THREAD – NWT6 – Part II: Lie group integrators

Brynjulf Owren

Department of Mathematical Sciences, NTNU, Trondheim, Norway

February 5, 8, 9 – 2021

About the lectures

- 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

Content

II.1 Introduction, motivation and historical remarks

Traditional view on solving ordinary differential equations (ODEs)

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

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

- 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.

Lie group integrators – first view – Example 1

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 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

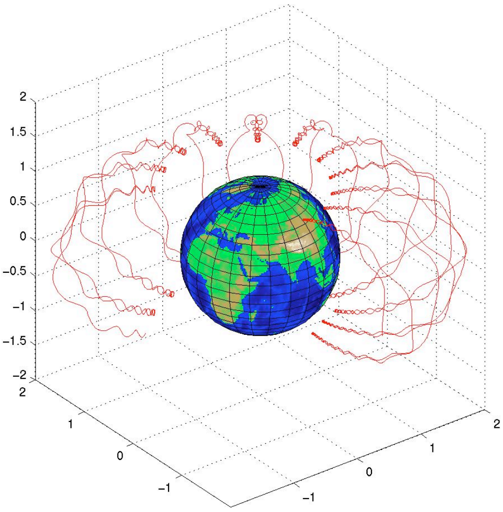
$$\dot{Y} = A(Y) \cdot Y, \quad A : M_{d,k} \to \mathfrak{so}(d)$$

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

Lie group integrators - Second view - Example 4

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

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

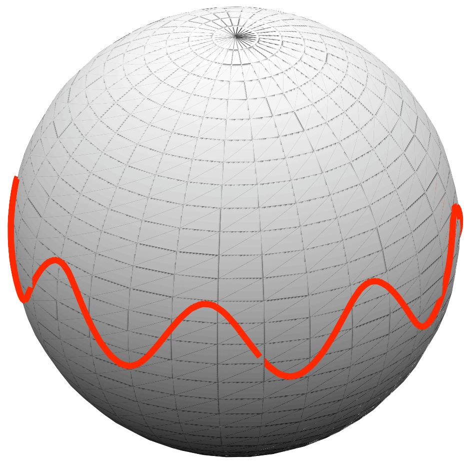
for a local known approximation $u^*(\cdot, t^*)$ let

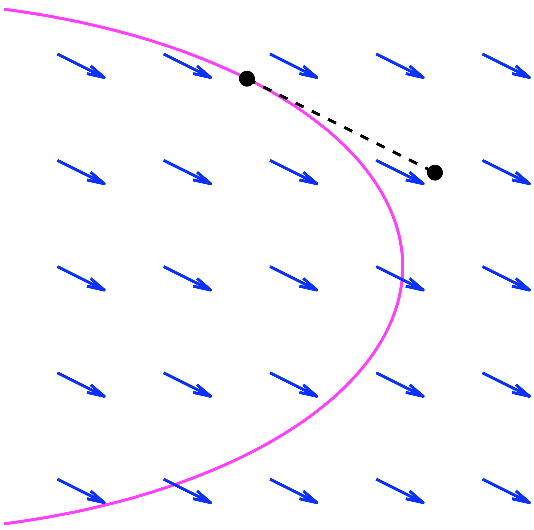
$$f(\mathbf{x}) = (\nu(\mathbf{x}) - \bar{\nu})\Delta u^*$$

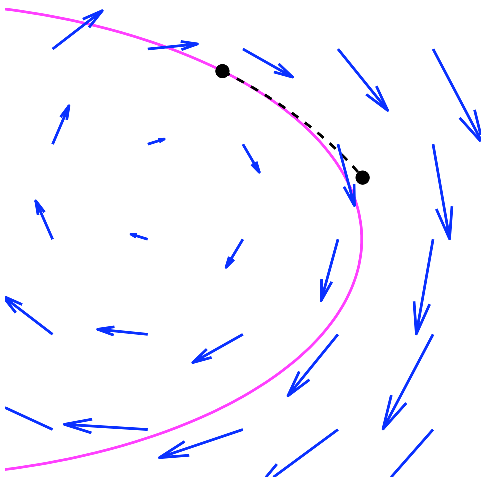
> Solving the simple PDE is an action by a Lie group

Summarising

- 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







The prototypical case

"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

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

This is called the Lie-Euler method.

Many questions remain to be answered

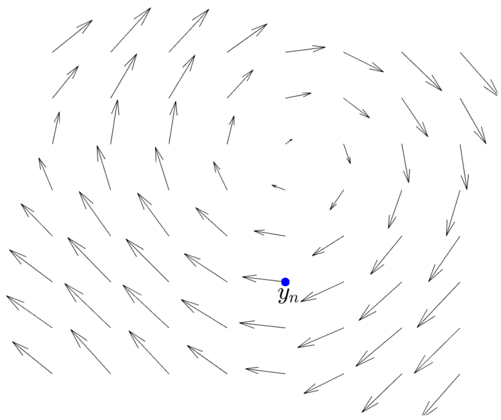
1. What is the form of $A(y)$ and y?
2. Are these all the problems we can solve?
3. What does it have to do with Lie groups and manifolds?
4. How can we get higher order of convergence?

II.2 A primer on Lie group methods

A new look on the Explicit Euler method

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

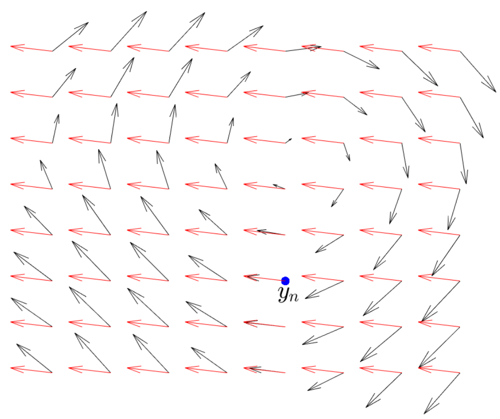
$$y_{n+1} = y_n + hf(y_n)$$



A new interpretation: Solve exactly the local problem

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

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



Free rigid body

We consider the Euler free rigid body

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

We can rewrite this as

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

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

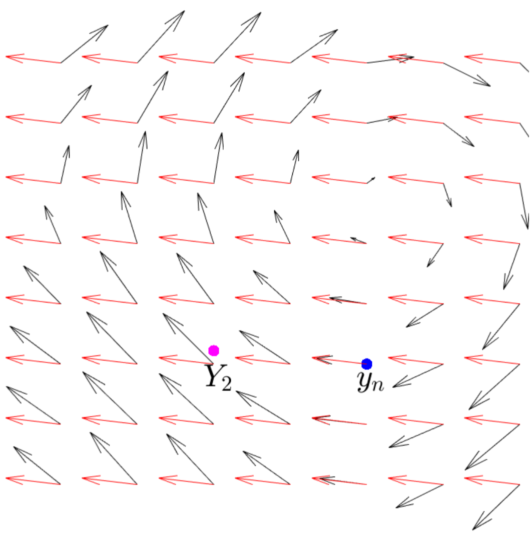
Remarks on Lie-Euler

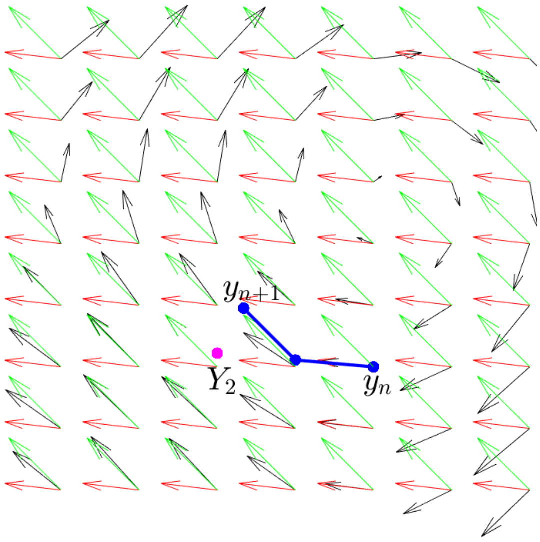
$$m_{n+1} = \mathrm{e}^{hA(m_n)} \cdot m_n$$

- The exponential here is the matrix exponential
- When a matrix A is skew-symmetric then $Q = \mathrm{e}^{A}$ 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.

Stepping it up to Runge–Kutta with two stages

Modified Euler





$$
\begin{aligned}
Y_1 &= y_n, \\
Y_2 &= y_n + hf(Y_1), \\
y_{n+1} &= y_n + \tfrac{h}{2}(f(Y_1) + f(Y_2))
\end{aligned}
$$

Geometric interpretation is ambiguous.
Details on next slide

Geometric interpretation of modified Euler method

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

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

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

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

Here

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

Explicit Runge–Kutta methods

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)

$$M_1 = m_n$$

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

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

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

$$M_i = \mathrm{e}^{ha_{i,i-1}M_{i-1}} \cdot \mathrm{e}^{ha_{i,i-2}M_{i-2}} \cdots \mathrm{e}^{ha_{i,1}M_1}\, m_n$$

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

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.

Examples of Lie group integrators for $\dot{m} = A(m) \cdot m$

First order methods with one stage

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

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

$$m_{n+1} = \mathrm{e}^{h(b_1 A_1 + h b_2 A_2)} m_n, \quad M_2 = \mathrm{e}^{h a_{21} A_1} m_n$$

or

$$m_{n+1} = \mathrm{e}^{h b_2 A_2} \mathrm{e}^{h b_1 A_1} m_n, \quad M_2 = \mathrm{e}^{h a_{21} A_1} m_n$$

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

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

A method of order 4 (of the first type) RKMK method

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

$$\begin{aligned} A_1 &= hA(m_n), \\ A_2 &= hA(\exp(\tfrac{1}{2}A_1) \cdot m_n), \\ A_3 &= hA(\exp(\tfrac{1}{2}A_2 - \tfrac{1}{8}[A_1, A_2]) \cdot m_n), \\ A_4 &= hA(\exp(A_3) \cdot m_n), \\ m_{n+1} &= \exp(\tfrac{1}{6}(A_1 + 2A_2 + 2A_3 + A_4 - \tfrac{1}{2}[A_1, A_4])) \cdot m_n. \end{aligned}$$

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

Commutator-free Lie group method

$$
\begin{aligned}
M_1 &= m_n \\
M_2 &= \exp(\tfrac{1}{2} h A_1) \cdot m_n \\
M_3 &= \exp(\tfrac{1}{2} h A_2) \cdot m_n \\
M_4 &= \exp(h A_3 - \tfrac{1}{2} h A_1) \cdot M_2 \\
m_{n+\frac{1}{2}} &= \exp(\tfrac{1}{12} h (3A_1 + 2A_2 + 2A_3 - A_4)) \cdot m_n \\
m_{n+1} &= \exp(\tfrac{1}{12} h (-A_1 + 2A_2 + 2A_3 + 3A_4)) \cdot m_{n+\frac{1}{2}}
\end{aligned}
$$

where $A_i = f(M_i)$.
Note: one exponential is saved in computing M_4 by making use of M_2.

Summary and remarks so far

- 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

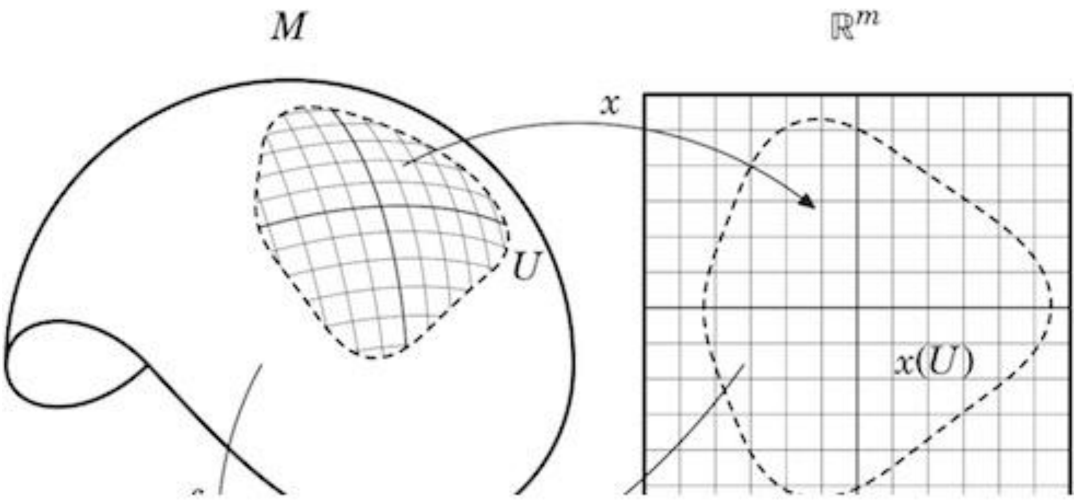
Several open questions remain.

II.3 Manifolds

Manifolds

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

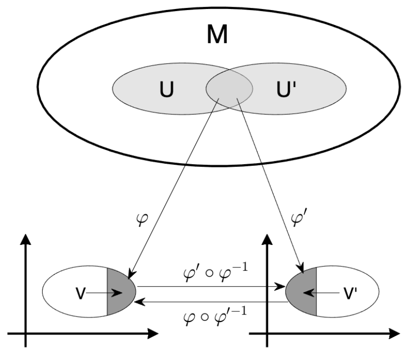
- $U \subset M$
- $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ is a bijective map
- $\varphi(m) = (x_1, \ldots, x_n)$ are called coordinates of the point m,



Compatible charts

(U, φ), (U', φ') 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

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

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

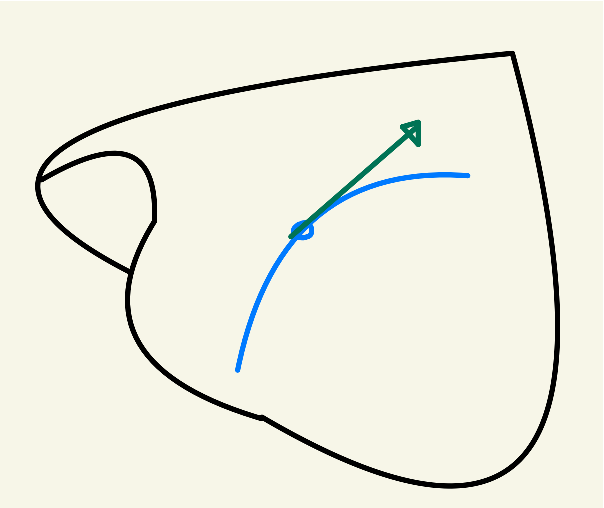
Tangent space

Two (for us) useful definitions

1. By curves
2. By derivations

$\gamma(t),\ t \in C^1(-\varepsilon, \varepsilon)$
$v_m = \dot{\gamma}(0)$, tangent 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 $v_m = \mathbf{v} \cdot \nabla$.

Examples – Manifolds and their tangent spaces

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)^T\gamma(t) = \dot{\gamma}(t)^T\gamma(t) + \gamma(t)^T\dot{\gamma}(t)$$
$$t = 0 \ \Rightarrow v^T + v = 0$$

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

Dual spaces and the cotangent space

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, $e_1, \ldots, e_d$
- Dual basis $\varepsilon_1, \ldots, \varepsilon_d$
  $\varepsilon_i(e_j) = \langle \varepsilon_i, e_j \rangle$

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

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

Tangent and cotangent bundles

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

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

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.

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

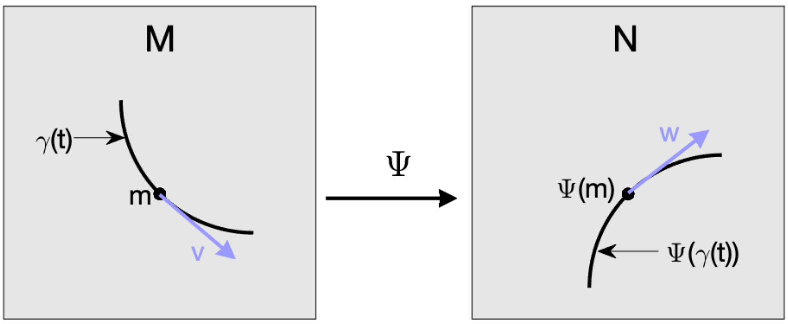
- Extra structure is needed to connect/compare $v \in T_m M$ and $v' \in T_{m'} M$.
- If it holds that e.g $TM = M \times V$ for some linear space V then the manifold is called trivial

Let M and N be manifolds and $\Psi : M \to N$ a map.
The tangent map $T\Psi_m : T_m M \to T_{\Psi(m)} N$ 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
$w := \dot{\sigma}(0) \in T_n N$.

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



Tangent bundle maps

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

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

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

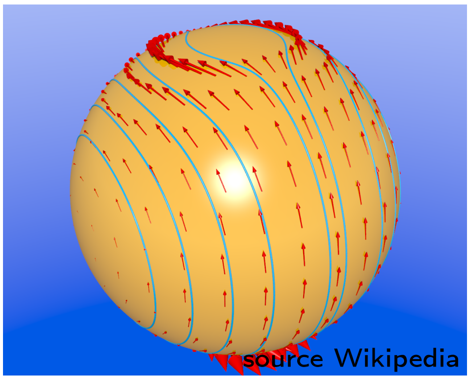
Show, by using the definition of the tangent map that $T\Psi$ is just the Jacobian matrix of Ψ, i.e.

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

Vector fields on manifolds

A vector field on M is a map

$$X : M \to TM \text{ s.t.}$$
$$\forall m : m \mapsto X|_m \in T_m M$$



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, $f : M \to \mathbb{R}$ then $Tf : TM \to T\mathbb{R}$. We have $T\mathbb{R} \equiv \mathbb{R} \times \mathbb{R}$, but often omit the first factor.

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

Vector fields as derivations of functions

- $f \in \mathcal{F}(M) := C^\infty(M)$
- Vector field $X \in \mathcal{X}(M)$
- $X : \mathcal{F}(M) \to \mathcal{F}(M)$ 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 of a vector field

Flow

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

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

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

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

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

Domain

exp only defined on open subset

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

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

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

Relatedness, push-forward and pull-back

Relatedness

Let $\Psi : M \to N$ 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

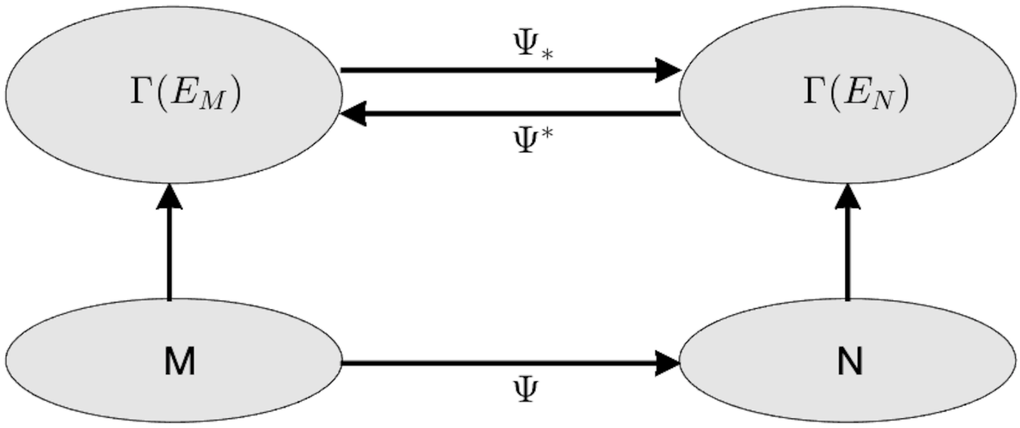
$$Y = \Psi_* X = T\Psi \circ X \circ \Psi^{-1} \in \mathcal{X}(N)$$

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

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

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

Push-forward and pull-back apply also to other objects



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

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

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

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

Lie-Jacobi bracket of vector fields

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 $\Psi : M \to N$
$X, Y \in \mathcal{X}(M)$. Then

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

Frames of vector fields

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

$$\mathrm{span}(E_1|_m, \ldots, E_d|_m\} = T_m M$$

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