Lagrangian and Hamiltonian systems over Lie groups appearing in statistical transformation models

Daisuke Tarama* and Jean-Pierre Françoise†

Abstract

The general framework of statistical transformation models is considered with the Fisher-Rao (semi-definite) metric and the Amari-Chentsov cubic tensor. The geodesic flows on compact semi-simple Lie groups with respect to a class of Fisher-Rao semi-definite metrics are studied and the Euler-Poincaré and the Lie-Poisson equations are explicitly found.
Key words Statistical transformation model, information geometry, Lie group, homogeneous space, geodesic flow, Lagrangian system, Hamiltonian system
MSC(2020): 37J39, 53B12, 53C30, 53D25, 70H33

1 Introduction

This is a brief summary of an on-going project on statistical transformation models in relation to dynamical systems theory carried out so far by the the authors. It is based on a talk delivered by the first author in the workshop "Submanifold Geometry and Lie Group Actions 2021" held on the 20th and 21st of March, 2022.

As is well known, one of the typical framework of the information geometry appears in the statistical inference on the smooth manifold M of data, equipped with a suitable volume form $\mathsf{d}\mathrm{vol}_M$, on which one considers a parametrized family of probability density functions $\rho(u,x)$ with the parameter $u \in U$ in another smooth manifold U. The *Fisher-Rao metric* is the (semi-definite) metric on U, which is in local coordinates $(u_1, \ldots, u_n)$ of U described as $\sum_{i,j} g_{ij} \mathsf{d}u^i \otimes \mathsf{d}u^j$ with

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

Here, $E[f]$ stands for the expectation of a random variable $f : M \to \mathbb{R}$ with respect to ρ. The importance of the Fisher-Rao metric lies on the famous Cramer-Rao Theorem, as its inverse gives a lower bound of the variant and covariant matrix, when it is positive-definite. (See e.g. [1, 2].)

Similarly, the *Amari-Chentsov tensor* is the cubic tensor on U which given in local coordinates $(u_1, \ldots, u_n)$ on U as $\sum_{i,j,k} c_{ijk} \mathsf{d}u^i \otimes \mathsf{d}u^j \otimes \mathsf{d}u^k$ with

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

*Department of Mathematical Sciences, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan. Partially supported by JSPS KAKENHI Grant Number 19K14540. E-mail: dtarama@fc.ritsumei.ac.jp
†Sorbonne Université, Laboratoire Jacques-Louis Lions, UMR 7598 CNRS, 4 Place Jussieu, 75252 Paris, France. E-mail: jean-pierre.francoise@sorbonne-universite.fr

This leads the notion of α-connections, in relation to which one of the most well-investigated geometric structures are the so-called dually flat structures. (See e.g. [1, 2, 3].)

In this summary, one considers a rather special case where the parameter manifold U is a Lie group G acting smoothly on the data manifold M and the family of probability density functions satisfies a certain invariance property. Such a statistical model is called a *statistical transformation model* and has been studied by Barndorff-Nielsen and his coauthors in [4]. (See also [3].) Although the statistical transformation models have been studied by many authors, one considers them in the present summary from the viewpoint of Lagrangian and Hamiltonian mechanics. The authors expect that the present research leads further developments over the earlier studies [8, 9, 10, 11] by Nakamura where the relation between information geometry and the (integrable) Hamiltonian systems are studied.

The structure of the subsequent sections is as follows:
Section 2 deals with the general framework of statistical transformation models and the description of the Fisher-Rao (semi-definite) metric and the Amari-Chentsov cubic tensor is given as left-invariant covariant tensors on the Lie group G.

In Section 3, one considers a specific family of probability density functions motivated by the study [5] on the double bracket equations for Toda lattice equations. The geodesic flow is considered on a certain homogeneous manifold $\mathcal{O} = G/H$ of G with respect to a Lie subgroup $H \subset G$ compatible with the Fisher-Rao semi-definite metric. The descriptions are given in view of geometric mechanics in the framework of Lagrangian and Hamiltonian systems. In accordance to these Lagrangian and Hamiltonian formalism, the explicit description of the Euler-Poincaré and Lie-Poisson equations are found. The Euler-Poincaré equation in Section 3 is also dealt with in the previous paper [15] by the authors.

2 General framework of statistical transformation models

In this section, we describe the general framework of statistical transformation models in the line of [4] and [3, §8.3].

Consider a smooth manifold M of data on which a Lie group G smoothly acts. The Lie algebra of G is denoted by $\mathfrak{g}$. We assume that the manifold M has an invariant volume form dvol_M with respect to the Lie group action. Now, suppose that we are given a family of probability density functions parametrized by the Lie group G:

$$\rho : G \times M \ni (g, x) \mapsto \rho(g, x) \in \mathbb{R}.$$

Besides the trivial condition $\displaystyle\int_M \rho(g, x)\mathsf{dvol}_M(x) = 1$ for any $g \in G$, we require the following invariance of the family of probability density functions with respect the action by G:

$$\rho(gh, x) = \rho(g, hx), \qquad \forall g, h \in G, \quad , \forall x \in M. \tag{2.1}$$

As is well known, one can think about the standard framework of information geometry, where the Fisher-Rao metric and the Amari-Chentsov cubic tensor play important rôles. See [1, 2, 3] for the general theories of information geometry. By virtue of the symmetry (2.1), these tensors on the Lie group G turn out to be left-invariant and they are hence reasonably described as explained below.

Using (2.1), we obtain the following formula:

$$X^M_x\left[\rho(g, x)\right] = -X^{(L)}_g\left[\rho(g, x)\right] \tag{2.2}$$

for arbitrary $g \in G$, $x \in M$, $X \in \mathfrak{g}$. Here, $X^M \in \mathfrak{X}(M)$ stands for the vector field on M induced by X through

$$X^M_x [f] = \left.\frac{\mathsf{d}}{\mathsf{d}t}\right|_{t=0} f\left(\exp(-tX) \cdot x\right), \quad x \in M,$$

where $f \in \mathcal{C}^\infty(M)$ is any smooth function on M, whereas $X^{(L)} \in \mathfrak{X}(G)$ is the left-invariant vector field corresponding to X.

For a random variable $f : M \to \mathbb{R}$, which is assumed to be a smooth function, we denote the expectation with respect to the density function ρ by

$$E[f] := \int_M f(x)\rho(g, x)\mathsf{d}\mathrm{vol}_M(x).$$

We can verify the following formulae by (2.2):

$$\begin{aligned} &E\left[X^M[\rho]\right] = 0, \\ &E\left[X^M[\rho]\, Y^M[\rho]\right] + E\left[X^M\left[Y^M[\rho]\right]\right] = 0 \end{aligned}$$

for any $X, Y \in \mathfrak{g}$. Note that, if one fixes X, Y, then $E\left[X^M[\rho]\, Y^M[\rho]\right] = E\left[X^{(L)}[\rho]\, Y^{(L)}[\rho]\right]$ does not depend on $g \in G$.

Definition 2.1. The positive-semi-definite bilinear form

$$\langle X, Y\rangle := E\left[X^M[\rho]\, Y^M[\rho]\right], \qquad X, Y \in \mathfrak{g},$$

on $\mathfrak{g}$ is called the *Fisher-Rao bilinear form* and the associated left-invariant $(0, 2)$-tensor, denoted by the same symbol $\langle\cdot,\cdot\rangle$, is called the *Fisher-Rao semi-definite metric* on G. □

Note that the Fisher-Rao bilinear form,as well as the Fisher-Rao semi-definite metric, is not necessarily positive-definite.

Similarly, the cubic multi-linear form

$$E\left[X^M[\rho]\, Y^M[\rho]\, Z^M[\rho]\right]$$

can be shown to be independent of $g \in G$.

Definition 2.2. The left-invariant $(0, 3)$ tensor field

$$C(X, Y, Z) := E\left[X^M[\rho]\, Y^M[\rho]\, Z^M[\rho]\right], \qquad X, Y, Z \in \mathfrak{g},$$

on G is called the *Amari-Chentsov cubic tensor* on G. □

Now, assume that the Fisher-Rao semi-definite metric is positive-definite. Then, we call it the Fisher-Rao Riemannian metric on G. In this case, we can naturally associate the Levi-Civita connection ∇ on G. Following the standard arguments in information geometry, we can consider the α-connection $\nabla^{(\alpha)}$ on G defined through

$$\left\langle \nabla^{(\alpha)}_X Y, Z\right\rangle = \langle\nabla_X Y, Z\rangle + \frac{\alpha}{2} C\left(X, Y, Z\right),$$

where $X, Y, Z \in \mathfrak{g}$ are arbitrary. Note that the connection $\nabla^{(\alpha)}$, as well as ∇, is left-invariant.

Even though the Fisher-Rao tensor $\langle\cdot,\cdot\rangle$ is not necessarily positive-definite in general, we can still consider the induced Riemannian metric, which we also call Fisher-Rao Riemannian metric, on the quotient manifold of G by a Lie subgroup under some additional conditions. Although the

Amari-Chentsov cubic tensor is of much interest, we focus on the geodesic flow with respect to the Fisher-Rao Riemannian metric below.

For simplicity, we now suppose that the Lie group G is a compact semi-simple Lie group and the Lie algebra $\mathfrak{g}$ is equipped with the Killing form κ which is negative-definite as is well known. Now, we further assume that there exists a Lie subgroup $H \subset G$ whose Lie algebra $\mathfrak{h}$ gives rise to the orthogonal decomposition $\mathfrak{g} = \mathfrak{h} \dot{+} \mathfrak{m}$ into the direct sum of two vector subspaces with respect to κ. Here, $\mathfrak{m}$ denotes the κ-orthogonal complement to $\mathfrak{h}$. If the Fisher-Rao semi-definite metric $\langle \cdot, \cdot \rangle$ is definite when restricted to $\mathfrak{m}$, then we can associate the positive-definite Riemannian metric on G/H.

Using a theorem in [12, Theorem 13.1], we can describe the induce Levi-Civita connection on G/H with respect to the Fisher-Rao Riemannian metric as

$$\nabla_X Y = \frac{1}{2} [X, Y]_{\mathfrak{m}} + U(X, Y), \quad X, Y \in \mathfrak{m}, \tag{2.3}$$

where $U(X, Y)$ is defined through $\langle U(X, Y), Z \rangle = \frac{1}{2} \left(\langle [X, Z]_{\mathfrak{m}}, Y \rangle + \langle X, [Y, Z]_{\mathfrak{m}} \rangle \right)$ for any $Z \in \mathfrak{m}$. Here, for any $W \in \mathfrak{g} = \mathfrak{h} \dot{+} \mathfrak{m}$, $W_{\mathfrak{m}}$ stands for its $\mathfrak{m}$-component.

3 Euler-Poincaré and Lie-Poisson equations for a class of geodesic flows

Now, we focus on the special case where the sample manifold M coincides with the compact semi-simple Lie group G, whose Lie algebra $\mathfrak{g}$ is equipped with the negative-definite Killing form κ, and the family of probability density functions ρ are concretely given as

$$\rho(g, h) := c \cdot \exp\left(F\left(gh\right)\right), \quad g, h \in G.$$

Here, $F(\theta) = \kappa\left(Q, \mathrm{Ad}_\theta N\right)$, $\theta \in G$, with Q, N being fixed elements in $\mathfrak{g}$. We write the Harr measure on G as $\mathsf{d}\mathrm{vol}_G$. The positive constant c is determined in order for ρ to be a probability density function:

$$1 = c \cdot \int_G \exp\left(F\left(gh\right)\right) \mathsf{d}\mathrm{vol}_G(h), \quad g \in G.$$

Remark 3.1. The motivation to introduce the above concrete probability density functions comes from the work [5] on the description of the Toda lattice equations as double bracket equations on the split real forms of complex semi-simple Lie algebras. In fact, the double bracket equation is obtained as the gradient flow with the function F being the cost function when restricted to special orbits called Jacobi orbits. ♦

Following the arguments in the previous section, the Fisher-Rao semi-definite bilinear form $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ can be calculated as

$$\langle X, Y \rangle = -c \cdot \kappa\left(\left[X, Q\right], \left[Y, N'\right]\right), \quad X, Y \in \mathfrak{g},$$

where $N' := \int_G \mathrm{Ad}_{gh} N \rho(g, h) \mathsf{d}\mathrm{vol}_G(h) = \int_G \mathrm{Ad}_h N \rho(e, h) \mathsf{d}\mathrm{vol}_G(h)$. Assuming that N' and Q are regular elements of the same Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$, whose corresponding Cartan subgroup is denoted by $H \subset G$, we see that the the Fisher-Rao bilinear form $\langle \cdot, \cdot \rangle$ is positive-definite on $\mathfrak{m} = \mathfrak{h}^{\perp_\kappa}$ and that the Fisher-Rao semi-definite metric $\langle \cdot, \cdot \rangle$ induces the Fisher-Rao Riemannian metric on $\mathcal{O} := G/H$.

To clarify the mechanical structure of the settings, we consider the Lagrangian and the Hamiltonian formalisms of the geodesic flow on $\mathcal{O}$ with respect to the induced Riemannian metric. It should be pointed out that these methods are essentially the same as the geodesics of the Levi-Civita connection (2.3) appearing in the last part of Section 2.

The Lagrangian formalism can be carried out through the Euler-Poincaré reduction of the Lagrangian system on $TG \cong G \times \mathfrak{g}$ with respect to the Lagrangian function

$$\mathcal{L}(X) := \frac{1}{2}\langle X, X\rangle = -\frac{c}{2}\kappa\left([X,Q],[X,N']\right), \quad X \in \mathfrak{g}. \tag{3.1}$$

Note that TG is identified with $G \times \mathfrak{g}$ through the left-translation and that the Lagrangian $\mathcal{L}$ in (3.1) is left-invariant. By the Euler-Poincaré reduction, the system can essentially be described by the Euler-Poincaré equation on the Lie algebra $\mathfrak{g}$ as in the following theorem.

Theorem 3.1. *The Euler-Poincaré equation for the geodesic flow with respect to the above Lagrangian $\mathcal{L}$ in* (3.1) *is written as*

$$\frac{\mathsf{d}}{\mathsf{d}t}[[X,Q]\,,N'] = c\cdot[X,[[X,Q]\,,N']]\,. \tag{3.2}$$

□

We see that the right-hand side of (3.2) coincides with the description of the covariant derivative $\nabla_X X$ for the Levi-Civita connection as described in (2.3), setting $Y = X$.

Remark 3.2. In [15], there is an unfortunate typo in the main theorem, where the expression of the Euler-Poincaré equation should be read as (3.2) in Theorem 3.1. ♦

See e.g. [6, 13] for the details on the Euler-Poincaré reduction.

On the Hamiltonian side, the geodesic flow on the homogeneous manifold $\mathcal{O} = G/H$ equipped with the Fisher-Rao Riemannian metric is given as the Hamiltonian system on $T^*G \cong G\times\mathfrak{g}^*$ with respect to the canonical symplectic form for the Hamiltonian function

$$\mathcal{H}(X) := \frac{1}{2c}\kappa\left(X_{\mathfrak{m}}, \mathrm{ad}_Q^{-1}\circ\mathrm{ad}_{N'}^{-1}X_{\mathfrak{m}}\right), \quad X \in \mathfrak{g}. \tag{3.3}$$

Here, the cotangent bundle T^*G is identified with the product $G \times \mathfrak{g}$ via the left-translations and the Killing form κ. In (3.3), $X_{\mathfrak{m}}$ stands for the $\mathfrak{m}$-component of $X \in \mathfrak{g} = \mathfrak{h}\dot{+}\mathfrak{m}$.

The Hamiltonian $\mathcal{H}$ in (3.3) is left-invariant and hence the Lie-Poisson reduction procedure applies. This yields the Lie-Poisson equation on $\mathfrak{g}^* \cong_\kappa \mathfrak{g}$ as in the following theorem.

Theorem 3.2. *The Lie-Poisson equation for the geodesic flow with respect to the above Hamiltonian $\mathcal{H}$ in* (3.3) *is written as*

$$\frac{\mathsf{d}}{\mathsf{d}t}X = \frac{1}{c}\cdot\left[X, \mathrm{ad}_Q^{-1}\circ\mathrm{ad}_{N'}^{-1}X_{\mathfrak{m}}\right], \quad X \in \mathfrak{g}. \tag{3.4}$$

□

See e.g. [6, 13] for the details on the Lie-Poisson reduction.

Note that the Hamiltonian system (3.4) has a resemblance to the Euler equation for the integrable geodesic flow on semi-simple Lie groups with respect to the left-invariant metric of rigid body type as was introduced by Mishchenko and Fomenko [7]. (See also [14] for the analysis around the equilibrium points of this integrable geodesic flow.) As we also observe the difference between the two systems, however, the integrability of (3.4) is unknown so far. The detailed analysis of the systems in (3.2) and (3.4) will be carried out in future works.

References

[1] S.-I. Amari, *Information Geometry and its Applications*, Applied Mathematical Sciences, **194**, Springer, Tokyo, 2016.

[2] N. Ay, J. Jost, Hông Vân Lê, L. Schwachhöfer, *Information geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, **64**, Springer, Cham, 2017.

[3] S.-I. Amari, H. Nagaoka, *Methods of information geometry*, translated from the 1993 Japanese original by D. Harada, Translations of Mathematical Monographs, **191**, American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000.

[4] O. E. Barndorff-Nielsen, P. Blæsild, P. S. Eriksen, *Decomposition and invariance of measures, and statistical transformation models*, Lecture Notes in Statistics, **58**, Springer-Verlag, New York, 1989.

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

[6] J. E. Marsden, T. S. Ratiu, *Introduction to Mechanics and Symmetry*, 2nd ed., Texts in Applied Mathematics, **17**, Springer-Verlag, New York-Heidelberg-Berlin, 2003.

[7] A. S. Mishchenko, A. T. Fomenko, Euler equations on finite-dimensional Lie groups. *Math. USSR-Izv.* **12**(2), 371-389, 1978. Translated from *Russian, Izvest. Akad. Nauk SSSR, Ser. Matem.* **42**, 1978, 396-415.

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

[9] Y. Nakamura, Neurodynamics and nonlinear integrable systems of Lax type, *Japan J. Indust. Appl. Math.*, **11**(1), 11-20, 1994.

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

[11] Y. Nakamura, A tau-function for the finite Toda molecule and information spaces, In: Y. Maeda, H. Omori, A Weinstein (eds.), *Symplectic geometry and quantization* (Sanda and Yokohama, 1993), 205-211, Contemp. Math., **179**, Amer. Math. Soc., Providence, RI, 1994.

[12] K. Nomizu, Affine connections on homogeneous spaces, *Amer. J. Math.*, **76**(1), 33-65, 1954.

[13] T. S. Ratiu et al., A Crash Course in Geometric Mechanics, In: J. Montaldi, T. Ratiu (eds.), *Geometric Mechanics and Symmetry: the Peyresq Lectures*, Cambridge University Press, Cambridge, 2005.

[14] 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.

[15] 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, 273-280, 2021.