Hamiltonian systems over Lie groups appearing in statistical transformation models

Daisuke TARAMA

Ritsumeikan University
E-mail: dtarama [at] fc.ritsumei.ac.jp

Joint project with Jean-Pierre Françoise (LJLL, Sorbonne Univ.)

Submanifold Geometry and Lie Group Actions 2021
21/03/2022

Contents

§1. Introduction

Information geometry and Hamiltonian systems 1

Some brief overview about information geometry and Hamiltonian systems:

- The Fisher-Rao metric, which may go back to early work by Fréchet [F], and the Amari-Chentsov cubic tensor have brought important differential-geometric studies in information geometry. See e.g. [A, AJLS].
- For dually flat manifolds, which can be formulated also as the Hessian geometry initiated by Koszul [K], the α-geodesics play a very important role, as pointed out by Amari.
- Nakamura studied the relations between information geometry and completely integrable Hamiltonian systems e.g. in [Na1, Na2, Na3].

Fisher-Rao metric and Amari-Chentsov tensor

M: manifold of samples, U: manifold of parameters,
$\rho : U \times M \to \mathbb{R}$: family of probability density functions
The Fisher-Rao metric is the (semi-definite) metric
$\sum_{i,j} g_{ij} du^i \otimes du^j$ on U ($(u_1, \ldots, u_n)$: local coordinates), where

$$g_{ij} = E\left[\frac{\partial \log \rho}{\partial u^i}\frac{\partial \log \rho}{\partial u^j}\right] = \int_M \frac{\partial \log \rho(u,x)}{\partial u^i}\frac{\partial \log \rho(u,x)}{\partial u^j}\rho(u,x)\mathrm{dvol}_M(x).$$

$\rightsquigarrow$ cf. Cramer-Rao Theorem
The Amari-Chentsov tensor is the cubic tensor
$\sum_{i,j,k} c_{ijk} du^i \otimes du^j \otimes du^k$ on U, where

$$c_{ijk} = E\left[\frac{\partial \log \rho}{\partial u^i}\frac{\partial \log \rho}{\partial u^j}\frac{\partial \log \rho}{\partial u^k}\right] = \int_M \frac{\partial \log \rho(u,x)}{\partial u^i}\frac{\partial \log \rho(u,x)}{\partial u^j}\frac{\partial \log \rho}{\partial u^k}\rho(u,x)\mathrm{dvol}_M(x).$$

$\rightsquigarrow$ cf. α-connections and dually flat structures.

Hamiltonian systems

(N, ω): symplectic manifold, i.e. ω: non-deg. closed two-form on N.
For a Hamiltonian $H \in \mathcal{C}^\infty(N)$, the Hamiltonian vector field Ξ_H is defined through $\iota_{\Xi_H} = -\mathrm{d}H$.
Poisson bracket $\{\cdot, \cdot\}$ is defined through
$\{F, G\} = -\Xi_F(G) = -\omega(\Xi_F, \Xi_G)$, where $F, G \in \mathcal{C}^\infty(M)$.

More generally, Hamiltonian vector field can be considered for $(N, \{\cdot, \cdot\})$: Poisson manifold.
For a Hamiltonian $H \in \mathcal{C}^\infty(N)$, the Hamiltonian vector field Ξ_H is defined through $\{H, F\} = -\Xi_H(F)$, $F \in \mathcal{C}^\infty(N)$.

Information geometry and Hamiltonian systems 2

The aim of the present talk:

- Hamiltonian systems would bring further interesting perspectives in information geometry. We further investigate the relation between information geometry and Hamiltonian (perhaps completely integrable) systems.
- We focus on the **transformation models** considered by Barndorff-Nielsen and his collaborators [BBE]. (See also the earlier works by Casalis [C1, C2] and the textbook by Amari and Nagaoka [AN].) The models give rise to Hamiltonian systems over **Lie groups** about which many interesting aspects may be discovered.

Information geometry and Hamiltonian systems 3

Plan of the present talk:

§2 The general framework of transformation models is explained.

§3 A specific case with the sample space being the Lie groups is considered.

§2. Transformation models

statistical transformation model

M : smooth manifold of samples, G: Lie group acting smoothly on M from the left:

$$G \times M \ni (g, x) \mapsto g \cdot x \in M.$$

dvol_M: invariant volume element on M w.r.t. the group action. Consider probability density functions $\rho : G \times M \to \mathbb{R}$ on M parameterized by elements in G:

$$\forall g \in G, \qquad \int_M \rho(g, x)\mathrm{dvol}_M(x) = 1.$$

Assume

$$\begin{aligned} &\forall (g, x) \in G \times M, && \rho(g, x) > 0; \\ &\forall g, h \in G, \forall x \in M, && \rho(g \cdot h, x) = \rho(g, h \cdot x). \end{aligned}$$

Expectation of random variables

For $f \in \mathcal{C}^\infty(M)$, consider the expectation w.r.t. ρ:

$$E[f] := \int_M f(x)\rho(g,x)\mathrm{dvol}_M(x)$$

which is a smooth function in $g \in G$. Note that $E[1] = 1$.

$\mathfrak{g}$: Lie alg. of G. For $X \in \mathfrak{g}$, we consider the induced vector field X^M on M which satisfies, at $g \in G$, $x \in M$,

$$X_x^M[\rho(g,x)] = -X_g^{(L)}[\rho(g,x)],$$

where $X_g^{(L)}$: left-invariant vector field on G induced by $X \in \mathfrak{g}$ evaluated at $g \in G$.

Fisher-Rao semi-definite metric 1

We have

$$E\left[X^M[\log \rho]\right] = 0,$$

and

$$E\left[X^M[\log \rho] \cdot Y^M[\log \rho]\right] = -E\left[X^M\left[Y^M[\log \rho]\right]\right]$$
$$= \int_M X_x^M\left[\log \rho(g,x)\right] Y_x^M\left[\log \rho(g,x)\right] \rho(g,x) \mathrm{dvol}_M(x)$$

is constant w.r.t. $g \in G$.

Definition

For $X, Y \in \mathfrak{g}$, the **Fisher-Rao bilinear form** on $\mathfrak{g}$ for the probability density functions ρ is defined through

$$\langle X, Y \rangle = E\left[X^M\left[\log \rho\right] Y^M\left[\log \rho\right]\right] = -E\left[X^M\left[Y^M[\log \rho]\right]\right].$$

Fisher-Rao semi-definite metric 2

Definition

The corresponding left-invariant $(0, 2)$-tensor $\langle \cdot, \cdot \rangle$ on G is called **Fisher-Rao semi-definite metric** w.r.t. the probability density functions ρ.

Note that Fisher-Rao metric is not necessarily definite.

Definition

The left-invariant $(0, 3)$-tensor

$$C(X, Y, Z) = E\left[X_x^M[\log \rho] \cdot Y_x^M[\log \rho] \cdot Z_x^M[\log \rho]\right]$$
$$= \int_M X_x^M[\log\rho(g, x)] \cdot Y_x^M[\log\rho(g, x)] \cdot Z_x^M[\log\rho(g, x)]\rho(g, x)\mathrm{dvol}_M(x)$$

is called **Amari-Chentsov cubic tensor**.

Fisher-Rao semi-definite metric 3

Assume that, for a Lie subgroup $H \subset G$ with $\mathrm{Lie}(H) = \mathfrak{h}$, we have the direct sum decomposition $\mathfrak{g} = \mathfrak{m} \dot{+} \mathfrak{h}$ and that the Fisher-Rao bilinear form $\langle \cdot, \cdot \rangle$ is definite on $\mathfrak{m}$.

If H is compact, G/H is reductive in the sense of Nomizu [No] and Fisher-Rao bilinear form $\langle \cdot, \cdot \rangle$ induces a G-inv. Riemannian metric on G/H. By [No, Theorem 13.1], the Levi-Civita connection ∇ on $(G/H, \langle \cdot, \cdot \rangle)$ is given as

$$\nabla_X Y = \frac{1}{2}[X, Y]_{\mathfrak{m}} + U(X, Y),$$

$$\langle U(X, Y), Z \rangle = \frac{1}{2}\left(\langle [X, Z]_{\mathfrak{m}}, Y \rangle + \langle X, [Y, Z]_{\mathfrak{m}} \rangle\right),$$

where $X, Y, Z \in \mathfrak{m}$, $W_{\mathfrak{m}}$: the $\mathfrak{m}$-component of $W \in \mathfrak{g}$.

Remarks

In the next section, we discuss a specific example for which the Fisher-Rao semi-definite metric on G induces the definite metric on G/H for a compact semi-simple Lie group G and a Cartan subgroup $H \subset G$ with a concrete family of probability density functions ρ.

The α-connection $\nabla^{(\alpha)}$ is defined through

$$\langle \nabla_X^{(\alpha)} Y, Z \rangle = \langle \nabla_X Y, Z \rangle - \frac{\alpha}{2} C(X, Y, Z),$$

where $X, Y, Z \in \mathfrak{m}$.

§3. Case $M = G$ and Euler-Poincaré equation

Case $M = G$

Assumption: $M = G$: compact semi-simple with the Killing form κ, dvol_G: Haar measure, $\mathfrak{h} \subset \mathfrak{g}$: Cartan subalgebra.

Taking into account the cost function considered in [BBR] to describe the generalized Toda lattice equations in double bracket form, consider the probability density functions

$$\rho(g, h) := c{\cdot}\exp\left(F(gh)\right), \quad \text{where}\, \forall\theta \in G, F(\theta) = \kappa\left(Q, \mathrm{Ad}_\theta N\right),$$

where $Q, N \in \mathfrak{g}$: fixed elements.
(c: normalization constant s.t. $\int_G \rho(g, h)\mathrm{dvol}_G(h) = 1$.)
We have

$$\langle X, Y\rangle = -c \cdot \kappa\left([X, Q], [Y, N']\right),$$

where $N' = c \cdot \int_G \mathrm{Ad}_h N \exp(\kappa(Q, \mathrm{Ad}_h N))\mathrm{dvol}_G(h)$.
We assume $Q, N' \in \mathfrak{h}$.

Motivation: the generalized Toda lattice 1

On the normal(split) real form $\mathfrak{g}_n$ of complex semi-simple Lie algebra, the generalized Toda lattice equation is described as

$$\dot{A} = [A, B],$$

where $A = \sum_{j=1}^{r} b_j h_j + \sum_{j=1}^{r} a_j \left(e_{\alpha_j} + e_{-\alpha_j} \right)$,
$B = \sum_{j=1}^{r} a_j \left(e_{\alpha_j} - e_{-\alpha_j} \right)$, $a_j > 0$. Here,
$\{ h_j, e_\alpha \mid j = 1, \cdots, r, \alpha \in \Delta \}$ is a Chevalley basis and
$h_1, \ldots, h_r$ are the simple roots.

Motivation: the generalized Toda lattice 2

By [BBR], it is shown that the generalized Toda lattice equation is equivalent to the double bracket equation

$$\dot{L} = [L, [L, N]],$$

when they are restricted to a special orbits. Here,

$$L = \sum_{j=1}^{r} \sqrt{-1} b_j h_j + \sum_{j=1}^{r} \sqrt{-1} a_j \left(e_{\alpha_j} + e_{-\alpha_j}\right).$$

In fact, the double bracket equation is nothing but the gradient flow for the function $F(\theta) = \kappa(Q, \mathrm{Ad}_\theta N)$, where N is in the orbit.

Euler-Poincaré equation

For the Cartan subgroup $H \subset G$ s.t. $\mathrm{Lie}(H) = \mathfrak{h}$, the Fisher-Rao bilinear form $\langle \cdot, \cdot \rangle$ is invariant w.r.t. Ad_H and induces Fisher-Rao Riemannian metric on $\mathcal{O} = G/H$. The geodesic equation can be formulated through Euler-Poincaré reduction w.r.t. the Lagrangian function

$$\mathcal{L}(X) := \frac{1}{2}\langle X, X \rangle = -c\kappa\left([X, Q], [X, N']\right), \quad X \in \mathfrak{g},$$

regarded as a left-invariant function on $TG \cong G \times \mathfrak{g}$.

Theorem

The Euler-Poincaré equation for the geodesic flow w.r.t. the above Fisher-Rao Riemannian metric is given as

$$\frac{\mathrm{d}}{\mathrm{d}t}[[X, Q], N'] = [X, [[X, Q], N']]. \qquad (1)$$

Euler-Poincaré reduction 1

We consider the tangent bundle TG to the Lie group G and a left-invariant Lagrangian $\widetilde{L}$ on TG.

Through the left-translation $L_g : G \ni a \mapsto ga \in G$, we have the left-trivialization of the tangent bundle:
$TG \ni (g, v_g) \to \left(g, (\mathrm{d}L_g)^{-1} v_g\right) \in G \times \mathfrak{g}$, where $v_g \in T_g G$.

The Lagrangian $\widetilde{L}$ induces a function L on $\mathfrak{g}$:
$\widetilde{L}(g, v_g) = L\left((\mathrm{d}L_g)^{-1} v_g\right)$.

Euler-Poincaré reduction 2

Theorem (Euler-Poincaré reduction)

For a smooth curve $I \ni t \mapsto g(t) \in G$, the followings are equivalent:

1. $g(t)$ *solves the Euler-Lagrange equation* $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{g}}(g(t), \dot{g}(t))\right) = \frac{\partial L}{\partial g}(g(t), \dot{g}(t))$.
2. $g(t)$ *is a critical point of the variation* $\delta \int_I \widetilde{L}(g(t), \dot{g}(t))\,\mathrm{d}t = 0$ *with fixed end points.*
3. $v(t) = (\mathrm{d}L_g)^{-1}\,\dot{g}(t) \in \mathfrak{g}$ *solves the Euler-Poincaré equation*
$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v} = \mathrm{ad}_v^*\frac{\partial L}{\partial v}.$$

Here, $\frac{\partial L}{\partial v}(v) \in \mathfrak{g}^*$ is the differential of L at $v \in \mathfrak{g}$.

Observations

Note that the Lagrangian $\mathcal{L}$ is invariant w.r.t. the right-action by H and the vector field $[X, [[X, Q], N']]$ in (1) can be restricted to $\mathfrak{m}$. On $\mathfrak{m}$, the Euler-Poincaré equation (1) can also be written as

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mathrm{ad}_Q^{-1} \circ \mathrm{ad}_{N'}^{-1} \left([X, [[X, Q], N']]\right).$$

The right-hand-side coincides with $\nabla_X X$ in §2. Thus, the Euler-Poincaré equation (1) can be restricted to $\mathcal{O} = G/H$ and it coincides with (the tangential part of) the geodesic equation on $\mathcal{O} = G/H$ w.r.t. the Fisher-Rao Riemannian metric. Although Theorem is formulated as a Lagrangian system, we can easily formulate the system in the Hamiltonian framework, as TG and T^*G are identified through the Killing form κ.

Lie-Poisson equation

The geodesic flow can also be formulated in the Lie-Poisson formalism. We identify the Lie algebra $\mathfrak{g}$ and its dual $\mathfrak{g}^*$ through the Killing form and take the Hamiltonian

$$\mathcal{H}(X) := \frac{1}{2c}\kappa\left(X_{\mathfrak{m}}, \mathrm{ad}_Q^{-1} \circ \mathrm{ad}_{N'}^{-1} X_{\mathfrak{m}}\right).$$

Here, $X_{\mathfrak{m}} \in \mathfrak{m}$ is the $\mathfrak{m}$-component of X.

Theorem

The Lie-Poisson equation for the geodesic flow w.r.t. the above Fisher-Rao Riemannian metric is given as

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{1}{c} \cdot \left[X, \mathrm{ad}_Q^{-1} \circ \mathrm{ad}_{N'}^{-1} X_{\mathfrak{m}}\right].$$

Reduction to $T^*(G/H)$

We consider the action of the Cartan subgroup $H \subset G$ on T^*G from right:

$$H \curvearrowright T^*G \overset{left}{\cong} G\times\mathfrak{g}^* \ni (g,\xi) \rightsquigarrow \left(gh^{-1}, \mathrm{Ad}^*_{h^{-1}}\xi\right) \in G\times\mathfrak{g}^* \overset{left}{\cong} T^*G.$$

The momentum mapping for the H-action is given by $\Phi : T^*G \overset{left}{\cong} G \times \mathfrak{g}^* \ni (g,\xi) \mapsto \xi_{\mathfrak{h}} \in \mathfrak{h}^*$. Applying Marsden-Weinstein reduction to the momentum manifold $\Phi^{-1}(0)$, we have the Hamiltonian system on $\Phi^{-1}(0)/H \cong T^*(G/H)$ with the canonical symplectic structure and the Hamiltonian $\mathcal{H}_0$.

Remark on Mishchenko-Fomenko geodesic flow

On a semi-simple Lie group G whose Lie algebra $\mathfrak{g}$, Mishchenko and Fomenko [MF] have introduced a class of left-invariant metrics whose geodesic flows are completely integrable.

The Lie-Poisson equation for this geodesic flow is described as follows:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = [X, \phi_{a,b,D}(X)], \qquad X \in \mathfrak{g}.$$

Here, a, b are regular elements in a fixed Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ and the operator $\phi_{a,b,D} : \mathfrak{g} \to \mathfrak{g}$ is given as $\phi_{a,b,D}(X) = \mathrm{ad}_b \circ \mathrm{ad}_a^{-1}(X_\mathfrak{m}) + D(X_\mathfrak{h})$, $X = X_\mathfrak{m} + X_\mathfrak{h}$, $X_\mathfrak{m} \in \mathfrak{m} = \mathfrak{h}^{\perp_\kappa}$, $X_\mathfrak{h} \in \mathfrak{h}$, $D : \mathfrak{h} \to \mathfrak{h}$: κ-symmetric operator.

There are many researches on this system. See e.g. [RT].

Remark from the sub-Riemannian point of view

The κ-complement $\mathfrak{m} \subset \mathfrak{g}$ to the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ gives rise to a left-invariant (bracket-generating) sub-Riemannian structure because of the relation $[\mathfrak{m}, \mathfrak{m}] = \mathfrak{h}$. The Fisher-Rao semi-definite metric $\langle \cdot, \cdot \rangle$ then induces a sub-Riemannian metric w.r.t. this sub-Riemannian structure. The Euler-Poincaré (or equivalently Lie-Poisson) equation is describing the flow for the normal geodesics for this sub-Riemannian metric.

Example on $SO(3)$

If $M = G = SO(3)$, set $X = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}$,

$Q = q \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $N' = n' \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

The Fisher-Rao metric is definite on $S^2 = SO(3)/H$ with $H \cong SO(2)$ being the rotations around x_3-axis.

Then, the Euler-Poincaré equation (1) is given as

$$\begin{cases} \dot{x}_1 &= \frac{1}{cqn'} x_2 x_3, \\ \dot{x}_2 &= -\frac{1}{cqn'} x_3 x_1, \\ \dot{x}_3 &= 0, \end{cases}$$

which is the Euler(-Arnol'd) equation for a symmetric free rigid body. (This is an integrable system.)

Conclusion

The talk is briefly summarized as follows:

- In §2, we have reviewed the general framework of transformation models.
- In §3, we have considered a concrete family of probability density functions in association with the cost function F earlier studied in [BBR] to investigate the generalized Toda lattice equation expressed as the gradient flow by the Brockett double bracket equation. We have rather considered the information geometry in relation to the probability density functions arising from the same function.

Perspectives

Some of the perspectives of Euler-Poincaré equation (1) are as follows:

- The Euler-Poincaré equation (1) is similar to the Brockett double bracket equation discussed e.g. in [BBR] and to Euler-Arnol'd equation for generalized free rigid body dynamics studied e.g. in [MF]. The integrability of (1) should be considered.
- The stability of equilibria for (1) should be investigated.
- Comparison of the geodesic flow (1) with the Souriau's thermodynamical formulation of information geometry and the associated Euler-Poincaré equations in [BG, M].

References 1

[A] S.-I. Amari, *Information Geometry and its Applications*. Springer, 2016.

[AN] S.-I. Amari, H. Nagaoka, *Methods of information geometry*. American Mathematical Society; Oxford University Press, 2000.

[AJLS] N. Ay, J. Jost, Hông Vân Lê, L. Schwachhöfer, *Information geometry*. Springer, 2017.

[BG] F. Barbaresco, F. Gay-Balmaz, Lie group cohomology and (multi)symplectic integrators: new geometric tools for Lie group machine learning based on Souriau geometric statistical mechanics, *Entropy* **22**, 498, 2020.

[BBE] O. E. Barndorff-Nielsen, P. Blaesild, P. S. Eriksen, *Decomposition and invariance of measures, and statistical transformation models*. Springer-Verlag, 1989.

[BBR] A. Bloch, R. W. Brockett, T. S. Ratiu, Completely integrable gradient flows. *Comm. Math. Phys.* **147**(1), 57-74, 1992.

[C1] M. Casalis, Familles exponentielles naturelles sur $\mathbb{R}^d$ invariantes par un groupe, *Int. Stat. Rev.* **59**(2), 241-262, 1991.

References 2

[C2] M. Casalis, Les familles exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique, *C. R. Acad. Sci. Paris* **312**, Série I, 537-540, 1991.

[F] M. Fréchet, Sur l'existence de certaines évaluations statistiques au cas de petits échantillons, *Rev. Int. Statist.* **11**, 182-205, 1943.

[I] H. Ishi, Special Issue "Affine differential geometry and Hesse geometry: a tribute and memorial to Jean-Louis Koszul," *Info. Geo.* 2021.

[K] J.-L. Koszul, Domaines bornés homogènes et orbites de groupes de transformations affines, *Bull. Soc. Math. France* **89**, 515-533, 1961.

[M] Ch.-M. Marle, On Gibbs states of mechanical systems with symmetries, preprint, arXiv: 2012.00582, 2021.

[MF] A. S. Mishchenko, A. T. Fomenko, Euler equations on finite-dimensional Lie groups. *Math. USSR-Izv.* **12**(2), 371-389, 1978.

References 3

[Na1] Y. Nakamura, Completely integrable gradient systems on the manifolds of Gaussian and Multinomial Distributions. *Japan J. Indust. Appl. Math.* **10**(2), 179-189, 1993.

[Na2] Y. Nakamura, Gradient systems associated with probability distributions. *Japan J. Indust. Appl. Math.* **11**(1), 21-30,1994.

[Na3] Y. Nakamura, A tau-function for the finite Toda molecule, and information spaces. Contemp. Math. **179**, 205-211, Amer. Math. Soc., 1994.

[No] K. Nomizu, Invariant affine connections on homogeneous spaces, Amer. J. Math., **76**(1), 33-65, 1954.

[RT] T. S. Ratiu, D. Tarama, Geodesic flows on real forms of complex semi-simple Lie groups of rigid body type, Res. Math. Sci. **7**, 32, 2020.

[TF] D. Tarama, J.-P. Françoise, Information Geometry and Hamiltonian Systems on Lie Groups, In: F. Nielsen, F. Barbaresco, eds., *Geometric Science of Information*, GSI 2021, **Lecture Notes in Computer Science**, **12829**. Springer, Cham, 2021.