 0$):

$$J_r(\xi) \approx I - \frac{1}{2}\mathrm{ad}_\xi + O(\|\xi\|^2)$$

$$J_l(\xi) \approx I + \frac{1}{2}\mathrm{ad}_\xi + O(\|\xi\|^2)$$

where ad_ξ is the adjoint matrix representation of ξ.

At $\xi = 0$: $J_l(0) = J_r(0) = I$

Inverse Jacobians (needed for Log derivatives):

$$J_r^{-1}(\xi) = I + \frac{1}{2}\mathrm{ad}_\xi + O(\|\xi\|^2)$$

Numerical Stability

Use Taylor series near $\omega = 0$ to avoid division by zero!

Why Jacobians Matter: Applications

1. **Time derivative of Lie algebra coordinates:** For $z = \mathrm{Log}(X)^\vee$:

$$\dot{z} = J_l^{-1}(z) \cdot \xi \quad \text{where } \dot{X} = X\xi^\wedge$$

Converts body velocity ξ to rate of change in coordinates $\dot{z}$.

2. **Uncertainty propagation:**

$$\Sigma_{\mathrm{Exp}(\xi^\wedge)} \approx J_l(\xi)\Sigma_\xi J_l(\xi)^T$$

3. **Optimization on manifolds:** For $\xi = \mathrm{Log}(X)^\vee$, perturb $X \to \mathrm{Exp}(\delta^\wedge)X$:

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathrm{Log}(\mathrm{Exp}(\epsilon\delta^\wedge)X)^\vee = J_l^{-1}(\xi)\delta$$

4. **Integration on Lie groups:** For $z(t) = \mathrm{Log}(X(t))^\vee$ and $\dot{X} = X\xi^\wedge$:

$$\dot{z}(t) = J_l^{-1}(z(t)) \cdot \xi(t)$$

Worked Example: $\mathrm{SE}(2)$ Jacobians

For $\xi = (2, 1, 0)^T$ *(pure translation): compute* ad_ξ *and* $J_l(\xi)$.

Solution: $\mathrm{SE}(2)$ Jacobians

$$\mathrm{ad}_\xi = \begin{bmatrix} \omega^\wedge & \boldsymbol{v}^\odot \\ \mathbf{0}^T & 0 \end{bmatrix} \text{ where } \omega = 0,\ \boldsymbol{v}^\odot = (v_y, -v_x)^T = (1, -2)^T$$

$$\mathrm{ad}_\xi = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(nilpotent: } \mathrm{ad}_\xi^2 = 0\text{)}$$

$$J_l = I + \tfrac{1}{2}\mathrm{ad}_\xi + \tfrac{1}{6}\mathrm{ad}_\xi^2 + \cdots = I + \tfrac{1}{2}\mathrm{ad}_\xi = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

For pure translation ($\omega = 0$): J_l is upper triangular with 1s on diagonal. $J_l^{-1} = \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Appendix Overview

The following slides contain additional reference material for self-study:

- **A: Derivative Theory** — Types of derivatives on manifolds (gradient vs differential, partial vs total, Lie derivative)
- **B: Jacobian Proofs** — Derivation of $\frac{d}{dt}\mathrm{Exp}(\xi^\wedge)$, beta function $\beta(\theta)$, series expansions, and closed-form Jacobians for rotation groups

These slides are not covered in lecture but are helpful for homework and deeper understanding.

Gradient vs. Differential: A Common Confusion

For $f : \mathbb{R}^n \to \mathbb{R}$, what's the difference between ∇f and df?

Gradient $\nabla f \in \mathbb{R}^n$: a **vector** (lives in the same space as x)

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Differential df: a **covector/1-form** (lives in the dual space T_p^*M)

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

They're related by the metric: $\nabla f = g^{-1} df$ where g is the metric tensor.

In Euclidean Space with Standard Metric

$g = I$, so ∇f and df have the same components. But on curved manifolds, they differ!

Partial Derivatives: Coordinate-Dependent

Definition

The **partial derivative** of $f : \mathbb{R}^n \to \mathbb{R}$ with respect to x_i is:

$$\frac{\partial f}{\partial x_i}(p) = \lim_{h \to 0} \frac{f(p + he_i) - f(p)}{h}$$

where e_i is the i-th standard basis vector.

Key properties:

- Requires a **coordinate system** (the basis $\{e_i\}$)
- Holds other variables fixed
- Is a **scalar** (not a vector or linear map)

On Manifolds

Partial derivatives depend on the choice of chart (local coordinates). They are **not** intrinsic to the manifold—different coordinates give different partials for the same geometric object.

Total Derivative vs. Partial Derivative

Consider $f(x(t), y(t), t)$ — how does f change with t?

Total derivative (chain rule):

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial t}$$

Partial derivative:

$$\frac{\partial f}{\partial t} \quad \text{(holding } x, y \text{ fixed)}$$

The difference:

- $\frac{\partial f}{\partial t}$: explicit dependence on t only
- $\frac{df}{dt}$: total change including implicit dependence through $x(t), y(t)$

In Mechanics

For $L(q, \dot{q}, t)$: $\frac{\partial L}{\partial t}$ is explicit time-dependence; $\frac{dL}{dt}$ includes motion along trajectory.

The Flow of a Vector Field

Before defining the Lie derivative, we need flows:

Definition

Given vector field X on M, its **flow** is the map $\phi_t : M \to M$ satisfying:

$$\frac{d}{dt}\phi_t(p) = X(\phi_t(p)), \quad \phi_0(p) = p$$

Intuition: $\phi_t(p)$ is where you end up after flowing along X for time t starting at p.

Properties:

- $\phi_0 = \mathrm{id}$ (identity map)
- $\phi_s \circ \phi_t = \phi_{s+t}$ (group property)
- $\phi_t^{-1} = \phi_{-t}$ (invertible)

Example

On $\mathbb{R}^2$, $X = \frac{\partial}{\partial x}$ (constant rightward field): $\phi_t(x, y) = (x + t, y)$

Understanding the Different Derivatives

Why so many "derivatives"? Each answers a different question:

1. **Partial derivative** $\frac{\partial f}{\partial x}$: How does f change when I vary *one coordinate*?
   Requires a coordinate system; other variables held fixed
2. **Directional derivative** $D_v f$: How does f change along direction v *at a point*?
   v is a single tangent vector; gives a number
3. **Differential/Pushforward** $df = f_*$: The *linear map* sending all tangent vectors to their images
   $df_p : T_pM \to T_{f(p)}N$; packages all directional derivatives
4. **Lie derivative** $\mathcal{L}_X f$: How does f change along the *flow* of vector field X?
   X is defined everywhere (not just at one point); follows integral curves
5. **Total derivative** $\frac{df}{dt}$: How does f change along a *specific curve* $\gamma(t)$?
   Chain rule: $\frac{df}{dt} = df(\dot{\gamma}) = D_{\dot{\gamma}} f$

When to Use Which Derivative

Situation	**Use**	**Result**
Scalar function, one variable	$\frac{\partial f}{\partial x_i}$	Scalar
Scalar function, specific direction	$D_v f = \nabla f \cdot v$	Scalar
Map between manifolds	$df = f_*$ (pushforward)	Linear map
Function along a flow	$\mathcal{L}_X f$	Scalar field
Vector field along a flow	$\mathcal{L}_X Y = [X, Y]$	Vector field
Along a parameterized curve	$\frac{d}{dt} f(\gamma(t))$	Scalar

Key Insight for Lie Groups

On Lie groups, we often have *natural* vector fields (left/right-invariant), so the **Lie derivative** and **Lie bracket** become essential tools—not just the point-wise directional derivative.

Proof: Derivative of the Exponential Map (1/3)

Goal: Compute $\frac{d}{dt}\mathrm{Exp}(\xi(t)^\wedge)$ where $\xi(t)$ is a curve in the Lie algebra.

Step 1: Start with the matrix exponential series

$$\mathrm{Exp}(\xi^\wedge) = e^{\xi^\wedge} = \sum_{k=0}^{\infty} \frac{(\xi^\wedge)^k}{k!} = I + \xi^\wedge + \frac{(\xi^\wedge)^2}{2!} + \frac{(\xi^\wedge)^3}{3!} + \cdots$$

Step 2: Differentiate term by term

For the k-th term, we need $\frac{d}{dt}(\xi^\wedge)^k$. But $\xi^\wedge$ and $\dot{\xi}^\wedge$ don't commute in general!

Product rule for non-commuting matrices:

$$\frac{d}{dt}(\xi^\wedge)^2 = \dot{\xi}^\wedge \xi^\wedge + \xi^\wedge \dot{\xi}^\wedge \neq 2\xi^\wedge \dot{\xi}^\wedge$$

This is where the complexity comes from!

Proof: Derivative of the Exponential Map (2/3)

Step 3: General term differentiation

For the k-th power:

$$\frac{d}{dt}(\xi^\wedge)^k = \sum_{j=0}^{k-1} (\xi^\wedge)^j \dot{\xi}^\wedge (\xi^\wedge)^{k-1-j}$$

Step 4: Collect terms in the series

$$\frac{d}{dt} e^{\xi^\wedge} = \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{j=0}^{k-1} (\xi^\wedge)^j \dot{\xi}^\wedge (\xi^\wedge)^{k-1-j}$$

Key observation: We can factor this as:

$$\frac{d}{dt} e^{\xi^\wedge} = e^{\xi^\wedge} \cdot \left(\text{something involving } \dot{\xi}^\wedge\right)$$

The "something" will be $(J_l(\xi)\dot{\xi})^\wedge$!

Proof: Derivative of the Exponential Map (3/3)

Step 5: Use the adjoint action identity

Key fact: $e^{\xi^\wedge}\eta^\wedge e^{-\xi^\wedge} = (\mathrm{Ad}_{e^{\xi^\wedge}}\eta)^\wedge = (e^{\mathrm{ad}_\xi}\eta)^\wedge$

After careful manipulation (see Chirikjian [5]), we get:

$$\frac{d}{dt}e^{\xi^\wedge} = e^{\xi^\wedge} \cdot \left(\int_0^1 e^{-s\,\mathrm{ad}_\xi}\,ds \cdot \dot{\xi}\right)^\wedge$$

Step 6: Identify the Jacobian

The integral is exactly $J_r(\xi)$:

$$J_r(\xi) = \int_0^1 e^{-s\,\mathrm{ad}_\xi}\,ds$$

But we want left multiplication, so use $J_l(\xi) = e^{\mathrm{ad}_\xi}J_r(\xi)$:

$$\boxed{\frac{d}{dt}\mathrm{Exp}(\xi^\wedge) = \mathrm{Exp}(\xi^\wedge) \cdot (J_l(\xi)\dot{\xi})^\wedge}$$

Why Does This Formula Make Sense?

$$\frac{d}{dt}\mathrm{Exp}(\xi(t)^\wedge) = \mathrm{Exp}(\xi^\wedge) \cdot (J_l(\xi)\dot{\xi})^\wedge$$

Sanity check 1: Abelian case — If G is abelian, $\mathrm{ad}_\xi = 0$, so $J_l(\xi) = I$:

$$\frac{d}{dt}e^{\xi^\wedge} = e^{\xi^\wedge} \cdot \dot{\xi}^\wedge \quad \text{(just like } \tfrac{d}{dt}e^{at} = a \cdot e^{at}\text{)}$$

Sanity check 2: At $\xi = 0$ — $J_l(0) = I$, so:

$$\left.\frac{d}{dt}\right|_{\xi=0} \mathrm{Exp}(\xi^\wedge) = \dot{\xi}^\wedge$$

The derivative at identity is exactly the Lie algebra element.
Sanity check 3: Dimensions match — LHS is tangent at $\mathrm{Exp}(\xi^\wedge) \in G$. RHS: left-translates a Lie algebra element to that point. ✓

Jacobian Series and Beta Function

Series expansion: From $J_l = \int_0^1 e^{s\,\mathrm{ad}_\xi}\,ds$:

$$J_l(\xi) = \sum_{k=0}^{\infty} \frac{\mathrm{ad}_\xi^k}{(k+1)!} = I + \frac{1}{2}\mathrm{ad}_\xi + \frac{1}{6}\mathrm{ad}_\xi^2 + \frac{1}{24}\mathrm{ad}_\xi^3 + \cdots$$

For rotation groups, define scalar functions (with $\theta = \|\boldsymbol{\omega}\|$):

$$\alpha(\theta) = \frac{\sin\theta}{\theta} = 1 - \frac{\theta^2}{6} + \frac{\theta^4}{120} - \cdots$$
$$\beta(\theta) = \frac{1-\cos\theta}{\theta^2} = \frac{1}{2} - \frac{\theta^2}{24} + \frac{\theta^4}{720} - \cdots$$
$$\gamma(\theta) = \frac{1-\alpha}{\theta^2} = \frac{1}{\theta^2}\left(1 - \frac{\sin\theta}{\theta}\right)$$

Closed forms for $\mathrm{SO}(3)$:

$$J_l(\boldsymbol{\omega}) = \alpha(\theta)I + \beta(\theta)\boldsymbol{\omega}^\wedge + \gamma(\theta)\boldsymbol{\omega}\boldsymbol{\omega}^T$$

Jacobian Inverse and Taylor Series

Inverse from series: Using $(I + A)^{-1} = I - A + A^2 - A^3 + \cdots$ for small A:

$$J_l^{-1}(\xi) = I - \frac{1}{2}\mathrm{ad}_\xi + \frac{1}{12}\mathrm{ad}_\xi^2 - \frac{1}{720}\mathrm{ad}_\xi^4 + \cdots$$

Note: The ad_ξ^3 term vanishes due to the Bernoulli numbers B_k pattern.

For $\mathrm{SO}(3)$: Define $\beta'(\theta) = \frac{\theta - \sin\theta}{\theta^3}$:

$$J_l^{-1}(\boldsymbol{\omega}) = I - \frac{1}{2}\boldsymbol{\omega}^\wedge + \beta'(\theta)(\boldsymbol{\omega}^\wedge)^2$$

Small angle approximation (for numerical stability):

$$\begin{aligned} \theta \to 0: \quad J_l &\approx I + \frac{1}{2}\mathrm{ad}_\xi \\ J_l^{-1} &\approx I - \frac{1}{2}\mathrm{ad}_\xi \end{aligned}$$

References I

[1] Brian C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2nd. Springer, 2015.

[2] Francesco Bullo and Andrew D Lewis. *Geometric Control of Mechanical Systems*. Springer, 2005.

[3] Timothy D Barfoot. *State Estimation for Robotics*. Cambridge University Press, 2017.

[4] Joan Solà, Jeremie Deray, and Dinesh Atchuthan. "A micro Lie theory for state estimation in robotics". In: *arXiv preprint arXiv:1812.01537* (2018).

[5] Gregory S Chirikjian. *Stochastic Models, Information Theory, and Lie Groups*. Vol. 2. Birkhäuser, 2011.