

THREAD – NWT6 – Part II: Lie group integrators

Brynjulf Owren

Department of Mathematical Sciences, NTNU, Trondheim, Norway

February 5, 8, 9 – 2021

- We aim at explaining the fundamentals of Lie group integrators
- For material which is more "background type" we use slides (made available to all participants)
- For material at the core of the topic we may write live on iPad
- There will be exercises every afternoon to practice the understanding of the material
- Supplementary material will be provided as documents

- ① *Introduction, motivation and historical remarks*
- ② *A primer on Lie group methods*
- ③ *Manifolds*
- ④ *Lie groups and Lie algebras*

II.1 Introduction, motivation and historical remarks

Traditionally one considered initial value problems for ODEs in a black box sense

$$\dot{y}(t) = f(t, y(t)), \quad y(0) = y_0 \quad f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

- No more detail or structure was assumed on $f(t, y)$
- Later one found reason to separate “stiff” and “nonstiff”
- Then came DAE problems, mix between ODEs and algebraic equations
- **Geometric** or **structure preserving** methods became important in the numerical analysis community from ca 1990.
- **Lie group integrators** is a subfield of Geometric integration that was studied systematically from the early 1990's.

Consider as an example the Euler equations for the free rigid body (angular momentum equations)

$$\dot{m}(t) = m(t) \times \mathbb{I}^{-1} m(t), \quad \mathbb{I} \text{ inertia tensor}$$

Then $\frac{d}{dt} \|m(t)\|^2 = \langle m(t), m(t) \times \mathbb{I}^{-1} m(t) \rangle = 0$

Constant $\|m(t)\|$ is associated to the sphere as a submanifold of \mathbb{R}^3 .

Evolution of the solution should effectively be by rotations

The space of rotations is a Lie group

Stiefel manifold

- Let $M_{d,k}$ be the manifold of $d \times k$ -matrices with orthonormal columns.
- and $\mathfrak{so}(d)$ the skew-symmetric $d \times d$ -matrices, $A^T = -A$.

Consider matrix-differential equation

$$\dot{Y} = A(Y) \cdot Y, \quad A : M_{d,k} \rightarrow \mathfrak{so}(d)$$

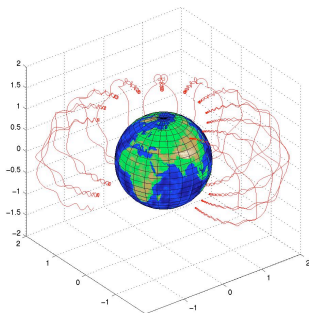
Invariant: $I(Y) = Y^T Y$.

Applications

- Computation of Lyapunov exponents
- Multi-variate data analysis
- Image/signal processing

If numerical solution $Y_n \mapsto Y_{n+1}$ can be evolved by $Y_{n+1} = Q_n Y_n$ and Q_n is an orthogonal $d \times d$ matrix, then $I(Y_{n+1}) = I(Y_n)$.

Northern light



Equation for particle movement (Carl Størmer)

$$\ddot{\mathbf{x}} = \dot{\mathbf{x}} \times \mathbf{d}(\mathbf{x})$$

\mathbf{x} particle position, $\mathbf{d}(\mathbf{x})$ earth magnetic field at \mathbf{x} .



Spiralling movement not easily followed by straight lines

Solving PDEs by means of “simpler” PDEs

We take as example a non-homogeneous heat equation

$$u_t = \nu(\mathbf{x})\Delta u$$

Fast solvers are available for the equation

$$u_t = \bar{\nu}\Delta u + f(\mathbf{x}).$$

The first problem can be approximated locally by the second, e.g. set

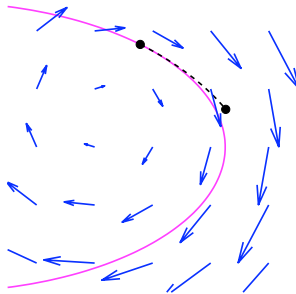
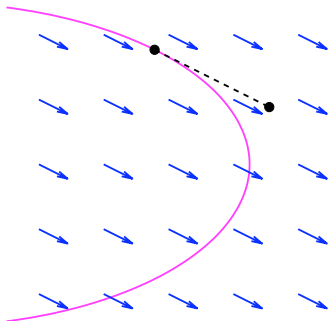
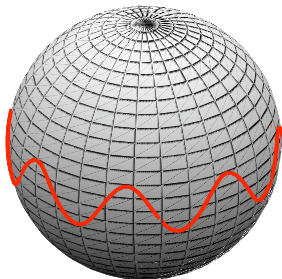
$$\bar{\nu} = \frac{\int \nu(\mathbf{x}) d\mathbf{x}}{\int d\mathbf{x}}$$

for a local known approximation $u^*(\cdot, t^*)$ let

$$f(\mathbf{x}) = (\nu(\mathbf{x}) - \bar{\nu})\Delta u^*$$

Solving the simple PDE is an action by a Lie group

- When the solution is known to be restricted to some manifold
- When it is useful to be able to move along curves rather than straight line segments



“Lie group equation”

$$\dot{y} = A(y) \cdot y, \quad A \text{ is a matrix.}$$

Given y_n , we could approximate this equation locally by the problem

$$\dot{y} = A(y_n) \cdot y \quad \Rightarrow \quad y_{n+1} = e^{hA(y_n)} \cdot y_n$$

This is called the **Lie-Euler** method.

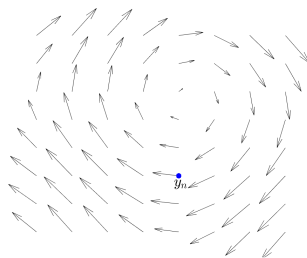
Many questions remain to be answered

- 1 What is the form of $A(y)$ and y ?
- 2 Are these all the problems we can solve?
- 3 What does it have to do with Lie groups and manifolds?
- 4 How can we get higher order of convergence?

II.2 A primer on Lie group methods

$$\dot{y} = f(y),$$

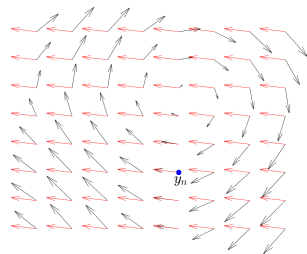
$$y_{n+1} = y_n + hf(y_n)$$



A new interpretation: **Solve exactly the local problem**

$$\dot{z} = f(y_n), \quad z(0) = y_n$$

Set $y_{n+1} = z(h)$, $t_{n+1} = t_n + h$.



We consider the Euler free rigid body

$$\dot{m} = m \times \mathbb{I}^{-1} m, \quad \mathbb{I} = \text{diag}(I_1, I_2, I_3)$$

We can rewrite this as

$$\begin{pmatrix} \dot{m}_1 \\ \dot{m}_2 \\ \dot{m}_3 \end{pmatrix} = \begin{pmatrix} 0 & \frac{m_3}{I_3} & -\frac{m_2}{I_2} \\ -\frac{m_3}{I_3} & 0 & \frac{m_1}{I_1} \\ \frac{m_2}{I_2} & -\frac{m_1}{I_1} & 0 \end{pmatrix} \cdot \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = A(m) \cdot m \quad (*)$$

Now $\|m(t)\| = \|m(0)\|$ for all t since $A(m)$ is skew-symmetric and

$$\frac{1}{2} \frac{d}{dt} (m^T m) = m^T \dot{m} = m^T A(m) m = 0$$

Inspired by the “New look” we could replace (*) by

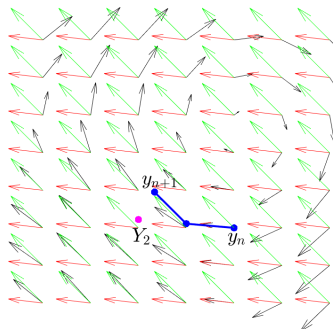
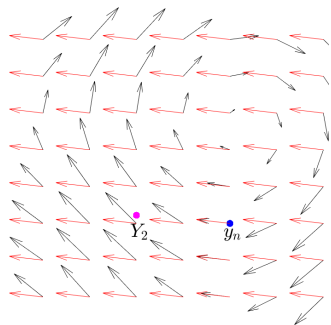
$$\dot{z} = A(m_n) z, \quad z(0) = m_n$$

Solution: $z(t) = e^{tA(m_n)} m_n$ so $m_{n+1} = e^{hA(m_n)} m_n$ (Lie-Euler)

$$m_{n+1} = e^{hA(m_n)} \cdot m_n$$

- The exponential here is the **matrix exponential**
- When a matrix A is skew-symmetric then $Q = e^A$ is **orthogonal**,
 $Q^T Q = I$
- Orthogonal matrices preserve the Euclidean norm $\|Qx\| = \|x\|$ for any vector x
- This means that the Lie-Euler method preserves $\|m\|$, i.e.
 $\|m_n\| = \|m_0\|$ for all n .

Modified Euler



$$Y_1 = y_n,$$

$$Y_2 = y_n + hf(Y_1),$$

$$y_{n+1} = y_n + \frac{h}{2}(f(Y_1) + f(Y_2))$$

Geometric interpretation
is ambiguous.

Details on next slide

Step 1.

Solve $\dot{z} = f(y_n)$, $z(0) = y_n$.

Set $Y_2 = z(h)$. Then

Interpretation 1

Evaluate $\bar{z} := z(h/2)$ from step 1

Solve $\dot{w} = f(Y_2)$, $w(0) = \bar{z}$

Set $y_{n+1} = w(h/2)$

FRB version

$$m_{n+1} = e^{\frac{h}{2}A_2} e^{\frac{h}{2}A_1} m_n$$

Here

$$A_1 = A(m_n), \quad A_2 = A(e^{hA_1} m_n)$$

Interpretation 2

Let $\bar{f} = \frac{1}{2}(f(Y_1) + f(Y_2))$

Solve $\dot{w} = \bar{f}$, $w(0) = \bar{y}_n$

Set $y_{n+1} = w(h)$

FRB version

$$m_{n+1} = e^{\frac{h}{2}(A_1 + A_2)} m_n$$

Runge–Kutta standard

$$Y_1 = y_n$$

$$Y_i = y_n + h \sum_{j=1}^{i-1} a_{ij} f(Y_j),$$

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i f(Y_i)$$

Runge–Kutta Lie (naive)

$$M_1 = m_n$$

$$M_i = \exp \left(h \sum_{j=1}^{i-1} a_{ij} A(M_j) \right) m_n,$$

$$m_{n+1} = \exp \left(h \sum_{i=1}^s b_i A(M_i) \right) m_n$$

Alternatively, in the Runge–Kutta Lie, we could have defined

$$M_i = e^{ha_{i,i-1}M_{i-1}} \cdot e^{ha_{i,i-2}M_{i-2}} \dots e^{ha_{i,1}M_1} m_n$$
$$m_{n+1} = e^{hb_s M_s} \cdot e^{hb_{s-1}M_{s-1}} \dots e^{hb_1 M_1} m_n$$

Hybrids between the two RK-Lie methods can also be considered.
For the first type, no method of order higher than $p = 2$ can be achieved.

First order methods with one stage

$$m_{n+1} = e^{hA(m_n)} \cdot m_n \quad (\text{Lie-Euler})$$

Explicit methods with two stages. Write $M_1 = m_n$, $A_i := A(M_i)$

$$m_{n+1} = e^{h(b_1 A_1 + hb_2 A_2)} m_n, \quad M_2 = e^{ha_{21} A_1} m_n$$

or

$$m_{n+1} = e^{hb_2 A_2} e^{hb_1 A_1} m_n, \quad M_2 = e^{ha_{21} A_1} m_n$$

Second order whenever $b_1 + b_2 = 1$ and $b_2 a_{21} = \frac{1}{2}$.

Important fact: $e^{A+B} \neq e^A e^B$ in general for matrices A and B .

We define: Matrix commutator $[A, B] = AB - BA$ for matrices A and B .

$$A_1 = hA(m_n),$$

$$A_2 = hA(\exp(\frac{1}{2}A_1) \cdot m_n),$$

$$A_3 = hA(\exp(\frac{1}{2}A_2 - \frac{1}{8}[A_1, A_2]) \cdot m_n),$$

$$A_4 = hA(\exp(A_3) \cdot m_n),$$

$$m_{n+1} = \exp(\frac{1}{6}(A_1 + 2A_2 + 2A_3 + A_4 - \frac{1}{2}[A_1, A_4])) \cdot m_n.$$

This is a generalisation of the “classical” Runge–Kutta method of order 4 found in all the text books.

Commutator-free Lie group method

$$M_1 = m_n$$

$$M_2 = \exp\left(\frac{1}{2}hA_1\right) \cdot m_n$$

$$M_3 = \exp\left(\frac{1}{2}hA_2\right) \cdot m_n$$

$$M_4 = \exp\left(hA_3 - \frac{1}{2}hA_1\right) \cdot M_2$$

$$m_{n+\frac{1}{2}} = \exp\left(\frac{1}{12}h(3A_1 + 2A_2 + 2A_3 - A_4)\right) \cdot m_n$$

$$m_{n+1} = \exp\left(\frac{1}{12}h(-A_1 + 2A_2 + 2A_3 + 3A_4)\right) \cdot m_{n+\frac{1}{2}}$$

where $A_i = f(M_i)$.

Note: one exponential is saved in computing M_4 by making use of M_2 .

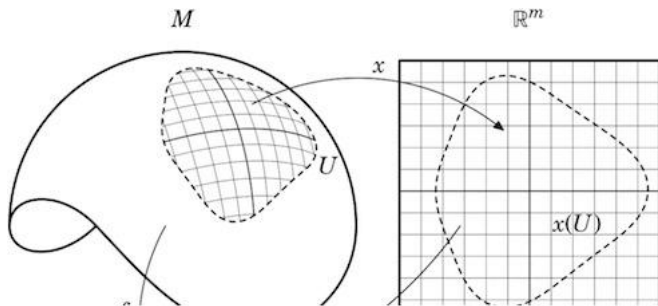
- We have consider **one** simple model problem of the type $\dot{m} = A(m) \cdot m$ where A is a matrix and m is a vector.
- We have naively generalised an interpretation of standard Runge–Kutta scheme, breaking them down into building blocks that consist of solving simpler differential equations exactly.
- We get away with this for methods of convergence order $p \leq 2$.
- For $p > 2$ we need to either include extra corrections (commutators) or compose together building blocks to get the right order

Several open questions remain.

II.3 Manifolds

A manifold is a set M with a collection of charts (U, φ) such that

- $U \subset M$
- $\varphi: U \rightarrow \varphi(U) \subset \mathbb{R}^n$ is a bijective map
- $\varphi(m) = (x_1, \dots, x_n)$ are called coordinates of the point m ,



$(U, \varphi), (U', \varphi')$ overlapping:

$$V = \varphi(U \cap U') \subset \mathbb{R}^n,$$

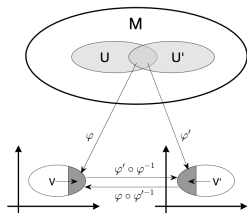
$$V' = \varphi'(U \cap U') \subset \mathbb{R}^n.$$

(U, φ) and (U', φ') compatible if

$$\varphi' \circ \varphi^{-1} : V \rightarrow V'$$

$$\varphi \circ (\varphi')^{-1} : V' \rightarrow V$$

are C^∞ .



Differentiable manifold

- 1 There is a collection of charts such that each $m \in M$ is a member of at least one chart
- 2 M is a union of compatible charts

Two (for us) useful definitions

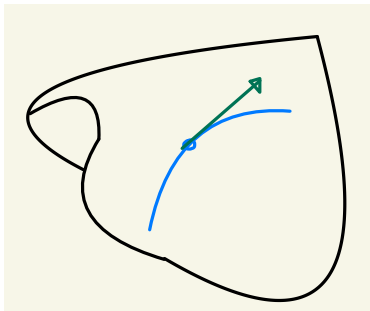
- 1 By curves
- 2 By derivations

$$\gamma(t), t \in C^1(-\varepsilon, \varepsilon)$$

$$v_m = \dot{\gamma}(0), \text{ tangent}$$

$$\text{vector at } m = \gamma(0).$$

Curve



Derivation acting on function germs. A tangent vector v_m can be seen as a linear operator acting on functions on M

- $v_m[\alpha f + \beta g] = \alpha v_m[f] + \beta v_m[g]$ (linearity)
- $v_m[fg] = v_m[f]g(m) + f(m)v_m[g]$ (derivation property)

Interpretation: $v_m[f]$ is the **directional derivative** of f in the direction of v_m at m . In coordinates $v_m = \mathbf{v} \cdot \nabla$.

The 2-sphere S^2 .

Vectors of unit norm,

$$\{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}.$$

Curve $\gamma(t)$ must satisfy

$$\sum_{i=1}^3 \gamma_i(t)^2 = 1$$

Differentiating wrt. t ,

$$\sum_{i=1}^3 \dot{\gamma}_i(t) \gamma_i(t) = 0$$

Suppose $\gamma(0) = r$ and $\dot{\gamma}(0) = v$,

$$T_r S^2 = \{v \in \mathbb{R}^3 : v \perp r\}$$

The Euclidean space \mathbb{R}^n .

Tangent space at x : $T_x \mathbb{R}^n \simeq \mathbb{R}^n$.

Curve $x + tv$ for any $v \in \mathbb{R}^n$.

Orthogonal $n \times n$ -matrices $\mathcal{O}(n)$.

Manifold contains identity matrix I .

Curve $\gamma(t)$ through I , i.e.

$$\gamma(0) = I, \quad \dot{\gamma}(0) = v.$$

Orthogonality: $\gamma(t)^T \gamma(t) = I \forall t$

$$\frac{d}{dt} \gamma(t)^T \gamma(t) = \dot{\gamma}(t)^T \gamma(t) + \gamma(t)^T \dot{\gamma}(t)$$

$$t = 0 \Rightarrow v^T + v = 0$$

$$T_I \mathcal{O}(n) = \{v \in \mathbb{R}^{n \times n} : v^T = -v\}$$

Linear space V , dual space V^*

- For $f \in V^*$, $u, v \in V$,

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

$$\alpha, \beta \in \mathbb{R}$$

- Duality pairing, write

$$f(v) = \langle f, v \rangle$$

- Basis for V , e_1, \dots, e_d

- Dual basis $\varepsilon_1, \dots, \varepsilon_d$

$$\varepsilon_i(e_j) = \langle \varepsilon_i, e_j \rangle$$

Cotangent space

- $T_m M$ is a linear space
- $T_m^* M$ its dual
- $v \in T_m M$ velocity vector
- $p \in T_m^* M$ momentum
- Kinetic energy

$$T = \frac{1}{2} \langle p, v \rangle$$

Smoothly glue together the (co)tangent spaces at each m

$$TM = \bigcup_{m \in M} T_m M, \quad T^*M = \bigcup_{m \in M} T_m^* M$$

Note

- These bundles are not (generally) linear spaces, but they are manifolds in their own right.
- Local coordinate charts are induced on TM from M .

$$\phi : M \supset U \rightarrow \phi(U) \longrightarrow \phi' : TM \supset TU \rightarrow (U \times V)$$

- Extra structure is needed to connect/compare $v \in T_m M$ and $v' \in T_{m'} M$.
- If it holds that e.g. $TM = M \times V$ for some linear space V then the manifold is called **trivial**

Maps and tangent maps

Let M and N be manifolds and $\Psi : M \rightarrow N$ a map.

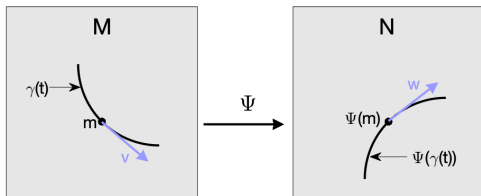
The **tangent** map $T\Psi_m : T_m M \rightarrow T_{\Psi(m)} N$ is defined via curves.

Let $n = \Psi(m)$.

$$\gamma(t) \in M, \quad \gamma(0) = m, \quad \dot{\gamma}(0) = v$$

The curve $\sigma(t) = \Psi(\gamma(t)) \in N$ satisfies $\sigma(0) = n$ and $w := \dot{\sigma}(0) \in T_n N$.

$$w := T\Psi_m(v)$$



One can extend the definition to all of TM , $T\Psi : TM \rightarrow TN$. Then only linear when restricted to fibers T_mM :

Exercise. Local coordinates (x_1, \dots, x_m) on M and similarly, (y_1, \dots, y_n) on N . Thus, $y = \Psi(x)$ can be expressed as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \Psi_1(x_1, \dots, x_m) \\ \vdots \\ \Psi_n(x_1, \dots, x_m) \end{bmatrix}$$

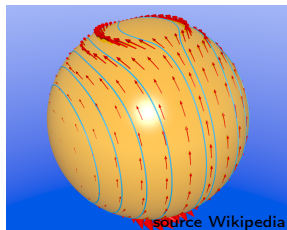
Show, by using the definition of the tangent map that $T\Psi$ is just the Jacobian matrix of Ψ , i.e.

$$T\Psi = \begin{bmatrix} \frac{\partial \Psi_1}{\partial x_1} & \cdots & \frac{\partial \Psi_1}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial \Psi_n}{\partial x_1} & \cdots & \frac{\partial \Psi_n}{\partial x_m} \end{bmatrix}$$

A **vector field** on M is a map

$$X : M \rightarrow TM \text{ s.t.}$$

$$\forall m : m \mapsto X|_m \in T_m M$$



A vector field on M is called a **section** of TM , write $X \in \mathcal{X}(M)$ or $X \in \Gamma(TM)$. We can also have sections on T^*M . They are called **differential one-forms**.

Example. If f is a function on M , $f : M \rightarrow \mathbb{R}$ then $Tf : TM \rightarrow T\mathbb{R}$. We have $T\mathbb{R} \equiv \mathbb{R} \times \mathbb{R}$, but often omit the first factor.

$$df|_m : T_m M \rightarrow \mathbb{R}, \quad df \in \Gamma(T^*M)$$

- $f \in \mathcal{F}(M) := C^\infty(M)$
- Vector field $X \in \mathcal{X}(M)$
- $X : \mathcal{F}(M) \rightarrow \mathcal{F}(M)$ as

$$X[f](m) = X|_m[f]$$

In coordinates x_1, \dots, x_d

$$X = \sum_{i=1}^d X_i(x) \partial x_i$$

Duality of vector fields and differential forms

If $X \in \Gamma(TM)$ and $\omega \in \Gamma(T^*M)$ then we have a point-wise pairing

$$\langle \omega, X \rangle \in \mathcal{F}(M)$$

For one-forms df one has

$$df(X) = \langle df, X \rangle = X[f]$$

Flow

$X \in \mathcal{X}$: vector field on M .
Consider the ODE

$$\dot{x}(t) = X(x(t)), \quad x(0) = x.$$

Solution $x(t), t \in (\alpha_x, \beta_x) \ni 0$.
Write

$$x(t) = \exp(tX)x$$

$\exp(X) : M \rightarrow M$ is called the
flow of X .

Domain

\exp only defined on open
subset

$$\exp(tX) : \mathcal{D}_t \rightarrow M$$

$$\mathcal{D}_t = \{x \in M : t \in (\alpha_x, \beta_x)\}$$

$$\bigcup_{t>0} \mathcal{D}_t = M$$

Relatedness

Let $\Psi : M \rightarrow N$ be a differentiable map and let X and Y be vector fields on M and N respectively. If

$$T\Psi \circ X = Y \circ \Psi$$

then we say that X is Ψ -related to Y .

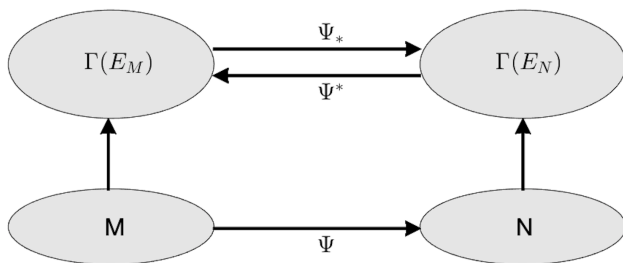
Push-forward. If Ψ is also invertible we can, for any vector field $X \in \mathcal{X}(M)$ define

$$Y = \Psi_* X = T\Psi \circ X \circ \Psi^{-1} \in \mathcal{X}(N)$$

Pull-back. $\Psi^* : \mathcal{X}(N) \rightarrow \mathcal{X}(M)$

$$X = \Psi^* Y = (T\Psi)^{-1} \circ Y \circ \Psi$$

Note. If $y(t) = \Psi(x(t))$, $\dot{x} = X(x)$, $\dot{y} = Y(y)$, then X is Ψ -related to Y .



Functions. $F : N \rightarrow \mathbb{R}$

$$\Psi^*F = F \circ \Psi \text{ i.e. } \Psi^*F(x) = F(\Psi(x))$$

Differential one-forms. Let ω be a one-form on N , $X \in \mathcal{X}(M)$.

$$\langle \Psi^*\omega, X \rangle = \langle \omega, T\Psi(X) \rangle$$

Let X and Y be vector fields on M . They act as derivations on functions $f \in \mathcal{F}(M)$. We define the vector field

$$Z = [X, Y] \in \mathcal{X}(M)$$

as the derivation

$$Z[f] = X[Y[f]] - Y[X[f]] \text{ for every } f \in \mathcal{F}(M)$$

In coordinates (x_1, \dots, x_d)

$$Z^i = \sum_{j=1}^d \left(X_j \frac{\partial Y^i}{\partial x_j} - Y_j \frac{\partial X^i}{\partial x_j} \right)$$

Pushforward homomorphism

Let $\Psi : M \rightarrow N$

$X, Y \in \mathcal{X}(M)$. Then

$$\Psi_*[X, Y] = [\Psi_*X, \Psi_*Y]$$

- The Lie-Jacobi bracket is bilinear and skew-symmetric
- One can prove that $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

A **frame** is a set of vector fields on a manifold M , say E_1, \dots, E_d such that for each $m \in M$

$$\text{span}(E_1|_m, \dots, E_d|_m) = T_m M$$

Clearly $d \geq \dim M$. Sometimes, frames are defined only locally, on a subset $U \subset M$ and one requires $d = \dim M$.

Representing arbitrary vector fields by using a frame.

Any smooth vector field on M can be represented by means of a frame and a set of d functions $f_i \in \mathcal{F}(M)$

$$F|_m = \sum_{i=1}^d f_i(m) E_i|_m$$

II.4 Lie groups and Lie algebras

A Lie group G is a differentiable manifold which also has the structure of a **group**

$$a, b, c \in G$$

$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$

There is $e \in G$ such that

$$a \cdot e = e \cdot a = a, \quad \forall a \in G$$

For every $a \in G$ there is $a^{-1} \in G$ s.t.

$$a \cdot a^{-1} = a^{-1} \cdot a = e$$

Smoothness

$(a, b) \mapsto a \cdot b$ is smooth

$a \mapsto a^{-1}$ smooth

Matrix Lie groups

$GL(n)$: invertible $g \in \mathbb{R}^{n \times n}$

$G \subset GL(n)$ closed subgroup

Commutative Lie groups

$g \cdot h = h \cdot g$ for all $g, h \in G$

$(\mathbb{R}^n, +)$ is an example

We always have “group product” = matrix-matrix multiplication
and neutral element = Identity matrix.

- 1 $SL(n) = \{A : \det(A) = 1\}$
- 2 $O(n) = \{Q : Q^T Q = I\}$ (Orthogonal matrices).
Subgroup $SO(n) = \{Q \in O(n) : \det(Q) = 1\}$.
- 3 $T(n) = \{\text{lower-triangular matrices}\}$
- 4 Special Euclidean group $SE(n)$.

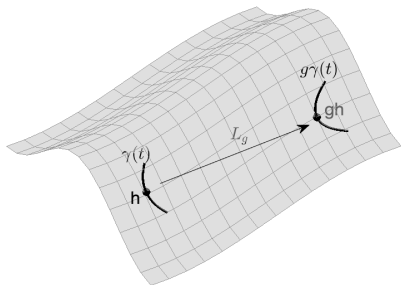
$$A = \begin{bmatrix} Q & a \\ 0 & 1 \end{bmatrix}, \quad Q \in SO(n), \quad a \in \mathbb{R}^n$$

Left and right product

$$g, h \in G$$

$$L_g : G \rightarrow G, L_g(h) = gh$$

$$R_g : G \rightarrow G, R_g(h) = hg$$



If $TL_g \circ X = X \circ L_g$ then X is a **left invariant vector field**.

If $TR_g \circ X = X \circ R_g$ then X is a **right invariant vector field**.

The linear space of left (right) invariant vector fields on G is called the **Lie algebra** of G .

A left invariant vector field X is given at e as $X|_e$, then

$$X|_g = TL_g \circ X|_e$$

- 1 Any tangent vector $v \in T_e G$ therefore uniquely identifies a left invariant vector field.
- 2 Any left invariant vector field X on G defines an element $v := X|_e \in T_e G$.
- 3 We can take $\mathfrak{g} := T_e G$ as an alternative definition of the Lie algebra.

Lie bracket on \mathfrak{g}

The Lie-Jacobi bracket on $\mathcal{X}(G)$ restricted to the space of left (or right) invariant vector fields provides a Lie bracket on \mathfrak{g} .

$$L_{g^*}[X, Y] = [L_{g^*}X, L_{g^*}Y] = [X, Y]$$

The adjoint representation of G

To the group element $g \in G$ there corresponds a map

$$A_g : G \rightarrow G, \quad A_g(h) = ghg^{-1} = L_g \circ R_{g^{-1}}(h)$$

Its tangent map $TA_g|_e$ is called **the adjoint representation of G** .

$$\text{Ad}_g(\xi) = TA_g|_e(\xi) = TL_g \circ TR_{g^{-1}}(\xi), \quad \xi \in \mathfrak{g}.$$

For matrix groups we simply write

$$\text{Ad}_g(\xi) = g\xi g^{-1}$$

Alternative definition of the Lie bracket. Let $g(t) \in G$ be such that $g(0) = e$, $\dot{g}(0) = \eta$. Then

$$[\eta, \xi] = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{g(t)}(\xi)$$

We define $\text{ad}_\xi : \mathfrak{g} \rightarrow \mathfrak{g}$ for any $\xi \in \mathfrak{g}$ as

$$\text{ad}_\xi(\eta) = [\xi, \eta]$$

- ad_ξ is a linear map
- The map is always singular, since $\text{ad}_\xi(\xi) = 0$
- Powers of ad_ξ are iterated commutators, e.g.

$$\text{ad}_\xi^2(\eta) = \text{ad}_\xi([\xi, \eta]) = [\xi, [\xi, \eta]] \text{ etc}$$

- In geometric mechanics, it often happens that the “battle field” is the dual \mathfrak{g}^* of a Lie algebra
- For instance, it has a natural structure as a Poisson manifold
- E.g. The heavy top has $\mathfrak{se}(3)^*$ as solution space
- The coadjoint representation can be central in these cases

Coadjoint representaton.

$$\text{Ad}_g^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*, \quad g \in G.$$

$$\langle \text{Ad}_g^*(\mu), \xi \rangle = \langle \mu, \text{Ad}_g(\xi) \rangle, \quad \text{for all } \mu \in \mathfrak{g}^*, \xi \in \mathfrak{g}$$

Example $SO(3)$. Vector representation: $\text{Ad}_g(\xi) = g\xi$, pairing $\langle \mu, \xi \rangle = \mu^T \xi$.

$$\langle \text{Ad}_g^*(\mu), \xi \rangle = \langle \mu, g\xi \rangle = \mu^T g\xi = \langle g^T \mu, \xi \rangle$$

so $\text{Ad}_g^*(\mu) = g^T \mu$.

Matrix Lie groups

Let $G \subset GL(n)$ be a Lie group of matrices. Lie algebra $\mathfrak{g} = T_1G$.
Take $g(t) = e^{t\xi} \in G$ so that $g(0) = I$ and $\dot{g}(0) = \xi$.

$$[\xi, \eta] = \left. \frac{d}{dt} \right|_{t=0} e^{t\xi} \eta e^{-t\xi} = \xi\eta - \eta\xi$$

How does T_1G look like for a given Lie group G ?

- Look at “defining condition”, such as $g^T g = I$ or $\det(g) = 1$
- Plug in a curve $g(t)$ such that $g(0) = I$.
- Differentiate and evaluate at $t = 0$
- Obtain relation for elements of Lie algebra

Example. Suppose $g(t)^T g(t) = I$. $\xi := \dot{g}(0)$. Differentiate

$$\dot{g}^T g + g^T \dot{g} = 0 \quad \Rightarrow \quad \xi^T + \xi = 0$$

$$SO(3) = \{g \in \mathbb{R}^{3 \times 3} : g^T g = I \text{ and } \det(g) = 1\}$$

Lie algebra

$$\mathfrak{so}(3) = \{\xi \in \mathbb{R}^{3 \times 3} : \xi^T + \xi = 0\}$$

Basis for $\mathfrak{so}(3)$

$$E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\hat{\xi} = \xi_1 E_1 + \xi_2 E_2 + \xi_3 E_3 \text{ and } \xi = (\xi_1, \xi_2, \xi_3)^T.$$

$$\hat{\xi} = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \quad (\text{hat map})$$

By using $\xi \in \mathbb{R}^3$ to represent $\hat{\xi}$, we can write ad_ξ as a matrix and we find

$$\text{ad}_\xi(\eta) = [\hat{\xi}, \hat{\eta}] = \hat{\zeta}, \quad \zeta = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}$$

We can also use the cross product for the commutator

$$\zeta = \text{ad}_\xi(\eta) = \xi \times \eta$$

Useful formulas

- $\widehat{g\xi} = g\hat{\xi}g^{-1} = \text{Ad}_g(\hat{\xi}), g \in SO(3), \xi \in \mathfrak{so}(3).$
- $g(\xi \times \eta) = (g\xi) \times (g\eta), g \in SO(3), \xi, \eta \in \mathfrak{so}(3).$

Unsurprisingly, the map $\exp : \mathfrak{g} \rightarrow G$ is just the flow of right- or left invariant vector fields.

$$X \in \mathfrak{g}, \quad \dot{x}(t) = X|_{x(t)}, \quad x(0) = e \quad \Rightarrow \quad x(1) = \exp(X) \in G$$

Properties

$$\exp((t + s)X) = \exp(tX) \exp(sX)$$

$$\exp(X + Y) \neq \exp(X) \exp(Y) \text{ (in general)}$$

\exp is neither one-to-one nor onto (in general)

- Find tangent space $T_I G \equiv \mathfrak{g}$. Construct arbitrary curve $A(t)$, $A(0) = I$ and let $v = \dot{A}(0) \in T_I G$.
- Define X_v and Y_v in $\mathcal{X}(G)$

$$X_v|_g = L_{g*}(v) = gv, \quad Y_v|_g = R_{g*}(v) = vg.$$

These are the left- and right invariant vector fields.

- Consider the matrix differential equation

$$\dot{g}(t) = X_v|_{g(t)} = g(t)v, \quad g(0) = e, \quad \Rightarrow \quad g(t) = e^{tv}$$

where the last expression is the **matrix exponential**

■

$$e^{tv} = \sum_{k=0}^{\infty} \frac{t^k}{k!} v^k$$

The derivative of the exponential map

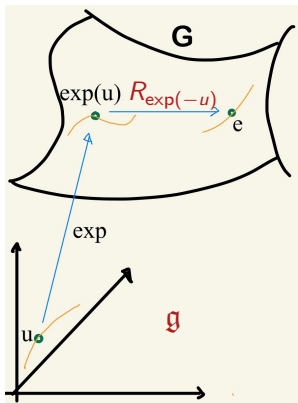
$$\exp : \mathfrak{g} \rightarrow G$$

$$T \exp_u : T_u \mathfrak{g} \rightarrow T_{\exp(u)} G$$

$T_u \mathfrak{g} \equiv \mathfrak{g}$ since \mathfrak{g} is a linear space.

Also $T_{\exp(u)} G \equiv \mathfrak{g}$ through

$$v \in T_{\exp(u)} G \longleftrightarrow w = TR_{\exp(-u)}(v) \in T_e G \equiv \mathfrak{g}$$



We define for $u \in \mathfrak{g}$ the linear map

$$\operatorname{dexp}_u : \mathfrak{g} \rightarrow \mathfrak{g}$$

$$\operatorname{dexp}_u(v) = TR_{\exp(-u)} \circ T \exp_u(v)$$

An infinite series for $\text{dexp}_u(v)$

Recall the ad -operator $\text{ad}_u : \mathfrak{g} \rightarrow \mathfrak{g}$

$$\text{ad}_u(v) = [u, v], \quad \text{ad}_u^2(v) = \text{ad}_u([u, v]) = [u, [u, v]], \quad \text{etc.}$$

One can prove that

$$\text{dexp}_u(v) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_u^k(v) = \left. \frac{e^z - 1}{z} \right|_{z=\text{ad}_u} (v)$$

We shall see later that we need the inverse $\text{dexp}_u^{-1}(v)$. This is obtained by inverting the function above, i.e.

$$\text{dexp}_u^{-1}(v) = \left. \frac{z}{e^z - 1} \right|_{z=\text{ad}_u} (v) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_u^k(v)$$

We shall see later that the series can be truncated after a final number of terms when used in Lie group integrators.

Let G be a Lie group and M a manifold. Lie group action $\Lambda : G \times M \rightarrow M$.

Lie group action $\Lambda : G \times M \rightarrow M$

$$\Lambda(e, m) = m, \quad \forall m \in M$$

$$\Lambda(gh, m) = \Lambda(g, \Lambda(h, m)), \quad \forall g, h \in G, m \in M$$

Remark

The definition above is a **left** action

A **right** action is also possible, in which case we have

$$\Lambda(gh, m) = \Lambda(h, \Lambda(g, m))$$

- ~ If G is a Lie group and $M = G$ then set $\Lambda(g, m) = gm$
- ~ If G is an $n \times n$ matrix group and $M = \mathbb{R}^n$, set $\Lambda(g, x) = gx$
- ~ Let M be the set of symmetric $n \times n$ -matrices and G the Lie group of orthogonal matrices. A group action is

$$\Lambda(g, m) = gmg^T$$

- ~ $G = SL(2)$ acting on the complex plane \mathbb{C} through

$$\Lambda(A, z) = \frac{az + b}{cz + d}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ad - bc = 1$$

The **adjoint representation** defines a **left** Lie group action by G on $M = \mathfrak{g}$

$$\Lambda(g, \xi) = \text{Ad}_g(\xi)$$

The **coadjoint representation** defines a **right** Lie group action by G on $M = \mathfrak{g}^*$

$$\Lambda(g, \mu) = \text{Ad}_g^*(\mu)$$

A **left** coadjoint action can be defined by replacing g by g^{-1}

$$\Lambda(g, \mu) = \text{Ad}_{g^{-1}}^*(\mu)$$

Transitive. For every $x, y \in M$ there is a $g \in G$ such that $y = \Lambda(g, x)$.

Free. For any $m \in M$, $\Lambda(g, m) = m \Rightarrow g = e$.

Orbit. An orbit \mathcal{O}_m containing m is $x = \{\Lambda(g, m), g \in G\}$.

Note. A group action is transitive when restricted to each of its orbits.

The stabilizer (or isotropy) subgroup of a group action is

$$H_x = \{h \in G : \Lambda(h, x) = x\}$$

Note. For a free group action any isotropy subgroup is the trivial group $H_x = \{e\}$.

Note. All isotropy groups are isomorphic to each other.

Spherical pendulum

- Configuration space S^2 , phase space TS^2
- Davide Murari: adjoint action by $SE(3)$ on $\mathfrak{se}(3) \equiv \mathbb{R}^6$

$$\text{Ad}_{(g,u)}(\xi, \eta) = (g\xi, u \times (g\xi) + g\eta)$$

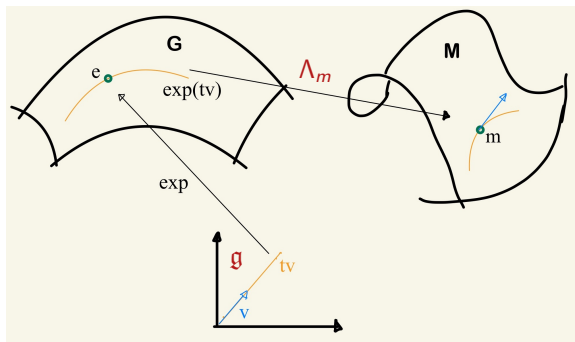
- can be restricted to

$$TS^2 \sim \{(q, v) \in \mathbb{R}^6 : |q| = 1, q^T v = 0\}$$

- So for any $q \in \mathbb{R}^3$, $|q| = 1$ and $v_q \in \mathbb{R}^3$, $q^T v = 0$ let

$$\Lambda((g, u), (q, v)) = (gq, u \times gv)$$

Infinitesimal generator of a Lie group action



M manifold

G Lie group

\mathfrak{g} Lie algebra

$v \in \mathfrak{g}$

$\Lambda_m(g) := \Lambda(g, m)$

Define the infinitesimal generator of the action as

$$\lambda_*(v)|_m = T\Lambda_m|_e(v)$$

$\lambda_*(v)$ is a vector field on M for every $v \in \mathfrak{g}$.

Free rigid body

Left coadjoint action by $SO(3)$ on $\mathfrak{so}(3)^*$

Represent $\mathfrak{so}(3)^*$ as 3-vectors μ

$$\Lambda(g, \mu) = \text{Ad}_{g^{-1}}^*(\mu) = g\mu$$

Transitive on coadjoint orbits: $\{\mu : |\mu| = \text{constant}\}$

The infinitesimal action is

$$\lambda_*(\xi)|_{\mu} = \hat{\xi}\mu = \xi \times \mu$$

The FRB equations are

$$\dot{\mu} = \lambda_*(-\mathbb{I}^{-1}\mu)|_{\mu}$$