

Sextas Matemáticas - UFRJ - 04/10/2024

## **Detecção de órbitas de representações de grupos de Lie compactos a partir de nuvens de pontos**

Henrique Ennes - DataShape/COATI (Sophia Antipolis)

Raphaël Tinarrage - EMAP/FGV (Rio de Janeiro)



Bernhard Riemann  
1826 - 1866



Sophus Lie  
1842 - 1899



Wilhelm Killing  
1847 - 1923



Felix Klein  
1849 - 1925



Élie Cartan  
1869 - 1951



Hermann Weyl  
1885 - 1955

## 1872, F. Klein, *Vergleichende Betrachtungen über neuere geometrische Forschungen*:

As geometrias não euclidianas devem ser estudadas por meio de suas simetrias (*Erlangen program*).

## Inverno 1873, S. Lie:

Um *grupo de Lie* é uma variedade equipada com uma estrutura de grupo. Ele possui uma *álgebra de Lie*, que permite trabalhar infinitesimalmente (correspondência grupo de Lie  $\sim$  álgebra de Lie).

## 1913, E. Cartan, *Teorema do peso máximo*:

As representações irredutíveis dos grupos de Lie são classificadas por seus pesos máximos.

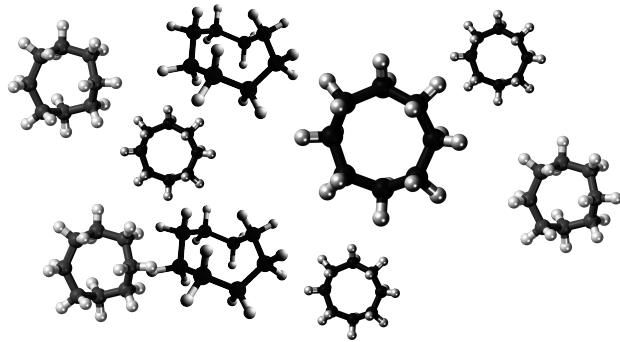
## 1935, V. Fock, *Zur theorie des wasserstoffatoms*:

Descrição do átomo de hidrogênio por meio da simetria  $SO(4)$  na equação de Schrödinger.

## 1939, *Teorema de Myers–Steenrod*:

O grupo de isometria de uma variedade riemanniana é um grupo de Lie.

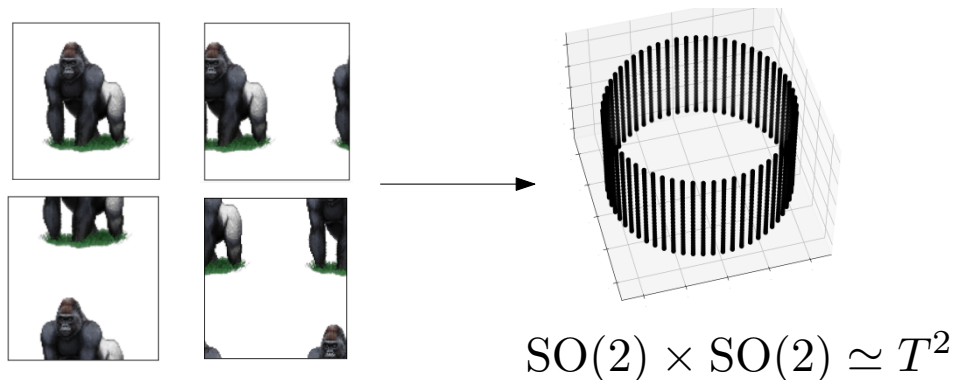
(1) Certos conjuntos reais exibem objetos simétricos.



O espaço de conformação da molécula  $C_8H_{16}$  é a união de **uma garrafa de Klein e uma esfera**.

[Martin, Thompson, Coutsiias & Watson, *Topology of cyclo-octane energy landscape*, 2010]

(2) As transformações euclidianas são regidas por representações de grupos de Lie.



Após aplicar permutação de pixels, as imagens (imersas no  $\mathbb{R}^{n \times n}$ ) se encontram numa **órbita de representação de grupo de Lie**.

(3) Simetrias nos sistemas hamiltonianos resultam em leis de conservação.

**Teorema de Noether (1915):** Se o hamiltoniano for invariante sob a ação de um grupo de Lie, então existe uma constante do movimento.

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}} \quad \frac{d\mathbf{q}}{dt} = \frac{\partial H}{\partial \mathbf{p}}$$



Emmy Noether  
1882 - 1935

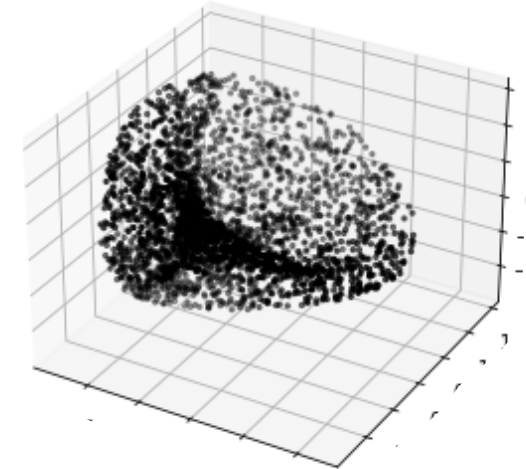
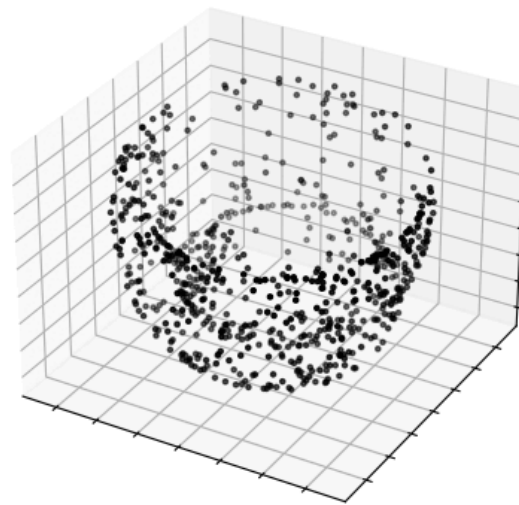
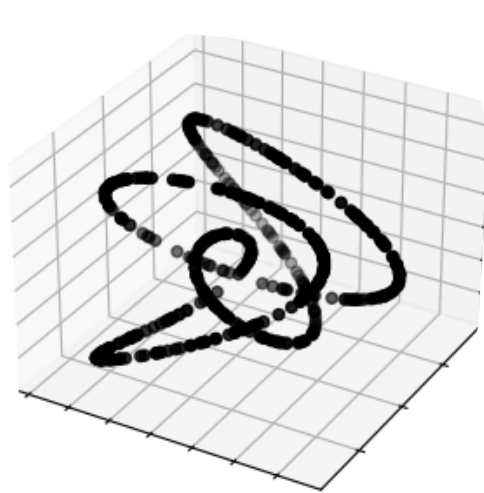
**Input:** Uma nuvem de pontos  $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^n$ .

**Output:** Um grupo de Lie compacto  $G$ , uma representação  $\phi$  no  $\mathbb{R}^n$ , uma órbita  $\mathcal{O}$  próxima a  $X$ .

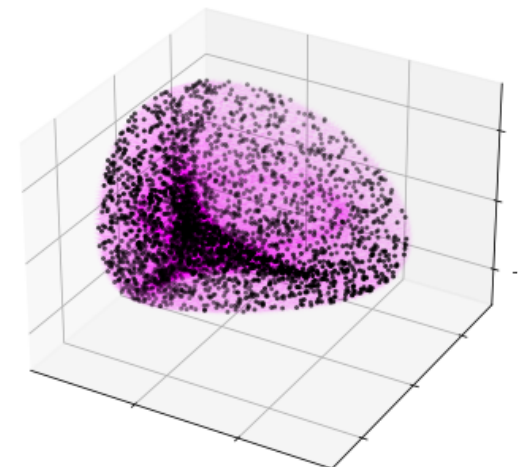
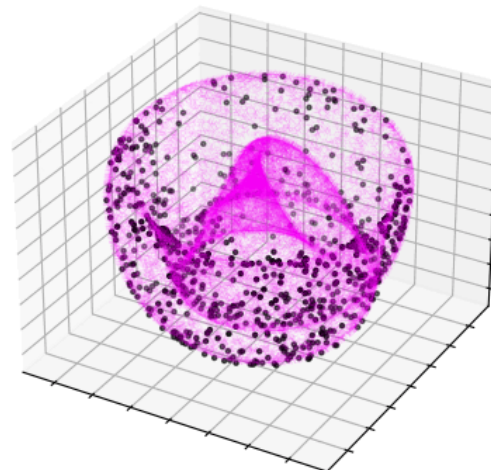
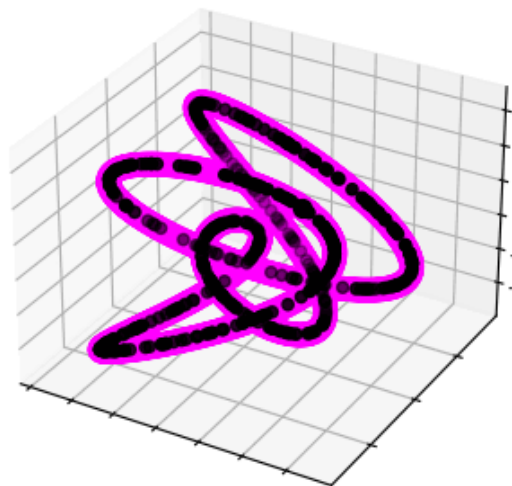
Órbita de  $SO(2)$  no  $\mathbb{R}^6$

Órbita de  $T^2$  no  $\mathbb{R}^6$

Órbita de  $SO(3)$  no  $\mathbb{R}^9$



**Input:**



**Output:**

1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
6. Aplicação: Espaço de conformação do ciclooctano
7. Garantias teóricas

**Definição:** Um **grupo de Lie**  $G$  é um grupo que também é uma variedade suave, e tal que o mapa de multiplicação  $(g, h) \mapsto gh$  e o mapa inverso  $g \mapsto g^{-1}$  são suaves.

**Exemplo:** Dado  $n \in \mathbb{N}$  positivo, temos os **grupos de matrizes**:

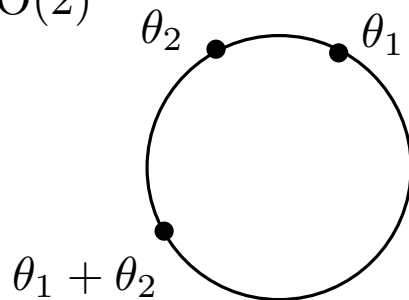
- $O(n)$  grupo ortogonal: o conjunto de matrizes ortogonais  $n \times n$  ( $A^T = A^{-1}$ )
- $SO(n)$  grupo especial ortogonal: matrizes ortogonais  $n \times n$  de determinante  $+1$
- $Sp(2n, \mathbb{C})$  grupo simplético: matrizes complexas simpléticas  $n \times n$
- $U(n)$  grupo unitário: matrizes complexas unitárias  $n \times n$  ( $A^* = A^{-1}$ )
- $SU(n)$  grupo especial unitário: matrizes unitárias complexas  $n \times n$  de determinante  $+1$

**Produtos** de grupos de Lie são grupos de Lie:

- $T^n$   $n$ -toro: o produto  $SO(2) \times \dots \times SO(2)$

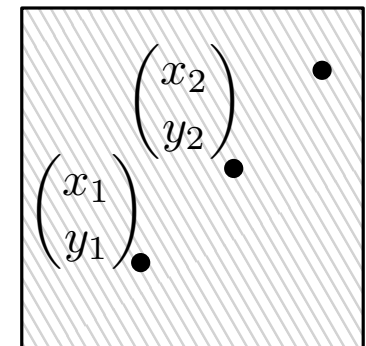
Estrutura de grupo em  $SO(2)$   
(o círculo)

$$\begin{pmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{pmatrix}$$



Estrutura de grupo em  $T^2$   
(o mundo do Pac-Man)

$$\begin{pmatrix} \cos x & -\text{sen } x & 0 & 0 \\ \text{sen } x & \cos x & 0 & 0 \\ 0 & 0 & \cos y & -\text{sen } y \\ 0 & 0 & \text{sen } y & \cos y \end{pmatrix}$$



$$\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}$$

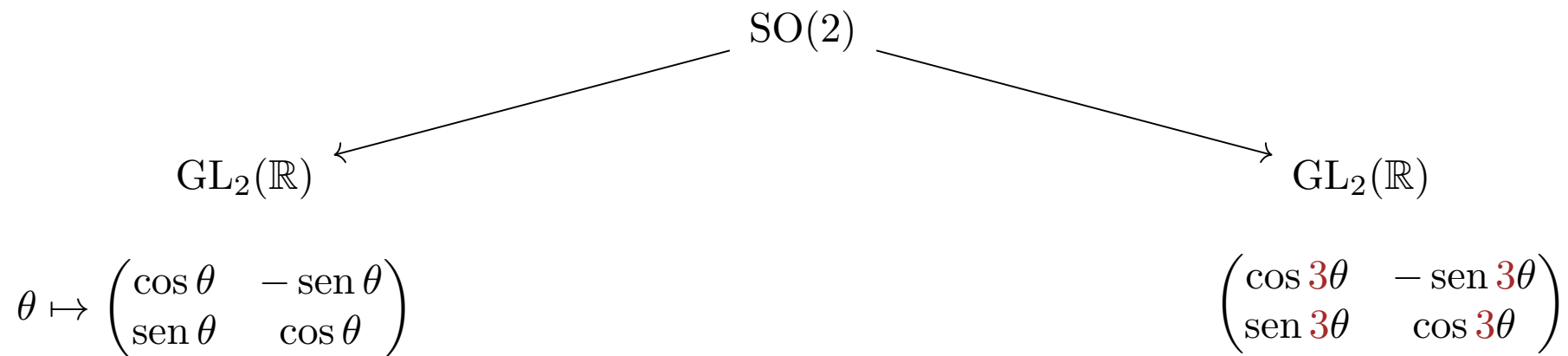
**Definição:** Uma **representação** de um grupo de Lie  $G$  em  $\mathbb{R}^n$  é um morfismo de grupo suave  $\phi: G \rightarrow \text{GL}_n(\mathbb{R})$  (as matrizes invertíveis  $n \times n$ ). Isto é, vale  $\phi(ab) = \phi(a)\phi(b)$ .

**Exemplo:** Obviamente, os grupos de Lie de matrizes vêm com uma **representação canônica**, pois já estão incluídos num espaço matricial.

$$\begin{aligned} \text{O}(n) &\hookrightarrow \text{GL}_n(\mathbb{R}) \\ \text{SO}(n) &\hookrightarrow \text{GL}_n(\mathbb{R}) \end{aligned}$$

$$\begin{aligned} \text{Sp}(2n, \mathbb{C}) &\hookrightarrow \text{GL}_n(\mathbb{C}) \hookrightarrow \text{GL}_{2n}(\mathbb{R}) \\ \text{U}(n) &\hookrightarrow \text{GL}_n(\mathbb{C}) \hookrightarrow \text{GL}_{2n}(\mathbb{R}) \\ \text{SU}(n) &\hookrightarrow \text{GL}_n(\mathbb{C}) \hookrightarrow \text{GL}_{2n}(\mathbb{R}) \end{aligned}$$

Entretanto, existem representações mais sofisticadas.



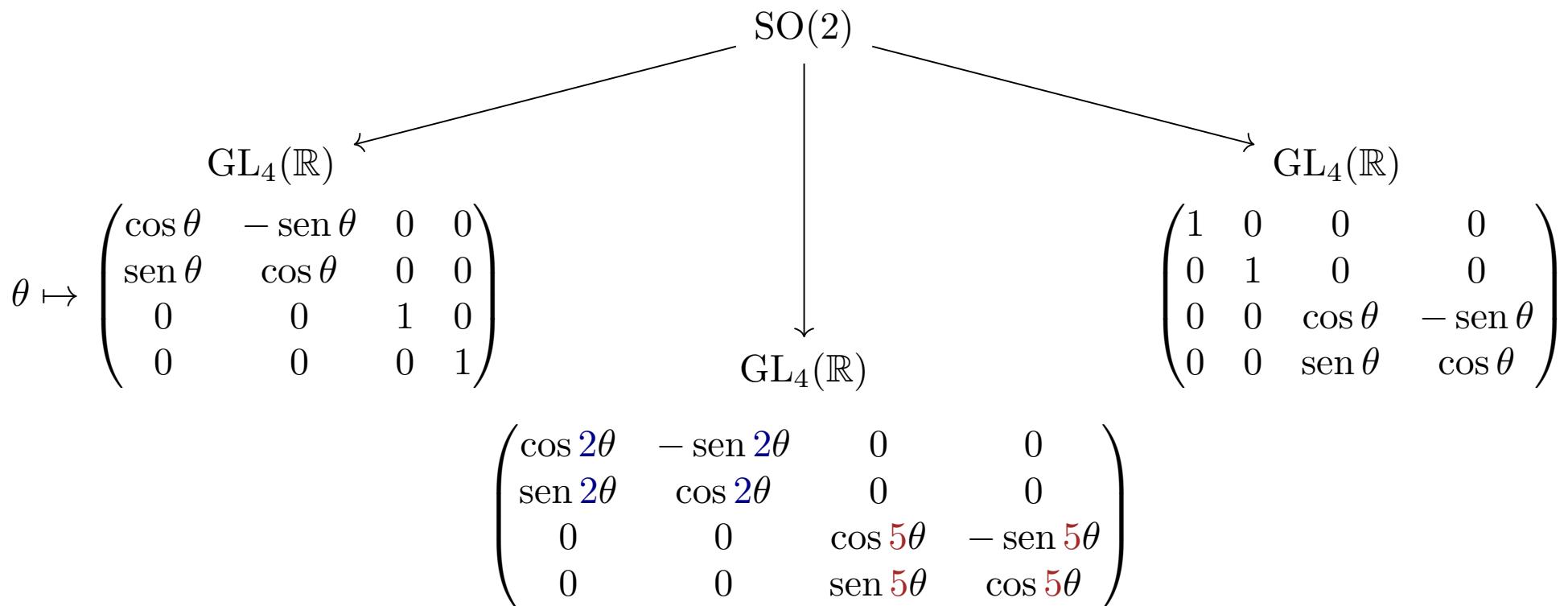
**Definição:** Uma **representação** de um grupo de Lie  $G$  em  $\mathbb{R}^n$  é um morfismo de grupo suave  $\phi: G \rightarrow GL_n(\mathbb{R})$  (as matrizes invertíveis  $n \times n$ ). Isto é, vale  $\phi(ab) = \phi(a)\phi(b)$ .

**Exemplo:** Obviamente, os grupos de Lie de matrizes vêm com uma **representação canônica**, pois já estão incluídos num espaço matricial.

$$\begin{aligned} O(n) &\hookrightarrow GL_n(\mathbb{R}) \\ SO(n) &\hookrightarrow GL_n(\mathbb{R}) \end{aligned}$$

$$\begin{aligned} Sp(2n, \mathbb{C}) &\hookrightarrow GL_n(\mathbb{C}) \hookrightarrow GL_{2n}(\mathbb{R}) \\ U(n) &\hookrightarrow GL_n(\mathbb{C}) \hookrightarrow GL_{2n}(\mathbb{R}) \\ SU(n) &\hookrightarrow GL_n(\mathbb{C}) \hookrightarrow GL_{2n}(\mathbb{R}) \end{aligned}$$

Entretanto, existem representações mais sofisticadas.





**Definição:** Uma **representação** de um grupo de Lie  $G$  em  $\mathbb{R}^n$  é um morfismo de grupo suave  $\phi: G \rightarrow \text{GL}_n(\mathbb{R})$  (as matrizes invertíveis  $n \times n$ ). Isto é, vale  $\phi(ab) = \phi(a)\phi(b)$ .

**Definição:** Duas representações  $\phi_1, \phi_2: G \rightarrow \text{GL}_n(\mathbb{R})$  são **equivalentes** se existir  $A \in \text{GL}_n(\mathbb{R})$  tal que  $\phi_2 = A\phi_1A^{-1}$ .

Elas são “iguais a menos de mudança de coordenadas”.

**Proposição:** As representações de  $\text{SO}(2)$  em  $\mathbb{R}^{2n}$  são classificadas por  $\mathbb{Z}^n / \mathfrak{S}_n$  ( $n$ -uplas a menos de permutação). Mais precisamente, a  $(\omega_1, \dots, \omega_n) \in \mathbb{Z}^n$  é associada  $\phi_{(\omega_1, \dots, \omega_n)}: \text{SO}(2) \rightarrow \text{GL}_{2n}(\mathbb{R})$ .

$$\phi_{(\omega_1, \dots, \omega_n)}(\theta) = \begin{pmatrix} R(\omega_1\theta) & & & \\ & R(\omega_2\theta) & & \\ & & \ddots & \\ & & & R(\omega_n\theta) \end{pmatrix} \quad \text{where} \quad R(\theta) = \begin{pmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{pmatrix}$$

**Definição:** Uma **representação** de um grupo de Lie  $G$  em  $\mathbb{R}^n$  é um morfismo de grupo suave  $\phi: G \rightarrow \text{GL}_n(\mathbb{R})$  (as matrizes invertíveis  $n \times n$ ). Isto é, vale  $\phi(ab) = \phi(a)\phi(b)$ .

**Definição:** Duas representações  $\phi_1, \phi_2: G \rightarrow \text{GL}_n(\mathbb{R})$  são **equivalentes** se existir  $A \in \text{GL}_n(\mathbb{R})$  tal que  $\phi_2 = A\phi_1A^{-1}$ .

Elas são “iguais a menos de mudança de coordenadas”.

**Proposição:** As representações de  $\text{SO}(2)$  em  $\mathbb{R}^{2n}$  são classificadas por  $\mathbb{Z}^n / \mathfrak{S}_n$  ( $n$ -uplas a menos de permutação). Mais precisamente, a  $(\omega_1, \dots, \omega_n) \in \mathbb{Z}^n$  é associada  $\phi_{(\omega_1, \dots, \omega_n)}: \text{SO}(2) \rightarrow \text{GL}_{2n}(\mathbb{R})$ .

$$\phi_{(\omega_1, \dots, \omega_n)}(\theta) = \begin{pmatrix} R(\omega_1\theta) & & & \\ & R(\omega_2\theta) & & \\ & & \ddots & \\ & & & R(\omega_n\theta) \end{pmatrix} \quad \text{where} \quad R(\theta) = \begin{pmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{pmatrix}$$

**Proposição:** As representações de  $T^2$  em  $\mathbb{R}^{2n}$  são classificadas por  $(\mathbb{Z}^n)^2 / \mathfrak{S}_n$  (matrizes  $2 \times n$  a menos de permutação das colunas).

As representações são estudadas por meio de **representações irredutíveis**.

**Definição:** Seja  $G \rightarrow GL_n(\mathbb{R})$  uma representação de  $G$  e  $x_0 \in \mathbb{R}^n$  um ponto. A **órbita** de  $x_0$  sob a ação de  $G$  é  $\mathcal{O} = \{\phi(g)x_0 \mid g \in G\}$ .

**Exemplo:** As órbitas de  $SO(2)$  são “círculos”.

- $SO(2) \rightarrow GL_2(\mathbb{R})$

$$\theta \mapsto \begin{pmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{pmatrix}$$

A órbita de  $(1, 0)$  é  $\mathcal{O} = \left\{ \begin{pmatrix} \cos \theta \\ \text{sen } \theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$

- $SO(2) \rightarrow GL_4(\mathbb{R})$

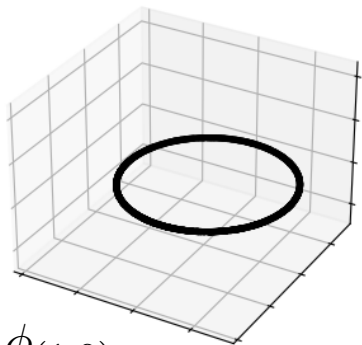
$$\theta \mapsto \begin{pmatrix} \cos \theta & -\text{sen } \theta & 0 & 0 \\ \text{sen } \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

A órbita de  $(1, 0, 1, 0)$  é  $\mathcal{O} = \left\{ \begin{pmatrix} \cos \theta \\ \text{sen } \theta \\ 1 \\ 0 \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$

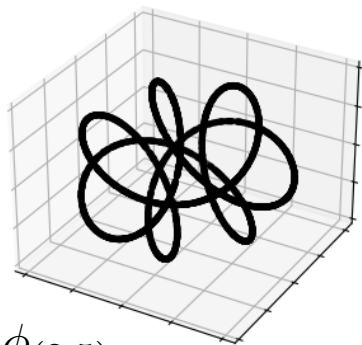
- $SO(2) \rightarrow GL_4(\mathbb{R})$

$$\theta \mapsto \begin{pmatrix} \cos 2\theta & -\text{sen } 2\theta & 0 & 0 \\ \text{sen } 2\theta & \cos 2\theta & 0 & 0 \\ 0 & 0 & \cos 5\theta & -\text{sen } 5\theta \\ 0 & 0 & \text{sen } 5\theta & \cos 5\theta \end{pmatrix}$$

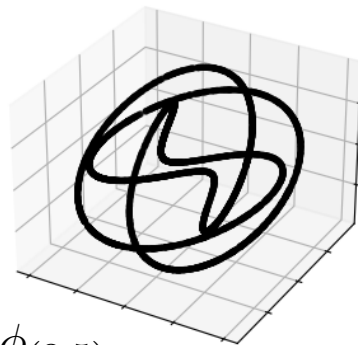
A órbita de  $(1, 0, 1, 0)$  é  $\mathcal{O} = \left\{ \begin{pmatrix} \cos 2\theta \\ \text{sen } 2\theta \\ \cos 5\theta \\ \text{sen } 5\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$



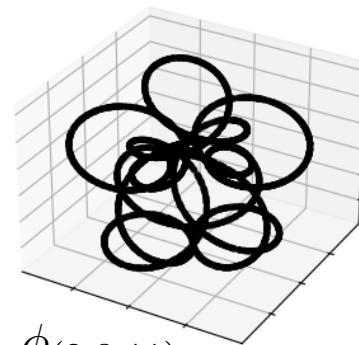
$\phi(1,0)$



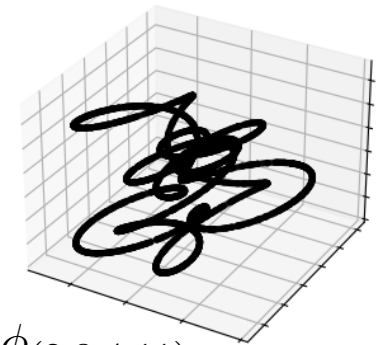
$\phi(2,5)$



$\phi(3,5)$



$\phi(2,3,11)$



$\phi(2,3,4,11)$

**Definição:** Seja  $G \rightarrow GL_n(\mathbb{R})$  uma representação de  $G$  e  $x_0 \in \mathbb{R}^n$  um ponto. A **órbita** de  $x_0$  sob a ação de  $G$  é  $\mathcal{O} = \{\phi(g)x_0 \mid g \in G\}$ .

**Exemplo:** As órbitas de  $SO(2)$  são “círculos”.

**Exemplo:** As órbitas de  $T^2$  são “toros”.

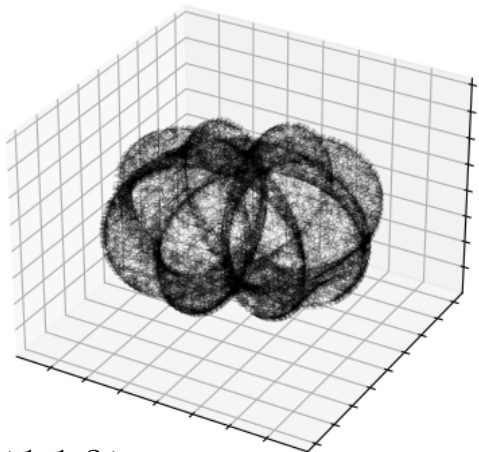
•  $T^2 \longrightarrow GL_6(\mathbb{R})$

$$\theta \mapsto \begin{pmatrix} \cos \theta & -\text{sen } \theta & 0 & 0 & 0 & 0 \\ \text{sen } \theta & \cos \theta & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos \theta & -\text{sen } \theta & 0 & 0 \\ 0 & 0 & \text{sen } \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos 3\theta & -\text{sen } 3\theta \\ 0 & 0 & 0 & 0 & \text{sen } 3\theta & \cos 3\theta \end{pmatrix}$$

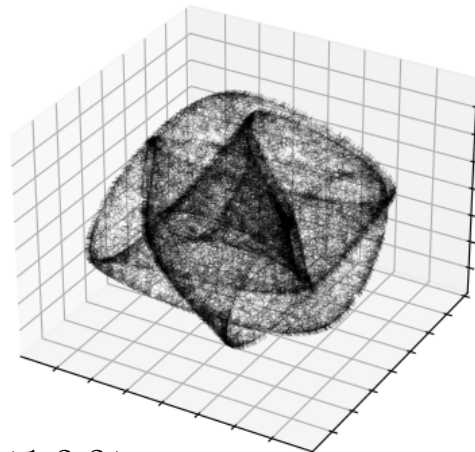
$$\mu \mapsto \begin{pmatrix} \cos \mu & -\text{sen } \mu & 0 & 0 & 0 & 0 \\ \text{sen } \mu & \cos \mu & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos 2\mu & -\text{sen } 2\mu & 0 & 0 \\ 0 & 0 & \text{sen } 2\mu & \cos 2\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos \mu & -\text{sen } \mu \\ 0 & 0 & 0 & 0 & \text{sen } \mu & \cos \mu \end{pmatrix}$$

A órbita de  $(1, 0, 1, 0, 1, 0)$  é

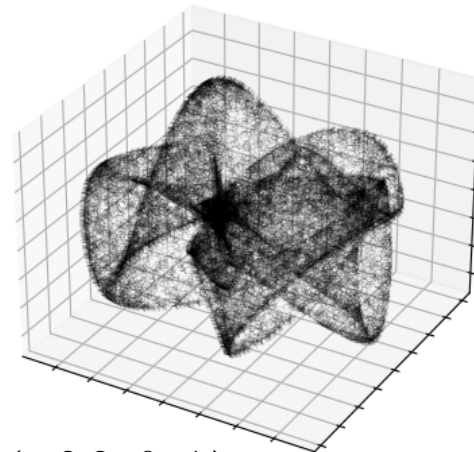
$$\mathcal{O} = \left\{ \begin{pmatrix} \cos \theta + \cos \mu \\ \text{sen } \theta + \text{sen } \mu \\ \cos \theta + \cos 2\mu \\ \text{sen } \theta + \text{sen } 2\mu \\ \cos 3\theta + \cos \mu \\ \text{sen } 3\theta + \text{sen } \mu \end{pmatrix} \mid (\theta, \mu) \in \mathbb{R}^2 \right\}$$



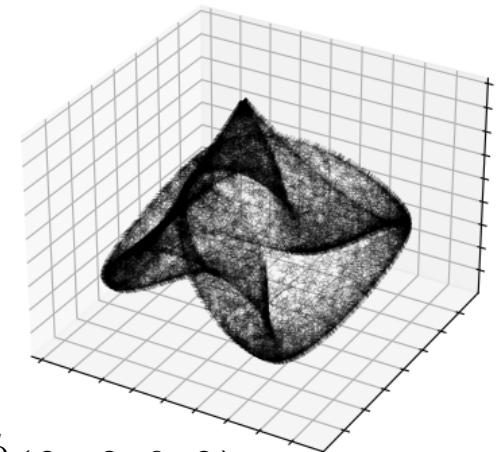
$$\phi \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \end{pmatrix}$$



$$\phi \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$



$$\phi \begin{pmatrix} -2 & 2 & 0 & 1 \\ -1 & 0 & -2 & 1 \end{pmatrix}$$



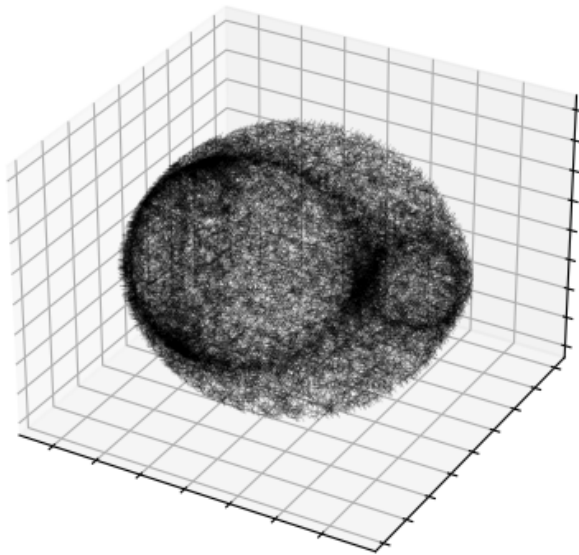
$$\phi \begin{pmatrix} 2 & -2 & 0 & 2 \\ 1 & 1 & -1 & 2 \end{pmatrix}$$

**Definição:** Seja  $G \rightarrow GL_n(\mathbb{R})$  uma representação de  $G$  e  $x_0 \in \mathbb{R}^n$  um ponto. A **órbita** de  $x_0$  sob a ação de  $G$  é  $\mathcal{O} = \{\phi(g)x_0 \mid g \in G\}$ .

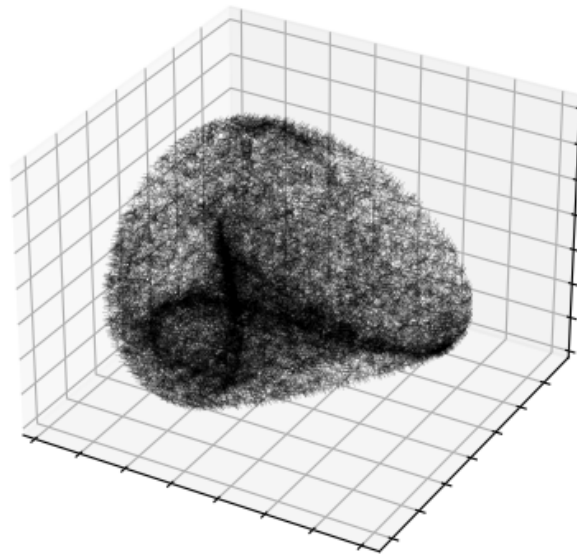
**Exemplo:** As órbitas de  $SO(2)$  são “círculos”.

**Exemplo:** As órbitas de  $T^2$  são “toros”.

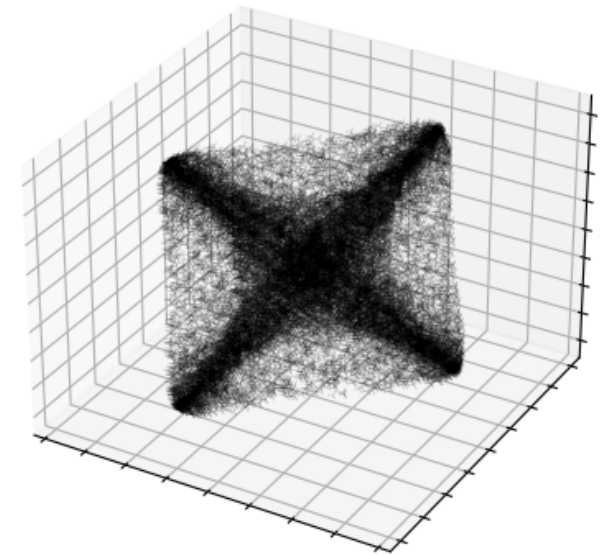
**Exemplo:** As órbitas de  $SO(3)$  e  $SU(2)$  são “esferas”.



$\psi_{(5)}$  in  $\mathbb{R}^5$



$\psi_{(3,4)}$  in  $\mathbb{R}^7$



$\psi_{(8)}$  in  $\mathbb{R}^8$

1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
6. Aplicação: Espaço de conformação do ciclooctano
7. Garantias teóricas

Descrição das trajetórias de três objetos, sujeitos apenas à influência da força gravitacional mútua.

## Sistema arbitrário:

[https://upload.wikimedia.org/wikipedia/commons/1/1c/Three-body\\_Problem\\_Animation\\_with\\_COM.gif](https://upload.wikimedia.org/wikipedia/commons/1/1c/Three-body_Problem_Animation_with_COM.gif)

## Sistema periódico:

[https://upload.wikimedia.org/wikipedia/commons/5/5a/5\\_4\\_800\\_36\\_downscaled.gif](https://upload.wikimedia.org/wikipedia/commons/5/5a/5_4_800_36_downscaled.gif)

(descobertas por Roger Broucke em 1975)

Soluções periódicas em forma fechada foram descobertas por Euler em 1740 e Lagrange em 1772.

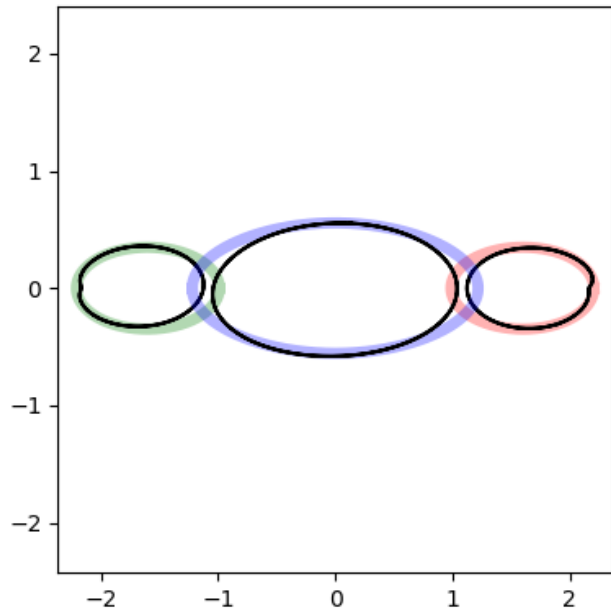
De acordo com um teorema de Poincaré (1889), não existe uma solução geral em forma fechada.

**Nossa tentativa:** dadas as trajetórias  $x_1, x_2, x_3: \mathbb{R} \rightarrow \mathbb{R}^2$  dos três corpos, forma a trajetórias total

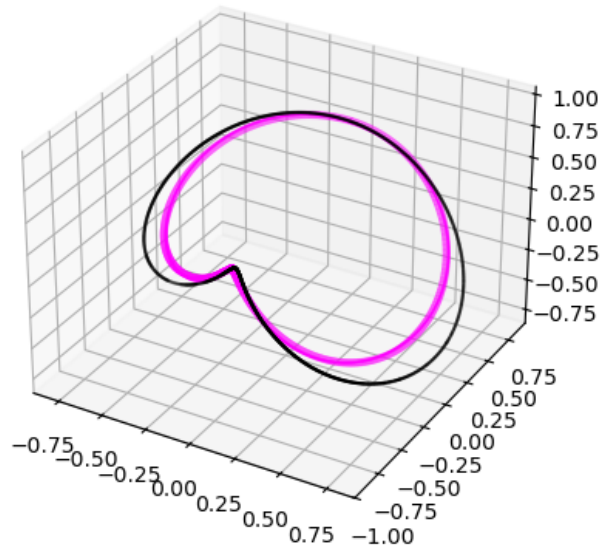
$$(x_1, x_2, x_3): \mathbb{R} \rightarrow \mathbb{R}^6$$

e vê se é uma órbita de uma representação de  $SO(2)$ .

Trajatórias dos corpos em  $\mathbb{R}^2$



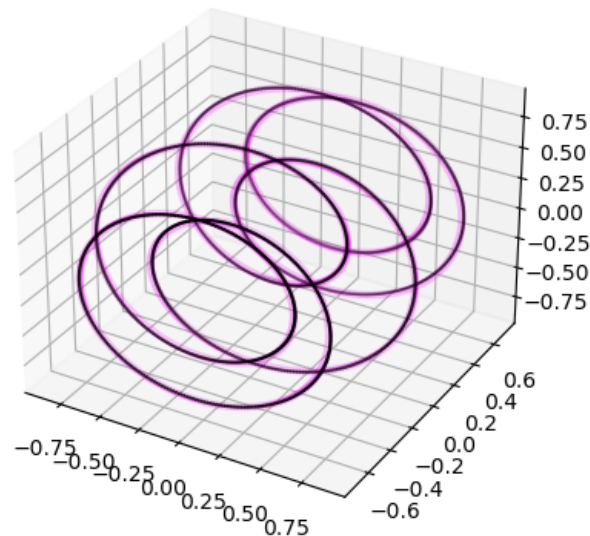
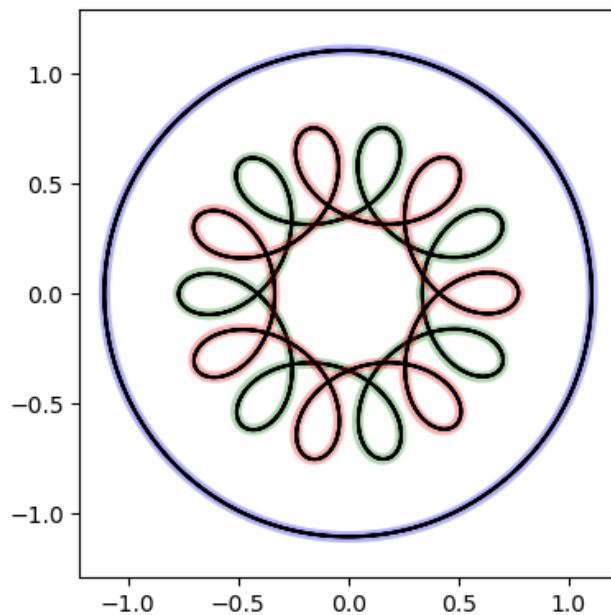
Trajatória total em  $\mathbb{R}^6$



Frequências da representação  
e distância de Hausdorff

(0, 1, 2)

0.22507

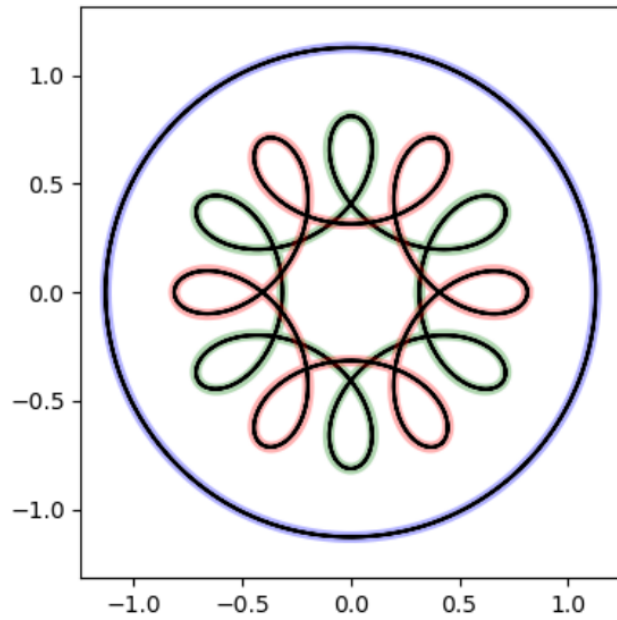


(0, 1, 6)

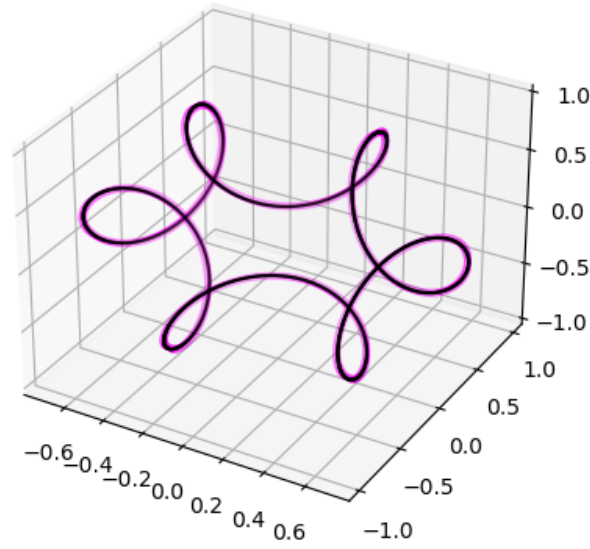
0.01176



Trajетórias dos corpos em  $\mathbb{R}^2$



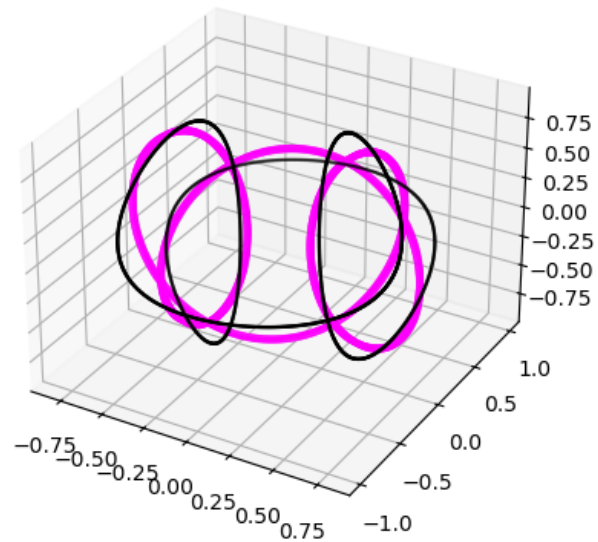
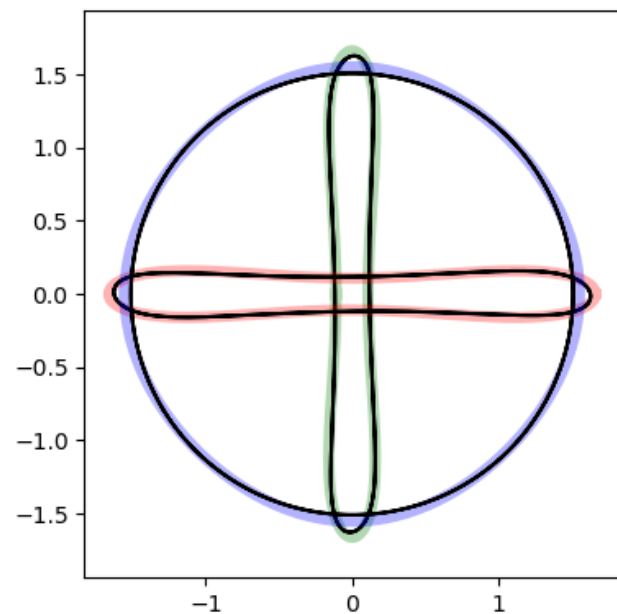
Trajетória total em  $\mathbb{R}^6$



Frequências da representação e distância de Hausdorff

(0, 1, 5)

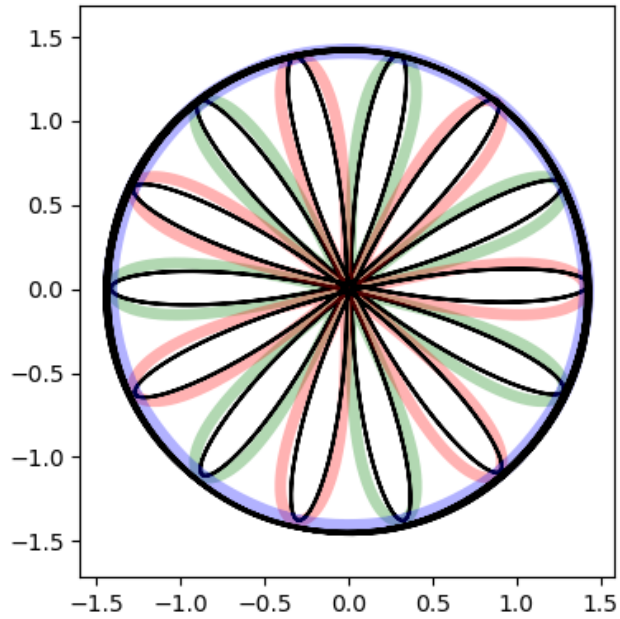
0.01323



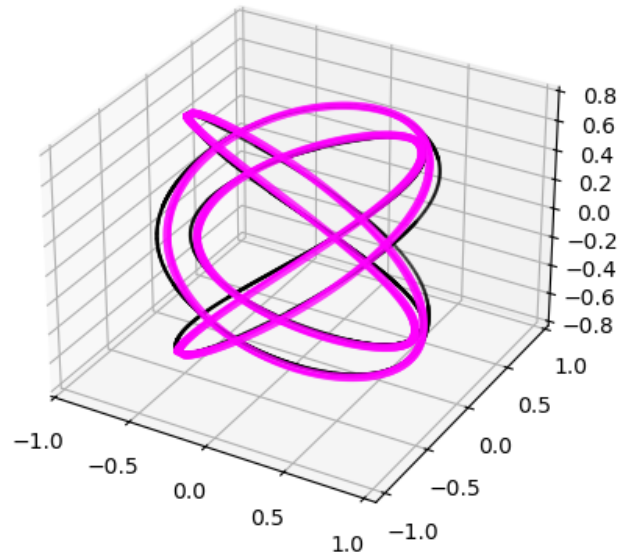
(0, 1, 3)

0.1767

Trajatórias dos corpos em  $\mathbb{R}^2$



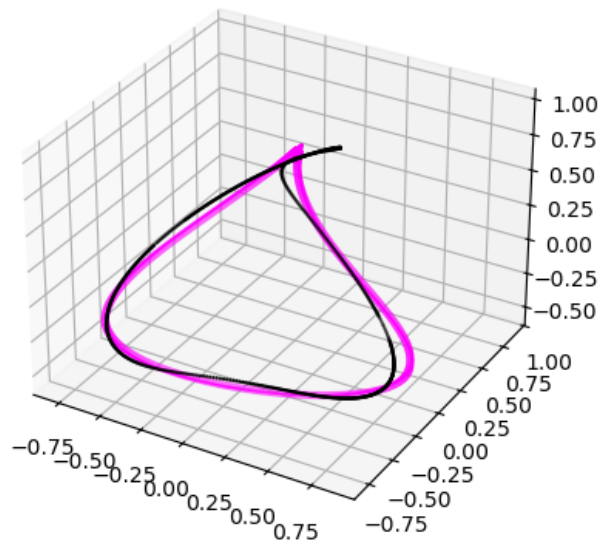
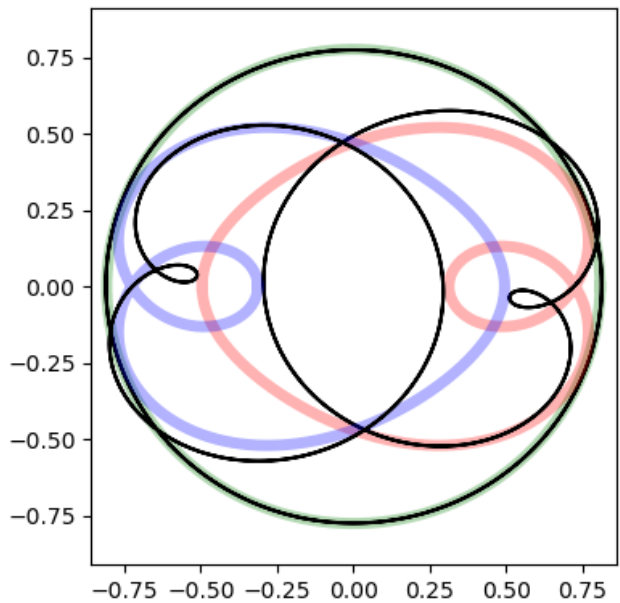
Trajatória total em  $\mathbb{R}^6$



Frequências da representação  
e distância de Hausdorff

$(0, 3, 4)$

0.06893



$(0, 1, 2)$

0.4521

1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
6. Aplicação: Espaço de conformação do ciclooctano
7. Garantias teóricas

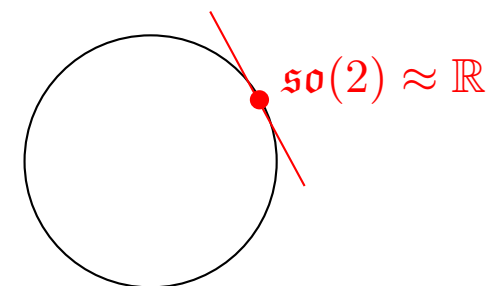
Seja  $G$  um grupo de Lie,  $0 \in G$  o elemento identidade e  $\mathfrak{g} = T_0G$  o **espaço tangente**.

Existe uma **aplicação exponencial**, denotada  $\exp: \mathfrak{g} \rightarrow G$ . Ela é suave. Quando  $G$  é conexo e compacto, ela é sobrejetora.

**Exemplo:** Para grupos de matrizes, o mapa exponencial é simplesmente a exponencial de matrizes.

$$\bullet \text{SO}(2) = \left\{ \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \mid t \in \mathbb{R} \right\} \xleftarrow{\exp} \mathfrak{so}(2) = \left\{ \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

$$\text{Temos } \exp \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$



$$\bullet \text{SO}(3) = \{A \in \text{GL}_3(\mathbb{R}) \mid A^T = A^{-1}, \det A = 1\} \xleftarrow{\exp} \mathfrak{so}(3) = \langle X_1, X_2, X_3 \rangle, \text{ onde}$$

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

**Observação:** Qualquer grupo de Lie compacto admite uma métrica Riemanniana (bi-invariante) para a qual a exponencial de Lie e a exponencial de Riemann coincidem.

**Definição:** Uma **álgebra de Lie** é um espaço vetorial  $\mathfrak{g}$  munido de um mapa bilinear  $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  que satisfaz a identidade de Jacobi.

É denotado por  $[A, B]$ , onde  $A, B \in \mathfrak{g}$ , e chamado de **colchete de Lie**.

**Proposição:** O espaço tangente  $\mathfrak{g} = T_0G$  de um grupo de Lie é uma álgebra de Lie.

**Exemplo:** No caso de grupos de matrizes, o colchete de Lie é simplesmente o comutador

$$[A, B] = AB - BA.$$

Por exemplo, em  $SO(3)$ , temos  $[X_1, X_2] = X_3$ ,  $[X_2, X_3] = X_1$  e  $[X_1, X_3] = -X_2$ , onde

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad X_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

A álgebra de Lie contém muitas informações sobre o grupo de Lie.

**Proposição:** Para grupos de Lie simplesmente conexos  $G_1$  e  $G_2$ , tem-se  $\mathfrak{g}_1 \simeq \mathfrak{g}_2 \implies G_1 \simeq G_2$ .

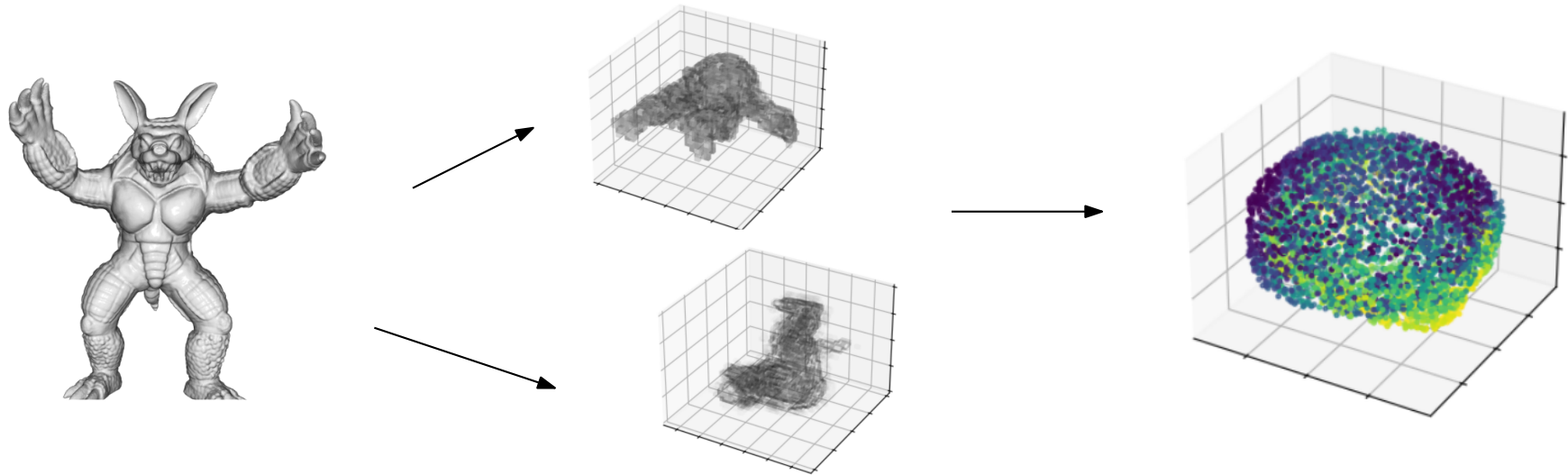
As álgebras de Lie permitem estudar representações de um ponto de vista infinitesimal.

**Proposição:** Dada uma representação  $\phi: G \rightarrow GL_n(\mathbb{R})$ , existe um morfismo  $d\phi: \mathfrak{g} \rightarrow \mathfrak{gl}_n(\mathbb{R})$  de álgebras de Lie, chamado **representação derivada**, de modo que o seguinte diagrama comuta:

$$\begin{array}{ccc}
 G & \xrightarrow{\phi} & GL_n(\mathbb{R}) \\
 \uparrow \text{exp} & & \uparrow \text{exp} \\
 \mathfrak{g} & \xrightarrow{d\phi} & \mathfrak{gl}_n(\mathbb{R}) = \text{matrizes } n \times n
 \end{array}$$

**Definição:** A imagem  $d\phi(\mathfrak{g}) \subset \mathfrak{gl}_n(\mathbb{R})$  é chamada de **álgebra de Lie imagem**. Ela desempenhará um papel fundamental em nosso problema.

1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
6. Aplicação: Espaço de conformação do ciclooctano
7. Garantias teóricas

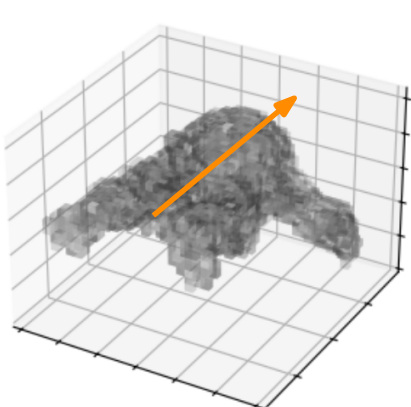


(1) Escolha uma imagem no  $\mathbb{R}^3$ ,  
i.e.,  $A \subset \mathbb{R}^{n \times n \times n}$ .

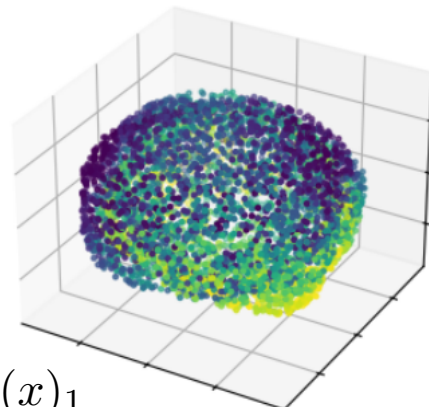
(2) Gere várias rotações dela,  
a fim de construir uma nuvem  
de pontos  $X \subset \mathbb{R}^{n \times n \times n}$ .

(3) Projete  $X$  no  $\mathbb{R}^n$  por meio  
de PCA.

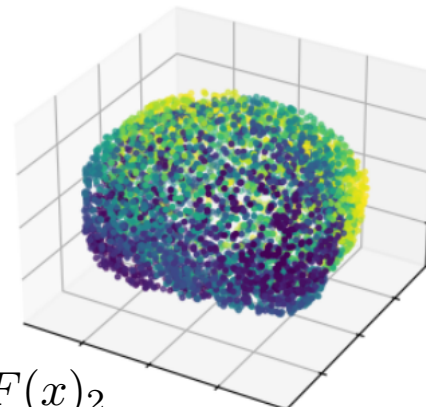
**Um problema de Machine Learning:** dado um  $x \in X$ , estima o vetor unitário  $F(x) \in \mathbb{R}^3$  que aponta para a cabeça do tatu.



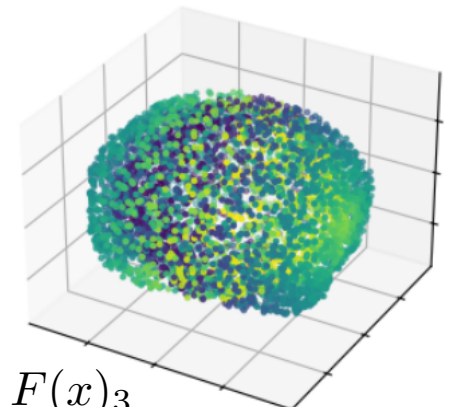
$F(x)_1$



$F(x)_2$



$F(x)_3$



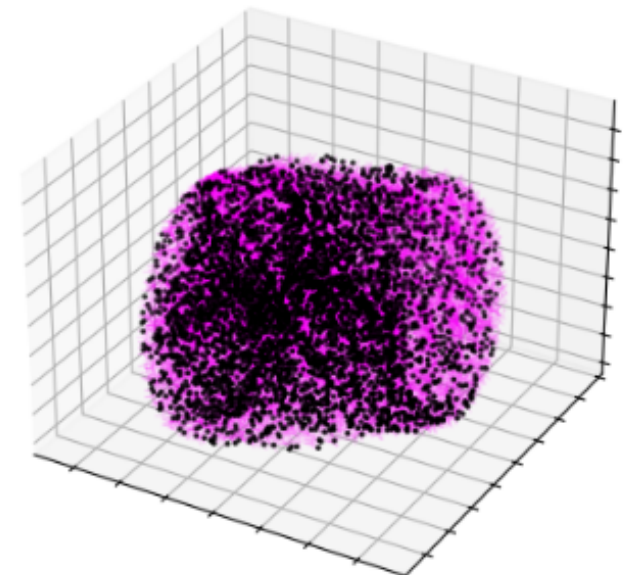


**Solução clássica:** reduz a dimensão de  $X$  por meio de PCA e treina um SVM.

Model	MSE on test data
SVM in dimension 3	0.4003
SVM in dimension 4	0.2496
SVM in dimension 5	0.1295
SVM in dimension 6	0.0380
SVM in dimension 7	0.0148
SVM in dimension 8	0.0119
SVM in dimension 9	0.0114
SVM in dimension 10	0.0122
SVM on orthogonal coordinates	<b>0.0066</b>

**Solução usando grupos de Lie:** Nosso algoritmo detecta uma órbita de representação de  $SO(3)$  em  $\mathbb{R}^8$  que está próxima a  $X$ .

Distância de Hausdorff:  $d_H(X, \mathcal{O}) \simeq 0.1909$ .

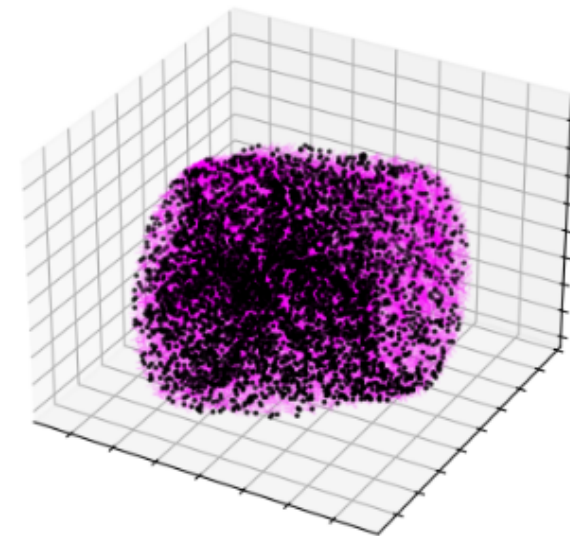


**Coordenadas ortogonais:** Vimos que, no  $\mathbb{R}^8$ ,  $X$  é bem aproximado por uma órbita  $\mathcal{O}$  de uma representação  $\phi$  de  $SO(3)$ . A órbita pode ser escrita como

$$\mathcal{O} = \{\phi(g) \cdot x_0 \mid g \in G\}.$$

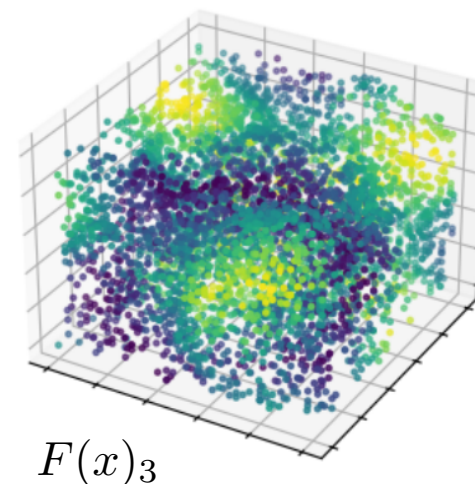
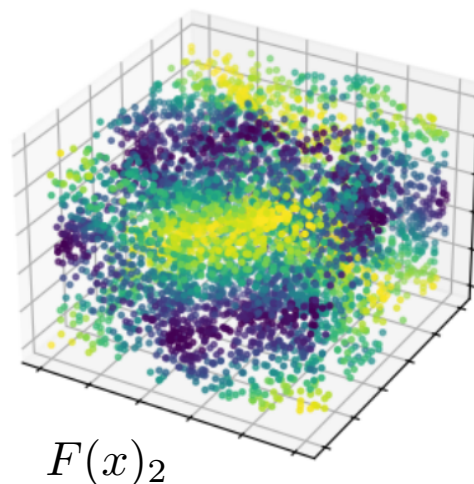
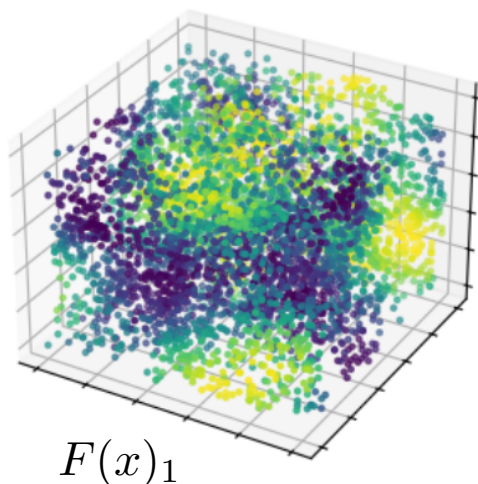
Aém disso, lembre-se do mapa exponencial:

$$\begin{array}{ccc} G & \xrightarrow{\phi} & GL_n(\mathbb{R}) \\ \uparrow \text{exp} & & \uparrow \text{exp} \\ \mathfrak{g} & \xrightarrow{d\phi} & \mathfrak{gl}_n(\mathbb{R}) \end{array}$$



Cada imagem de  $x \in \mathbb{R}^8$  pode ser associada a um elemento de  $c \in \mathfrak{g}$ , por meio de

$$\min_{c \in \mathfrak{so}(3)} \|x - \phi(\exp(c)) \cdot x_0\|.$$



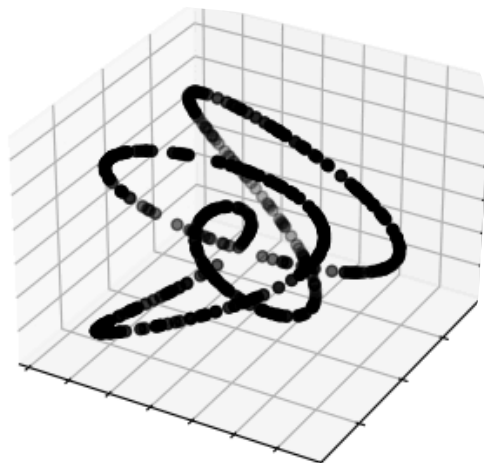
1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
- 5. Descrição do algoritmo**
6. Aplicação: Espaço de conformação do ciclooctano
7. Garantias teóricas

# Formulação do problema: ponto de vista infinitesimal 18/28 (1/3)

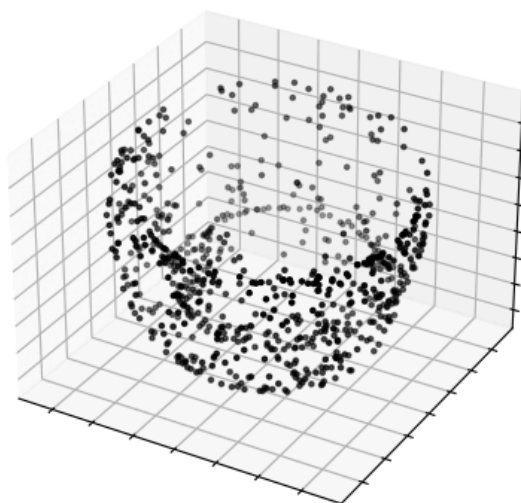
**Input:** Uma nuvem de pontos  $X = \{x_1 \dots, x_N\} \subset \mathbb{R}^n$ .

**Output:** Uma representação  $\phi$  de grupo de Lie compacto  $G$  em  $\mathbb{R}^n$  e uma órbita  $\mathcal{O}$  próxima a  $X$ .

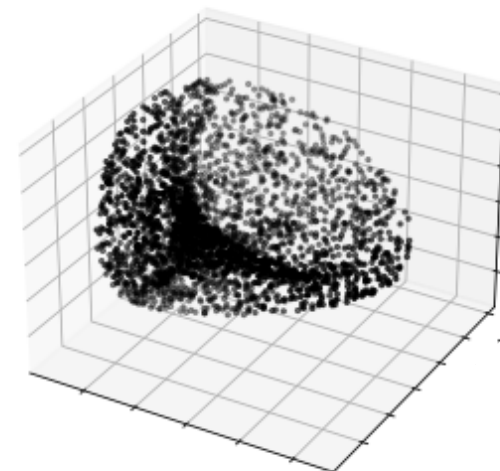
Órbita de  $SO(2)$  no  $\mathbb{R}^6$



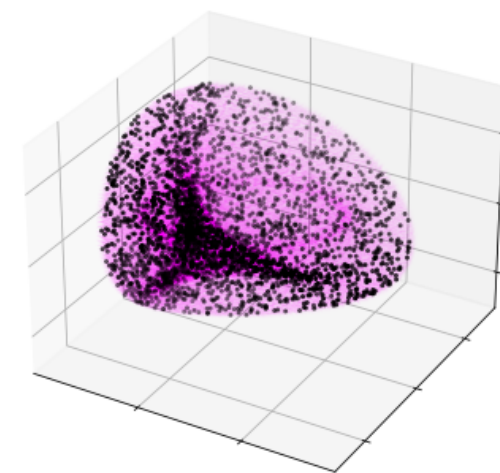
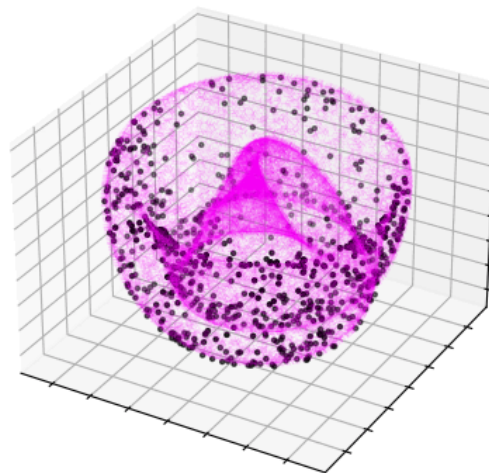
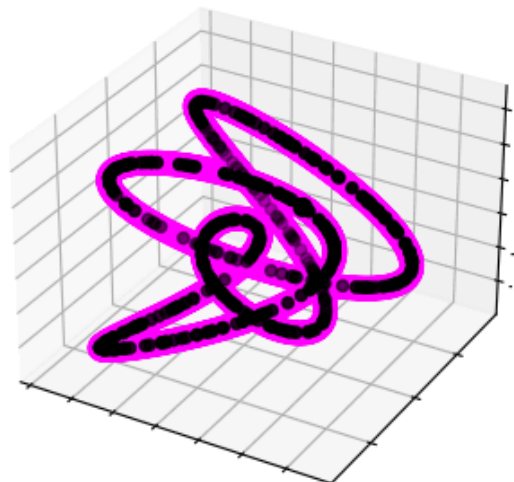
Órbita de  $T^2$  no  $\mathbb{R}^6$



Órbita de  $SO(3)$  no  $\mathbb{R}^9$



**Input:**



**Output:**

# Formulação do problema: ponto de vista infinitesimal 18/28 (2/3)

**Input:** Uma nuvem de pontos  $X = \{x_1 \dots, x_N\} \subset \mathbb{R}^n$ .

**Output:** Uma representação  $\phi$  de grupo de Lie compacto  $G$  em  $\mathbb{R}^n$  e uma órbita  $\mathcal{O}$  próxima a  $X$ .

**Idea:** Obter a melhor órbita  $\mathcal{O}$  minimizando o erro quadrático médio.

**Problema:** Não está claro como calcular a projeção de  $X$  em  $\mathcal{O}$ .

**Outra ideia:** Lembre-se da representação derivada

$$\begin{array}{ccccccc}
 G & \xrightarrow{\phi} & \phi(G) \subset \text{Sym}(\mathcal{O}) & \subset & \text{GL}_n(\mathbb{R}) \\
 \uparrow \text{exp} & & \uparrow \text{exp} & & \uparrow \text{exp} & & \uparrow \text{exp} \\
 \mathfrak{g} & \xrightarrow{d\phi} & \mathfrak{h} \subset \mathfrak{sym}(\mathcal{O}) & \subset & \mathfrak{gl}_n(\mathbb{R})
 \end{array}$$

Em vez de estimar  $\phi$ , busque a álgebra imagem  $\mathfrak{h} = d\phi(\mathfrak{g})$ .  $\mathcal{O}$  é obtido exponenciando-a:

$$\mathcal{O} = \phi(G) \cdot x = \exp(\mathfrak{h}) \cdot x = \{ \exp(A)x \mid A \in \mathfrak{h} \} \quad \text{onde } x \in \mathcal{O} \text{ é arbitrário.}$$

**Grupo de simetria:**

$$\text{Sym}(\mathcal{O}) = \{ P \in \text{GL}_n(\mathbb{R}) \mid P\mathcal{O} = \mathcal{O} \}$$

**Álgebra de simetria:**

$$\mathfrak{sym}(\mathcal{O}) = \{ P \in \mathfrak{gl}_n(\mathbb{R}) \mid \exp(P) \in \text{Sym}(\mathcal{O}) \}$$

**Hipótese:** Supomos que  $\mathfrak{h} = \mathfrak{sym}(\mathcal{O})$ .

**Lie-PCA** é um algoritmo desenvolvido recentemente, que permite estimar  $\mathfrak{h}$  a partir de  $X$ .

O resultado, denotado  $\hat{\mathfrak{h}} = \text{span}\{\hat{\mathfrak{h}}_1, \dots, \hat{\mathfrak{h}}_d\}$ , é um subespaço linear de  $\mathfrak{gl}_n(\mathbb{R})$  de dimensão  $d$ .

**Operador Lie-PCA:**  $\Lambda: M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$  é definido como

$$\Lambda(A) = \frac{1}{N} \sum_{1 \leq i \leq N} \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle]$$

onde

- $\hat{\Pi}[N_{x_i} X]$  são estimativas de matrizes de projeção sobre os espaços normais  $N_{x_i} \mathcal{O}$ ,
- $\Pi[\langle x_i \rangle]$  são as matrizes de projeção nas linhas  $\langle x_i \rangle$ .

Definimos  $\hat{\mathfrak{h}}$  como o subespaço gerado pelos primeiros autovetores  $\hat{\mathfrak{h}}_1, \dots, \hat{\mathfrak{h}}_d$  de  $\Lambda$ .

**Operador Lie-PCA:**  $\Lambda: M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$  é definido como

$$\Lambda(A) = \frac{1}{N} \sum_{1 \leq i \leq N} \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle]$$

onde

- $\hat{\Pi}[N_{x_i} X]$  são estimativas de matrizes de projeção sobre os espaços normais  $N_{x_i} \mathcal{O}$ ,
- $\Pi[\langle x_i \rangle]$  são as matrizes de projeção nas linhas  $\langle x_i \rangle$ .

Definimos  $\hat{\mathfrak{h}}$  como o subespaço gerado pelos primeiros autovetores  $\hat{\mathfrak{h}}_1, \dots, \hat{\mathfrak{h}}_d$  de  $\Lambda$ .

**Derivação de Lie-PCA:** Baseia-se no fato de que  $\mathfrak{sym}(\mathcal{O}) = \{A \in M_n(\mathbb{R}) \mid \forall x \in \mathcal{O}, Ax \in T_x \mathcal{O}\}$ , onde  $T_x \mathcal{O}$  denota o espaço tangente de  $\mathcal{O}$  em  $x$ . Em outras palavras,

$$\mathfrak{sym}(\mathcal{O}) = \bigcap_{x \in \mathcal{O}} S_x \mathcal{O} \quad \text{onde} \quad S_x \mathcal{O} = \{A \in M_n(\mathbb{R}) \mid Ax \in T_x \mathcal{O}\},$$

Usando apenas a nuvem de pontos  $X = \{x_1, \dots, x_N\}$ , consideramos

$$\bigcap_{i=1}^N S_{x_i} \mathcal{O} = \ker \left( \sum_{i=1}^N \Pi[(S_{x_i} \mathcal{O})^\perp] \right),$$

Além disso, os autores mostram que  $\Pi[(S_{x_i} \mathcal{O})^\perp](A) = \Pi[N_{x_i} \mathcal{O}] \cdot A \cdot \Pi[\langle x_i \rangle]$ . Podemos então

$$\Lambda(A) = \frac{1}{N} \sum_{i=1}^N \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle]$$

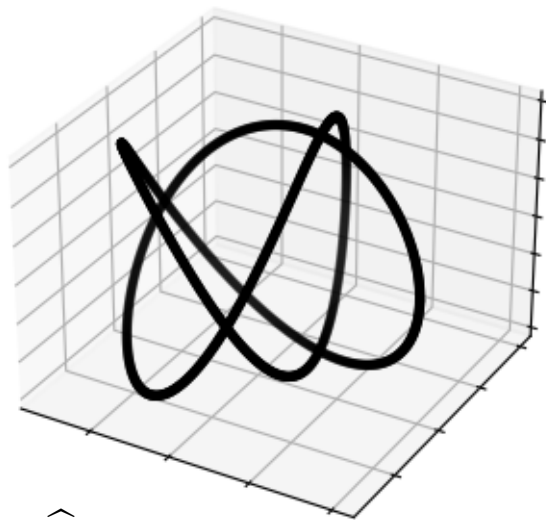
onde  $\hat{\Pi}[N_{x_i} X]$  é uma estimativa de  $\Pi[N_{x_i} \mathcal{O}]$  calculada a partir da observação  $X$ .



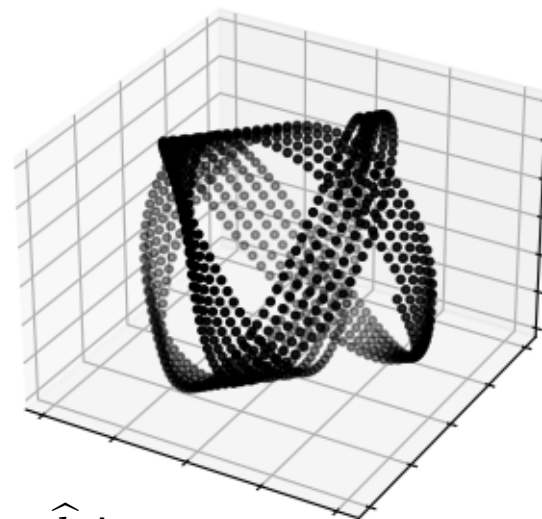
Via Lie-PCA, obtemos  $\hat{\mathfrak{h}}$ , um subespaço linear de  $M_n(\mathbb{R})$  de dimensão  $d$ . É uma estimativa de  $\mathfrak{h}$ .

**Problema:** O subespaço  $\hat{\mathfrak{h}}$  é estimado como se fosse um subespaço linear. Ele pode não ser uma álgebra de Lie (para  $A, B \in \hat{\mathfrak{h}}$ , devemos ter  $AB - BA \in \hat{\mathfrak{h}}$ ).

Exponenciar uma álgebra que não é de Lie, ou uma álgebra de Lie que não vem de um grupo compacto, pode gerar grandes erros.



$\hat{\mathfrak{h}}$  exato



$\hat{\mathfrak{h}}$  inexato

As a consequence of the Poincaré separation theorem, the orthogonal projection onto the span of any  $\ell$  of the bottom eigenvectors of  $\Sigma$  gives an optimizer of this program. In particular, the span of any such eigenvectors gives a worthy estimate for  $\mathfrak{sym}(M)$ . One may be inclined to round this estimate to the nearest member of  $L(\ell, d)$ , but we do not attempt this here.

Desejamos projetar  $\widehat{\mathfrak{h}}$  na álgebra de Lie mais próxima. Trabalhamos em  $\mathfrak{so}(n)$ , as matrizes anti-simétricas  $n \times n$ . Ele tem dimensão  $n(n+1)/2$ . Ele é dotado do produto interno e norma

$$\langle A, B \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} b_{i,j} \quad \text{and} \quad \|A\| = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2}.$$

## Variedade de Stiefel de álgebras de Lie

Veja os subespaços  $d$ -dimensionais de  $\mathfrak{so}(n)$  como matrizes  $n(n-1)/2 \times d$ .

$\mathcal{V}^{\text{Lie}}(d, \mathfrak{so}(n))$  é definido como o **conjunto de  $d$ -uplas** de vetores  $(A_1, \dots, A_d)$  de  $\mathfrak{so}(n)$  normalizados e ortogonais, com a condição de álgebra de Lie:  $\forall i, j \in [1, \dots, n], A_i A_j - A_j A_i \in \langle A_1, \dots, A_d \rangle$ .

O problema é

$$\min \left\{ \sum_{i=1}^d \|\widehat{\mathfrak{h}}_i - A_i\|^2 \mid (A_1, \dots, A_d) \in \mathcal{V}^{\text{Lie}}(d, \mathfrak{so}(n)) \right\}$$

## Variedade de Grassmann de álgebras de Lie

Veja os subespaços  $d$ -dimensionais de  $\mathfrak{so}(n)$  como matrizes  $n(n-1)/2 \times n(n-1)/2$ .

$\mathcal{G}^{\text{Lie}}(d, \mathfrak{so}(n))$  é definido como o **conjunto de matrizes de projeção ortogonal** de posto  $d$  em  $\mathfrak{so}(n)$  com a condição de álgebra de Lie:  $\forall i, j \in [1, \dots, n], P(Pe_i \cdot Pe_j - Pe_j \cdot Pe_i) = Pe_i \cdot Pe_j - Pe_j \cdot Pe_i$  em que  $(e_1, \dots, e_{n(n+1)/2})$  é uma base ortonormal de  $\mathfrak{so}(n)$ .

O problema é

$$\min \left\{ \|\text{proj}[\widehat{\mathfrak{h}}] - P\| \mid P \in \mathcal{G}^{\text{Lie}}(d, \mathfrak{so}(n)) \right\}$$

Escrito explicitamente em forma matricial,

## Variedade de Stiefel de álgebras de Lie

$$\min \sum_{i=1}^d \|\widehat{\mathfrak{h}}_i - A_i\|^2 \text{ tal que } \begin{cases} \forall i \in [1 \dots, d], & A_i \text{ é uma matriz } (n \times n), \\ \forall i \in [1 \dots, d], & A_i^\top = -A_i, \\ \forall i, j \in [1 \dots, d], & \sum_{k=1}^d \langle A_k, A_i A_j - A_j A_i \rangle^2 = \|A_i A_j - A_j A_i\|^2. \end{cases}$$

## Variedade de Grassmann de álgebras de Lie

$$\min \|\text{proj}[\widehat{\mathfrak{h}}] - P\| \text{ tal que } \begin{cases} P \text{ é uma matriz } (n(n+1)/2 \times n(n+1)/2), \\ P^2 = P, \\ P^\top = P, \\ \text{rank}(P) = d, \\ \forall i, j \in [1 \dots, d], & P(Pe_i \cdot Pe_j - Pe_j \cdot Pe_i) = Pe_i \cdot Pe_j - Pe_j \cdot Pe_i. \end{cases}$$

- Problema:**
- (1) Esses programas parecem intratáveis (contêm a classificação das álgebras de Lie),
  - (2) Uma álgebra de Lie em  $\mathfrak{so}(n)$  pode nem mesmo vir de um grupo de Lie compacto.

**Idea:** Fixe um grupo de Lie compacto  $G$  e restrinja a variedade de Stiefel  $\mathcal{V}^{\text{Lie}}(d, \mathfrak{so}(n))$  e de Grassmannian  $\mathcal{G}^{\text{Lie}}(d, \mathfrak{so}(n))$  às álgebras de Lie que são imagens de  $G$ .

Doravante,  $G$  é um grupo de Lie compacto fixo de dimensão  $d$ .

## Variedade de Stiefel de álgebras imagem de $G$

$\mathcal{V}(G, \mathfrak{so}(n))$  é definido como o conjunto de  $(A_1, \dots, A_d) \in \mathcal{V}^{\text{Lie}}(d, \mathfrak{so}(n))$  para os quais existe uma representação ortogonal  $\phi: G \rightarrow \text{SO}(n)$  de modo que  $d\phi(\mathfrak{g})$  seja gerado por  $(A_1, \dots, A_d)$ .

**Lemma:** As componentes conexas de  $\mathcal{V}(G, \mathfrak{so}(n))$  são em correspondência com as **classes de equivalência de órbita** de representações de  $G$  no  $\mathbb{R}^n$ .

**Definição:** Dizemos que duas representações  $\phi, \phi': G \rightarrow \text{GL}_n(\mathbb{R})$  são **órbita-equivalentes** se existir uma matriz  $M \in \text{M}_n(\mathbb{R})$  tal que  $d\phi(\mathfrak{g}) = Md\phi'(\mathfrak{g})M^{-1}$ . Isto é, suas órbitas são conjugadas. Denotaremos por  $\text{orb}(G, n)$  um conjunto de representantes das classes de equivalência de órbita.

## Variedade de Grassmann de álgebras imagem de $G$

$\mathcal{G}(G, \mathfrak{so}(n))$  é definido como o conjunto de elementos  $P \in \mathcal{G}^{\text{Lie}}(d, \mathfrak{so}(n))$  para os quais existe uma representação ortogonal  $\phi: G \rightarrow \text{SO}(n)$  de modo que  $P$  é a matriz de projeção em  $d\phi(\mathfrak{g})$ .

**Lemma:** As componentes conexas de  $\mathcal{G}(G, \mathfrak{so}(n))$  também são em correspondência com as classes de equivalência de órbita de representações de  $G$  em  $\mathbb{R}^n$ .

Doravante,  $G$  é um grupo de Lie compacto fixo de dimensão  $d$ .

## Variedade de Stiefel de álgebras imagem de $G$

$\mathcal{V}(G, \mathfrak{so}(n))$  é definido como o conjunto de  $(A_1, \dots, A_d) \in \mathcal{V}^{\text{Lie}}(d, \mathfrak{so}(n))$  para os quais existe uma representação ortogonal  $\phi: G \rightarrow \text{SO}(n)$  de modo que  $d\phi(\mathfrak{g})$  seja gerado por  $(A_1, \dots, A_d)$ .

**Lemma:** As componentes conexas de  $\mathcal{V}(G, \mathfrak{so}(n))$  são em correspondência com as **classes de equivalência de órbita** de representações de  $G$  no  $\mathbb{R}^n$ .

**Definição:** Dizemos que duas representações  $\phi, \phi': G \rightarrow \text{GL}_n(\mathbb{R})$  são **órbita-equivalentes** se existir uma matriz  $M \in \text{M}_n(\mathbb{R})$  tal que  $d\phi(\mathfrak{g}) = Md\phi'(\mathfrak{g})M^{-1}$ . Isto é, suas órbitas são conjugadas. Denotaremos por  $\text{orb}(G, n)$  um conjunto de representantes das classes de equivalência de órbita.

**Lemma:** Para qualquer  $(A_1, \dots, A_d) \in \mathcal{V}(G, \mathfrak{so}(n))$ , existe  $p \geq 1$ , uma  $p$ -upla  $(B^1, \dots, B^p) \in \text{orb}(G, n)$  e duas matrizes  $O \in \text{O}(n)$  e  $P \in \text{O}(d)$  tais que, para todos os  $i \in [1 \dots d]$ ,

$$A_i = \sum_{j=1}^d P_{j,i} O \text{diag}(B_j^k)_{k=1}^p O^\top.$$

Em particular, o subespaço  $\langle A_1, \dots, A_d \rangle \subset \mathfrak{so}(n)$  é gerado pelas matrizes

$$O \text{diag}(B_1^k)_{k=1}^p O^\top, \quad O \text{diag}(B_2^k)_{k=1}^p O^\top, \quad \dots, \quad O \text{diag}(B_p^k)_{k=1}^p O^\top.$$

**Corolário:** O problema  $\min \left\{ \|\text{proj}[\widehat{\mathfrak{h}}] - P\| \mid P \in \mathcal{G}(G, \mathfrak{so}(n)) \right\}$  é equivalente a

$$\min \left\| \text{proj}[\widehat{\mathfrak{h}}] - \text{proj}[\langle O \text{diag}(B_i^k)_{k=1}^p O^\top \rangle_{i=1}^d] \right\| \quad \text{tal que} \quad \begin{cases} (B^1, \dots, B^p) \in \text{orb}(G, n), \\ O \in O(n). \end{cases}$$

**Observação:** Esse é um problema discreto-contínuo.

Se decompõe em  $N$  problemas de minimização sobre  $O(n)$ , onde  $N$  é o cardinal de  $\text{orb}(G, n)$ .

Na prática, efetuamos uma descida de gradiente em  $O(n)$  (line-search com retração QR).

**Observação:** Para aplicar esse resultado na prática, é preciso ter acesso a uma descrição explícita de  $\text{orb}(G, n)$ . Resolvemos os casos de  $SO(2)$ ,  $T^d$ ,  $SO(3)$  e  $SU(2)$ .

**Input:** Uma nuvem de pontos  $X = \{x_1 \dots, x_N\} \subset \mathbb{R}^n$ .

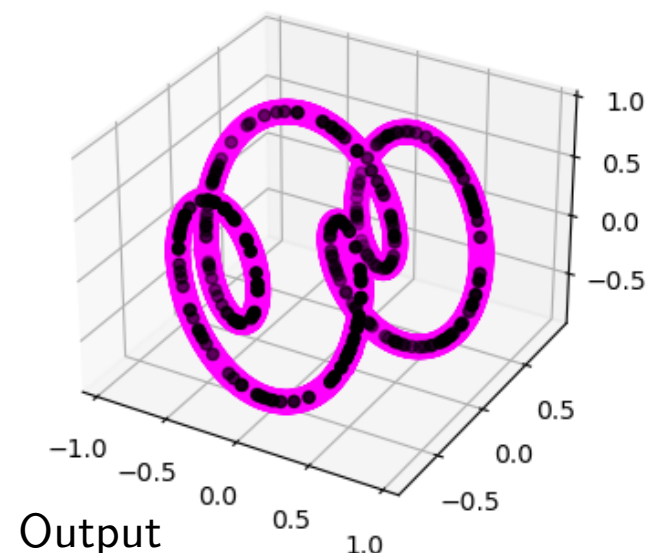
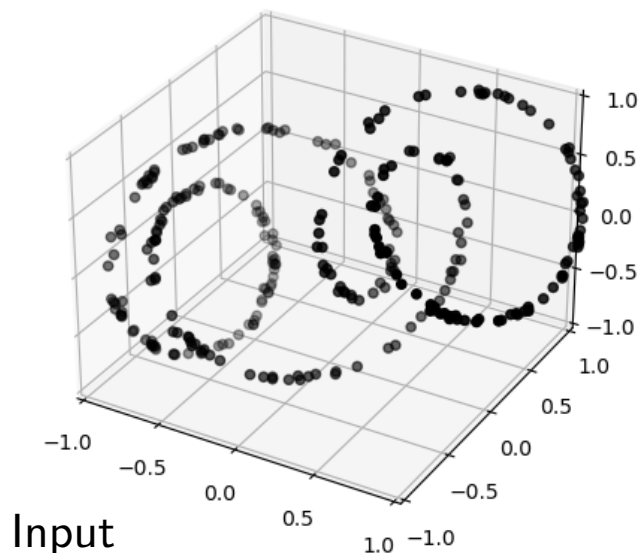
**Output:** Uma representação **ortogonal**  $\hat{\phi}$  de um grupo de Lie compacto  $G$  em  $\mathbb{R}^n$  e uma órbita  $\hat{\mathcal{O}}$  próxima a  $X$ .

**Etapa 1: Ortonormalização** Aplique redução de dimensão e ortonormalização.

**Etapa 2: Lie-PCA** Diagonalize o operador Lie-PCA  $\Lambda: M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$ .

**Etapa 3: Álgebra de Lie mais próxima** Estime  $\hat{\mathfrak{h}}$  por meio de uma otimização sobre  $O(n)$ .

**Etapa 4: Geração da órbita** Construa  $\hat{\mathcal{O}}_x = \exp(\hat{\mathfrak{h}}) \cdot x$  e verifique que ela esteja próxima a  $X$ .

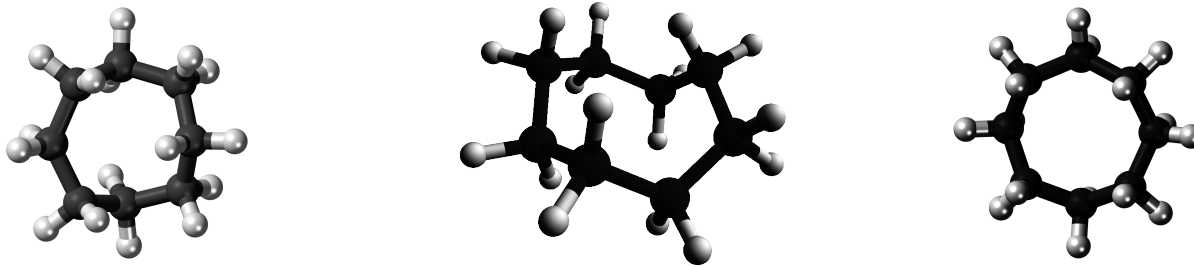


1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
- 6. Aplicação: Espaço de conformação do ciclooctano**
7. Garantias teóricas



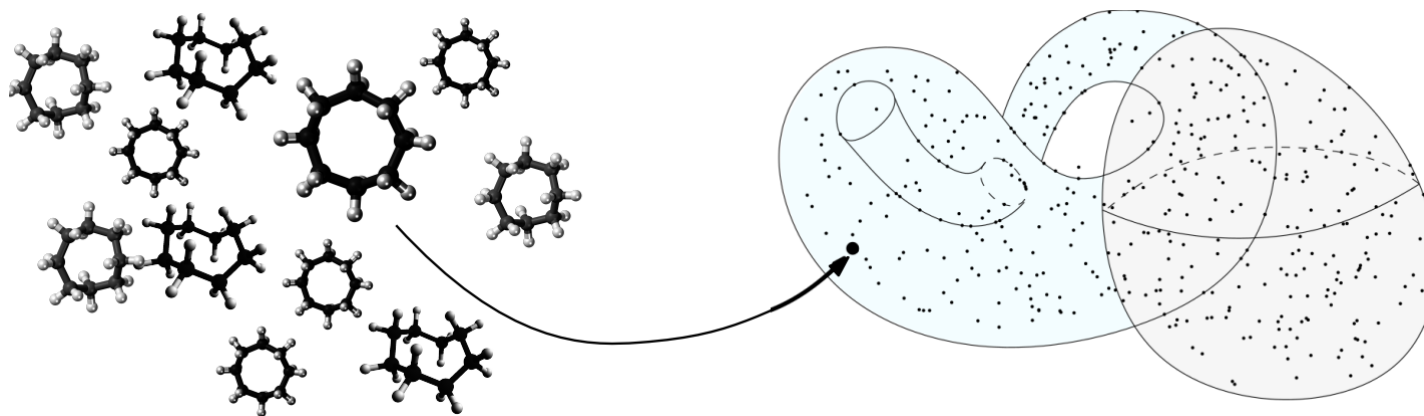
Toda molécula admite muitas conformações possíveis, chamadas de **confômeros**.

A molécula de ciclooctano  $C_8H_{16}$  contém 24 átomos. Cada átomo tem 3 coordenadas espaciais. Portanto, um confômero pode ser resumido por um ponto em  $\mathbb{R}^{72}$  ( $3 \times 24 = 72$ ).



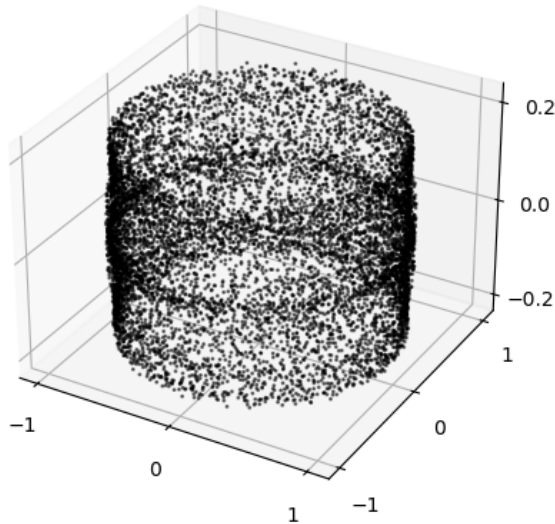
Ao considerar muitos confômeros, obtém-se uma nuvem de pontos, denotada  $X \subset \mathbb{R}^{72}$ . Ela está próxima à *união entre uma garrafa de Klein e uma esfera*.

[Martin, Thompson, Coutsiás & Watson, [Topology of cyclo-octane energy landscape](#), 2010]



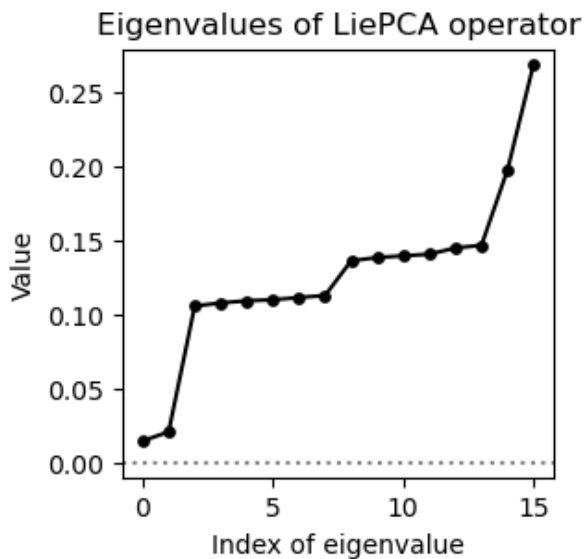
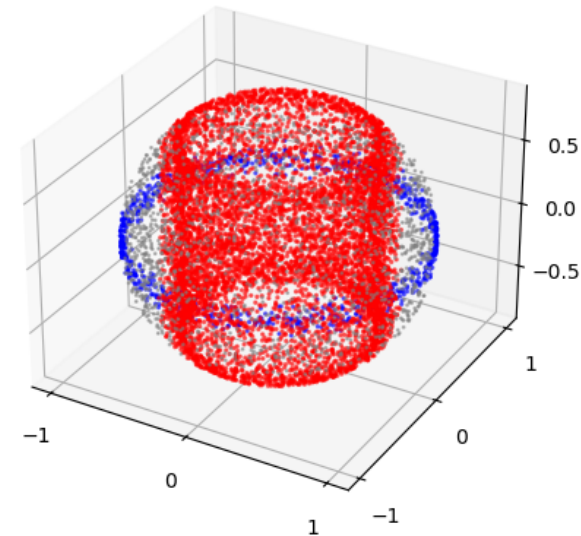
**Nossa tentativa:** verifique se  $X$  não é, na verdade, uma órbita de um grupo de Lie.

**Confômeros não alinhados:** Em nosso primeiro experimento, geramos 10.000 confômeros de ciclooctano, sem alinhá-los.



Projetado na dimensão 3, vemos um cilindro cercado por um círculo.

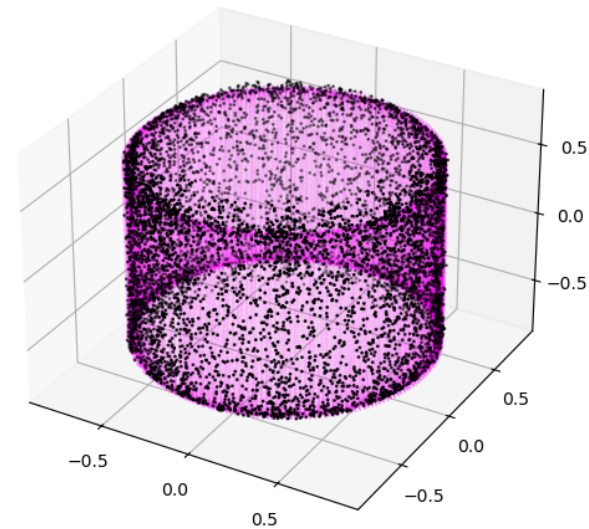
$X$  é projetado no  $\mathbb{R}^4$  e ortonormalizado. Após descartar 15% dos outliers (cinza), dois clusters aparecem. Ficamos com o vermelho.



Lie-PCA apresenta dois autovalores pequenos, sugerindo um grupo de simetria de dimensão 2.

LieDetect encontra uma órbita de uma representação de  $T^2$  no  $\mathbb{R}^4$  que coincide com os dados.

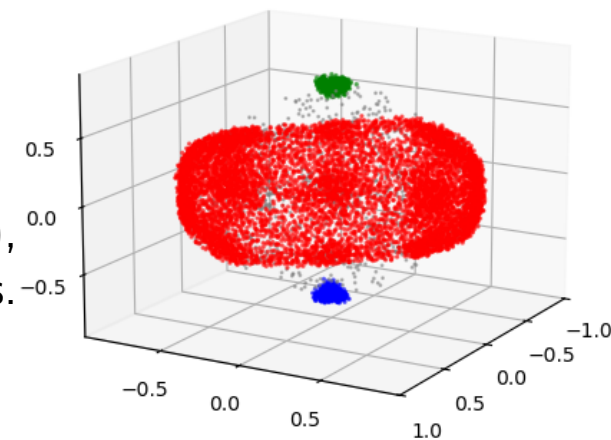
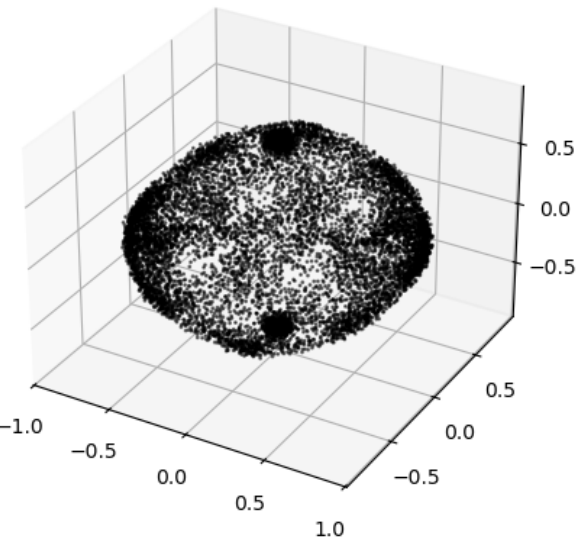
Distância de Hausdorff  $d_H(X, \mathcal{O}) \simeq 0.2$ .



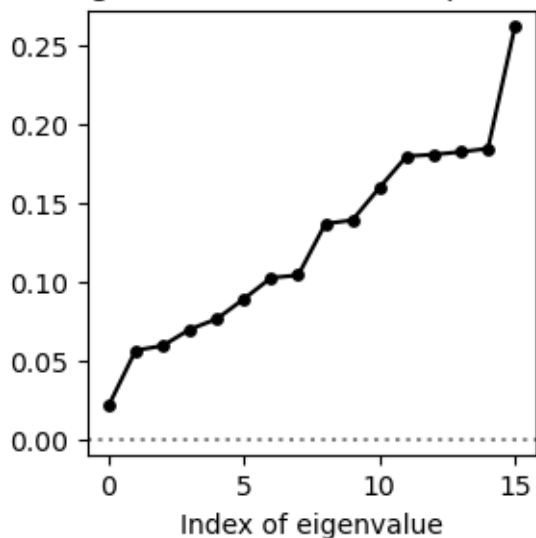
**Confômeros alinhados:** Agora, geramos 10.000 confômeros alinhados.

Projetado na dimensão 3, vemos três componentes: uma superfície e dois clusters.

Após descartar 10% dos outliers (cinza), os pontos são agrupados em três classes. Ficamos com o vermelho.

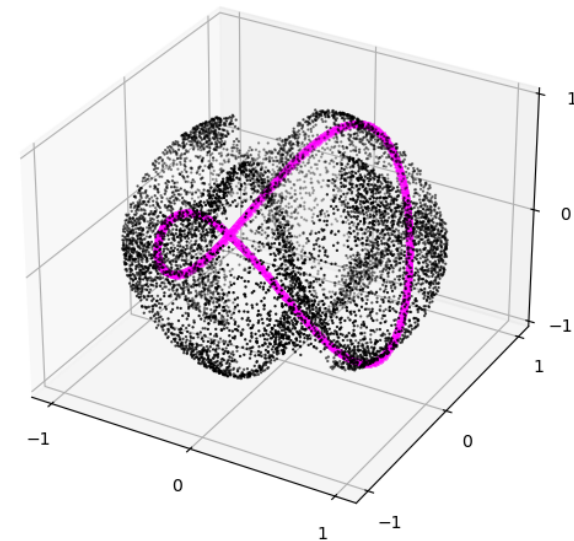


Eigenvalues of LiePCA operator



Lie-PCA apresenta um autovalor pequeno, sugerindo um grupo de simetria de dimensão 1.

LieDetect encontra uma ação de  $SO(2)$  que estabiliza a nuvem de pontos. Distância média:  $d_H(X, \mathcal{O}_x) \simeq 0.1$ .



1. Representações e órbitas
2. Aplicação: Problema de três corpos
3. Correspondência grupo de Lie - álgebra de Lie
4. Aplicação: Rotação de corpos rígidos
5. Descrição do algoritmo
6. Aplicação: Espaço de conformação do ciclooctano
- 7. Garantias teóricas**

**Input:**  $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^n$  e  $G$  grupo de Lie compacto

**Modelo:**  $X$  amostrado próximo a uma órbita  $\mathcal{O}$  de uma representação  $\phi: G \rightarrow \mathbb{R}^n$

**Etapa 1: Ortonormalização** via  $X \leftarrow \sqrt{\Sigma[X]^+} \cdot \Pi_{\Sigma[X]}^{>\epsilon} \cdot X$ .

com  $\Sigma[X]$  matriz de covariância e  $\Pi_{\Sigma[X]}^{>\epsilon}$  matriz de projeção nos eixos  $> \epsilon$ .

**Etapa 2: Diagonalize** o operador  $\Lambda: A \mapsto \frac{1}{N} \sum_{i=1}^N \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle]$

com  $A \in M_n(\mathbb{R})$  e  $\hat{\Pi}[N_{x_i} X]$  estimativa da projeção no espaço normal de  $X$ .

**Etapa 3: Resolve**  $\arg \min_{\hat{h}} \|\Pi[\langle A_i \rangle_{i=1}^d] - \Pi[\hat{h}]\|$  com  $(A_i)_{i=1}^d$  primeiros autovetores de  $\Lambda$

com  $\hat{h} \in \mathcal{G}(\mathfrak{g}, \mathfrak{so}(n))$  variedade de Grassmann das subálgebras de Lie imagem de  $G$ .

**Etapa 4: Construa**  $\hat{\mathcal{O}}_x = \{\exp(A)x \mid A \in \hat{h}\}$

onde  $x \in X$  é um ponto arbitrário.

**Objetivo:** Mostre que  $\hat{\mathcal{O}}_x$  está próximo de  $\mathcal{O}$

**Input:**  $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^n$  e  $G$  grupo de Lie compacto

$\mu$  medida no  $\mathbb{R}^n$ . E.g.,  $\mu_X$  medida empírica sobre  $X$

**Modelo:**  $X$  amostrado próximo a uma órbita  $\mathcal{O}$  de uma representação  $\phi: G \rightarrow \mathbb{R}^n$

$\mu_{\mathcal{O}}$  medida uniforme sobre  $\mathcal{O}$

**Etapa 1: Ortonormalização** via  $X \leftarrow \sqrt{\Sigma[X]^+} \cdot \Pi_{\Sigma[X]}^{\geq \epsilon} \cdot X$ .

$$\mu \leftarrow \sqrt{\Sigma[\mu]^+} \cdot \Pi_{\Sigma[\mu]}^{\geq \epsilon} \cdot \mu.$$

**Etapa 2: Diagonalize** o operador  $\Lambda: A \mapsto \frac{1}{N} \sum_{i=1}^N \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle]$

$$\Lambda[\mu]: A \mapsto \int_{i=1}^N \hat{\Pi}[N_{x_i} X] \cdot A \cdot \Pi[\langle x_i \rangle] d\mu$$

**Etapa 3: Resolve**  $\arg \min_{\hat{h}} \|\Pi[\langle A_i \rangle_{i=1}^d] - \Pi[\hat{h}]\|$  com  $(A_i)_{i=1}^d$  primeiros autovetores de  $\Lambda$

$$\arg \min_{\hat{h}} \|\Pi[\langle A_i \rangle_{i=1}^d] - \Pi[\hat{h}]\| \text{ com } (A_i)_{i=1}^d \text{ primeiro autovetores de } \Lambda[\mu]$$

**Etapa 4: Construa**  $\hat{\mathcal{O}}_x = \{ \exp(A)x \mid A \in \hat{h} \}$

$$\mu_{\hat{\mathcal{O}}_x} = \exp(\hat{h}) \cdot \mu$$

**Objetivo:** Mostre que  $\hat{\mathcal{O}}_x$  está próximo de  $\mathcal{O}$

Mostre que  $W_2(\mu_{\hat{\mathcal{O}}_x}, \mu_{\mathcal{O}}) \leq W_2(\mu, \mu_{\mathcal{O}})$

Por que trabalhar com Wasserstein e não com Hausdorff?

- Formalismo natural para grupos de Lie (média com a medida de Haar)
- Possibilita ruído e pontos anômalos
- O PCA local não é estável em Hausdorff

**Observação:** Nosso objetivo é obter uma desigualdade explícita  $W_2(\mu_{\hat{\theta}_x}, \mu_{\mathcal{O}}) \leq W_2(\mu, \mu_{\mathcal{O}})$ . Isso é diferente de outros formalismos estatísticos. Em particular, não há lei dos grandes números/concentração.

**Teorema:** Seja  $G$  um grupo de Lie compacto de dimensão  $d$ ,  $\mathcal{O}$  uma órbita de uma representação quase fiel  $\phi: G \rightarrow \mathbb{R}^n$ , potencialmente não ortogonal, e  $l$  sua dimensão. Seja  $\mu_{\mathcal{O}}$  a medida uniforme sobre  $\mathcal{O}$  e  $\mu_{\tilde{\mathcal{O}}}$  a medida sobre a órbita ortonormalizada.

Além disso, seja  $X \subset \mathbb{R}^n$  uma nuvem de pontos finita e  $\mu_X$  a medida empírica. Sejam  $\hat{\phi}$ ,  $\hat{h}$  e  $\mu_{\hat{\mathcal{O}}}$  o resultado do algoritmo. Sob hipóteses técnicas, vale que  $\hat{\phi}$  é equivalente a  $\phi$  e

$$\|\Pi[\hat{h}] - \Pi[\text{sym}(\mathcal{O})]\|_{\text{F}} \leq 9d \frac{\rho}{\lambda} \left( r + 4 \left( \frac{\tilde{\omega}}{r^{l+1}} \right)^{1/2} \right)$$

$$W_2(\mu_{\hat{\mathcal{O}}}, \mu_{\tilde{\mathcal{O}}}) \leq \frac{1}{\sqrt{2}} \frac{W_2(\mu_X, \mu_{\mathcal{O}})}{\sigma_{\min}} + 3\sqrt{dn} \left( \frac{\rho}{\lambda} \right)^{1/2} \left( r + 4 \left( \frac{\tilde{\omega}}{r^{l+1}} \right)^{1/2} \right)^{1/2}$$

onde

- $\rho = \left( 16l(l+2)6^l \right) \frac{\max(\text{vol}(\tilde{\mathcal{O}}), \text{vol}(\tilde{\mathcal{O}})^{-1})}{\min(1, \text{reach}(\tilde{\mathcal{O}}))}$
- $\sigma_{\max}^2, \sigma_{\min}^2$  o primeiro e último autovalor da matriz de covariância  $\Sigma[\mu_{\mathcal{O}}]$
- $\tilde{\omega} = 4(n+1)^{3/2} \left( \frac{\sigma_{\max}^3}{\sigma_{\min}^3} \right) \left( \omega(v+\omega) \right)^{1/2}$  onde  $\omega = \frac{W_2(\mu_{\mathcal{O}}, \mu_X)}{\sigma_{\min}}$  e  $v = \left( \frac{\mathbb{V}[\|\mu_{\mathcal{O}}\|]}{\sigma_{\min}^2} \right)^{1/2}$
- $r$  é o raio da PCA local ( estimação de espaços tangentes)
- $\lambda$  o primeiro autovalor diferente de zero do operador Lie-PCA ideal  $\Lambda_{\mathcal{O}}$



**Teorema:** Seja  $G$  um grupo de Lie compacto de dimensão  $d$ ,  $\mathcal{O}$  uma órbita de uma representação quase fiel  $\phi: G \rightarrow \mathbb{R}^n$ , potencialmente não ortogonal, e  $l$  sua dimensão. Seja  $\mu_{\mathcal{O}}$  a medida uniforme sobre  $\mathcal{O}$  e  $\mu_{\tilde{\mathcal{O}}}$  a medida sobre a órbita ortonormalizada.

Além disso, seja  $X \subset \mathbb{R}^n$  uma nuvem de pontos finita e  $\mu_X$  a medida empírica. Sejam  $\hat{\phi}$ ,  $\hat{h}$  e  $\mu_{\hat{\mathcal{O}}}$  o resultado do algoritmo. Sob hipóteses técnicas, vale que  $\hat{\phi}$  é equivalente a  $\phi$  e

$$\|\Pi[\hat{h}] - \Pi[\text{sym}(\mathcal{O})]\|_F \leq 9d \frac{\rho}{\lambda} \left( r + 4 \left( \frac{\tilde{\omega}}{r^{l+1}} \right)^{1/2} \right)$$

trade-off viés e variância ao estimar espaços tangentes

$$W_2(\mu_{\hat{\mathcal{O}}}, \mu_{\tilde{\mathcal{O}}}) \leq \frac{1}{\sqrt{2}} \frac{W_2(\mu_X, \mu_{\mathcal{O}})}{\sigma_{\min}} + 3\sqrt{dn} \left( \frac{\rho}{\lambda} \right)^{1/2} \left( r + 4 \left( \frac{\tilde{\omega}}{r^{l+1}} \right)^{1/2} \right)^{1/2}$$

$$\lesssim W_2(\mu, \mu_{\mathcal{O}})^{1/4(l+3)}$$

onde

- $\rho = \left( 16l(l+2)6^l \right) \frac{\max(\text{vol}(\tilde{\mathcal{O}}), \text{vol}(\tilde{\mathcal{O}})^{-1})}{\min(1, \text{reach}(\tilde{\mathcal{O}}))}$
- $\sigma_{\max}^2, \sigma_{\min}^2$  o primeiro e último autovalor da matriz de covariância  $\Sigma[\mu_{\mathcal{O}}]$
- $\tilde{\omega} = 4(n+1)^{3/2} \left( \frac{\sigma_{\max}^3}{\sigma_{\min}^3} \right) \left( \omega(v + \omega) \right)^{1/2}$  onde  $\omega = \frac{W_2(\mu_{\mathcal{O}}, \mu_X)}{\sigma_{\min}}$  e  $v = \left( \frac{\mathbb{V}[\|\mu_{\mathcal{O}}\|]}{\sigma_{\min}^2} \right)^{1/2}$
- $r$  é o raio da PCA local ( estimação de espaços tangentes)
- $\lambda$  o primeiro autovalor diferente de zero do operador Lie-PCA ideal  $\Lambda_{\mathcal{O}}$

**Hipóteses técnicas:** Defina as quantidades

$$\omega = \frac{W_2(\mu_{\mathcal{O}}, \mu_X)}{\sigma_{\min}}, \quad v = \left( \frac{\mathbb{V}[\|\mu_{\mathcal{O}}\|]}{\sigma_{\min}^2} \right)^{1/2},$$

$$\tilde{\omega} = 4(n+1)^{3/2} \left( \frac{\sigma_{\max}^3}{\sigma_{\min}^3} \right) \left( \omega(v+\omega) \right)^{1/2}, \quad \rho = \left( 16l(l+2)6^l \right) \frac{\max(\text{vol}(\tilde{\mathcal{O}}), \text{vol}(\tilde{\mathcal{O}})^{-1})}{\min(1, \text{reach}(\tilde{\mathcal{O}}))},$$

$$\gamma = (4(2d+1)\sqrt{2})^{-1} \cdot \lambda \cdot \Gamma(G, n, \omega_{\max}) \quad (\text{constante de rigidez das subálgebras de Lie})$$

Suponha que  $\omega$  seja suficientemente pequeno, de modo a satisfazer

$$\omega < \left( \left( v^2 + \frac{1}{2} \right)^{1/2} - v \right) / \left( 3(n+1) \frac{\sigma_{\max}^2}{\sigma_{\min}^2} \right), \quad \tilde{\omega} \leq \min \left\{ \left( \frac{1}{6\rho} \right)^{3(l+1)}, \frac{\gamma^{l+3}}{16}, \left( \frac{\gamma}{(6\rho)^2} \right)^{l+1} \right\}.$$

Escolha dois parâmetros  $\epsilon$  e  $r$  nos seguintes conjuntos não vazios:

$$\epsilon \in \left( (2v+\omega)\omega\sigma_{\min}^2, \frac{1}{2}\sigma_{\min}^2 \right], \quad r \in \left[ (6\rho)^2 \cdot \tilde{\omega}^{1/(l+1)}, (6\rho)^{-1} \right] \cap \left[ (4/\gamma)^{2/(l+1)} \cdot \tilde{\omega}^{1/(l+1)}, \gamma \right].$$

Além disso, suponha que

- os problemas de minimização são calculados com exatidão,
- $\mathfrak{sym}(\mathcal{O})$  é gerado por matrizes cujos espectros vêm de vetores integrais primitivos de coordenadas de no máximo  $\omega_{\max}$ ,
- $G = \text{Sym}(\mathcal{O})$ .

# Conclusão

- Primeiro algoritmo para encontrar o **tipo de representação** (não apenas como subespaço linear)
- Implementação para  $G = SO(2)$ ,  $T^d$ ,  $SO(3)$  e  $SU(2)$
- Pode ser adaptado a outro grupo de Lie compacto desde que seja fornecida uma descrição explícita de suas representações
- Experimentos em análise de imagens, análise harmônica, sistemas físicos, aprendizado de máquina e redes neurais equivariantes no <https://github.com/HLovisiEnnes/LieDetect>

Limitações:

- As otimizações sobre  $O(n)$  são computacionalmente caras e instáveis
- O algoritmo não lida com órbitas entrelaçadas
- Restrito a **representações** de grupos de Lie

Próximo objetivo:

- Detecção de **ações** por meio da representação induzida no espaço dos campos vetoriais.

$$\begin{array}{ccc} G & \xrightarrow{\phi} & \text{Diff}(\mathcal{M}) \\ \exp \uparrow & & \uparrow \exp \\ \mathfrak{g} & \xrightarrow{d\phi} & \mathcal{X}(\mathcal{M}) \end{array}$$