Fisher Information Geometry of The Barycenter Map

Mitsuhiro Itoh (Institute of Mathematics, University of Tsukuba, Japan)*1
Hiroyasu Satoh (Nippon Institute of Technology, Japan)*2

1. Introduction

We would like to report Fisher information geometry of the barycenter map associated with normalized Busemann function \(B\) of an Hadamard manifold \(X\), simply connected non-positively curved manifold and to present an application to Riemannian geometry of \(X\) from viewpoint of Fisher information geometry. This report is an improvement of [ItSat’13] together with a fine investigation of the barycenter map.

2. The barycenter map

Let \(\mathbb{P}(\mu) = \mathcal{P}(\partial X, d\theta)\) be a probability measure on the ideal boundary \(\partial X\) of \(X\). A point \(x \in X\) is called a barycenter of \(\mu\), when \(x\) is a critical point of the \(\mu\)-average Busemann function on \(X\);

\[B_{\mu}(y) = \int_{\partial X} B_{\theta}(y) d\mu(\theta), \ y \in X.\]

Denote by \(\mathcal{P}^+ = \mathcal{P}^+(\partial X, d\theta)\) the space of probability measures \(\mu = f(\theta) d\theta\) defined on \(\partial X\) satisfying \(\mu \ll d\theta\) and with continuous density \(f = f(\theta) > 0\). A point \(x \in X\) is a barycenter of a measure \(\mu\) if and only if the \(\mu\)-average one form \(dB_{\mu}(\cdot) = \int_{\partial X} dB_{\theta}(\cdot) d\mu(\theta)\) vanishes at \(x\).

We follow the idea given by [DoEa], [BeCoGa’95].

Theorem 2.1 ([ItSat’14-2]). The function \(B\) admits for any \(\mu \in \mathcal{P}^+\) a barycenter, provided (i) \(X\) satisfies the axiom of visibility and (ii) \(B_{\theta}(x)\) is continuous in \(\theta \in \partial X\).

\(X\) is said to satisfy the axiom of visibility, when any two ideal points \(\theta_1, \theta_2\) of \(\partial X\), \(\theta_1 \neq \theta_2\), can be joined by a geodesic in \(X\) (see [EO]). In [BeCoGa’95] the existence theorem is verified under the conditions that (i) \(B_{\theta}\) satisfies \(\lim_{x \to \theta_1} B_{\theta}(x) = +\infty\), when \(\theta_1 \neq \theta\) and (ii) \(B_{\theta}(\cdot)\) is continuous with respect to \(\theta\). The condition (i) can be replaced by the axiom of visibility (refer to [BGS]) to obtain Theorem 2.1.

For the uniqueness we have

Theorem 2.2 ([ItSat’14-2]). Assume (i) and (ii) in Theorem 2.1. If, for some \(\mu_0 \in \mathcal{P}^+\) the \(\mu_0\)-average Hessian

\[(\nabla d \mathbb{B}_{\mu_0})_x(\cdot, \cdot) = \int_{\theta \in \partial X} (\nabla d B_{\theta})_x(\cdot, \cdot) d\mu_0(\theta)\]

is positive definite on \(T_x X\) at any \(x \in X\), then the existence of barycenter is unique for any \(\mu \in \mathcal{P}^+\).

So, we obtain a map, called the barycenter map

\[\text{bar} : \mathcal{P}^+ = \mathcal{P}^+(\partial X, d\theta) \to X, \mu \mapsto x,\]

where \(x\) is a barycenter of \(\mu\).

Notice that the differentiability of \(\mathbb{B}_\mu\) is guaranteed when the Hessian of \(B_{\theta}\) is uniformly bounded with respect to \(\theta\).

*1 e-mail: itohm@math.tsukuba.ac.jp
*2 e-mail: hiroyasu@nit.ac.jp
§3. A fibre space structure of $\mathcal{P}^+$ over $X$ and Fisher information metric

It is easily shown that the map $\bar{\mu}$ is regular at any $\mu$, that is, the differential map

$$d\bar{\mu}: T_\mu \mathcal{P}^+ \to T_y X$$

is surjective (see [BeCoGa’96]). Moreover the map $\bar{\mu}$ is itself surjective and hence it yields a fibre space projection with fibre $\bar{\mu}^{-1}(x)$ over $x \in X$,

$$\mathcal{P}^+ (\partial X, d\theta) \downarrow \bar{\mu} \downarrow X$$

provided $X$ carries Busemann-Poisson kernel $P(x, \theta) d\theta = \exp\{-qB_\theta(x)\}$, the fundamental solution of Dirichlet problem at the boundary $\partial X$, namely, Poisson kernel represented by $B_\theta(x)$ in an exponential form ($q = q(X) > 0$ is the volume entropy of $X$). An Hadamard manifold admitting Busemann-Poisson kernel turns out to be asymptotically harmonic ([Led], [ItSat’11]), since $\Delta B_\theta$ is constant for any $\theta$.

The tangent space $T_\mu \bar{\mu}^{-1}(x)$ of $\bar{\mu}^{-1}(x)$ is characterized as

$$\{ \tau \in T_\mu \mathcal{P}^+ | \int_{\theta \in \partial X} (dB_\theta)_x(U)d\tau(\theta) = 0, \forall U \in T_x X \}$$

so one gets

**Proposition 3.1.** $\tau \in T_\mu \mathcal{P}^+$ belongs to $T_\mu \bar{\mu}^{-1}(x)$ if and only if

$$G_\mu (\tau, N_\mu(U)) = 0, \forall U \in T_x X$$

where $G_\mu$ is the Fisher information metric of $\mathcal{P}^+$ at $\mu$ and $N_\mu : T_x X \to T_\mu \mathcal{P}^+$ is a linear map defined by

$$N_\mu : T_x X \to T_\mu \mathcal{P}^+ \quad U \mapsto (dB_\theta)_x(U)d\mu(\theta).$$

From this we have

**Proposition 3.2.** At any $\mu \in \mathcal{P}^+$ the tangent space $T_\mu \mathcal{P}^+$ admits an orthogonal direct sum decomposition into the vertical and horizontal subspaces as

$$T_\mu \mathcal{P}^+ = T_\mu \bar{\mu}^{-1}(x) \oplus \text{Im}N_\mu, \ x = \bar{\mu}(\mu),$$

with $\text{dim} \text{Im}N_\mu = \text{dim} X$.

**Definition 3.1** ([AN], [Fr] and [ItSat’11]). A positive definite inner product $G_\mu$ on the tangent space $T_\mu \mathcal{P}^+$ is defined by

$$G_\mu (\tau, \tau_1) = \int_{\theta \in \partial X} \frac{d\tau}{d\mu}(\theta)\frac{d\tau_1}{d\mu}(\theta)d\mu(\theta), \ \tau, \tau_1 \in T_\mu \mathcal{P}^+.$$ 

The collection $\{G_\mu | \mu \in \mathcal{P}^+\}$ provides a Riemannian metric on $\mathcal{P}^+$, called Fisher information metric $G$. 
As the $G$ is viewed as a Riemannian metric on an infinite dimensional manifold $\mathcal{P}^+$, the Levi-Civita connection $\nabla$ is given (see p.276, [Fr])

$$\nabla_{\tau_1 \tau} = -\frac{1}{2} \left( \frac{d\tau}{d\mu}(\theta) \frac{d\tau_1}{d\mu}(\theta) - \int \frac{d\tau}{d\mu}(\theta) \frac{d\tau_1}{d\mu}(\theta) d\mu(\theta) \right) \mu,$$

(3)

at a point $\mu \in \mathcal{P}^+$ for constant vector fields $\tau, \tau_1$ on $\mathcal{P}^+$.

The space $\mathcal{P}^+$ with the metric $G$ has then constant sectional curvature $\frac{1}{4}$ (refer to Satz 2, §1, [Fr]). By using the formula (3) we have

**Theorem 3.3.** Let $\gamma(t)$ be a geodesic in $\mathcal{P}^+$ of $\gamma(0) = \mu$ and $\gamma'(0) = \tau \in T_\mu \mathcal{P}^+$, a unit tangent vector. Then $\gamma(t)$ is described as

$$\gamma(t) = \left\{ \cos^2 \frac{t}{2} + 2 \cos \frac{t}{2} \sin \frac{t}{2} \frac{d\tau}{d\mu}(\theta) + \sin^2 \frac{t}{2} \left( \frac{d\tau}{d\mu}(\theta) \right)^2 \right\} d\mu(\theta)$$

(4)

Note that the geodesic lies inside of $\mathcal{P}^+$ as far as the density is positive with respect to $\theta \in \partial X$.

**Corollary 3.4.** Every geodesic in $\mathcal{P}^+$ is periodic of period $2\pi$. The length $\ell$ of a geodesic segment joining two measures $\mu$ and $\mu_1$ of $\mathcal{P}^+$ is given by

$$\cos \frac{\ell}{2} \leq \int_{\partial X} \sqrt{\frac{d\mu_1}{d\mu}(\theta) d\mu(\theta)} = \int_{\partial X} \sqrt{\frac{d\mu}{d\mu_1}(\theta) d\mu_1(\theta)}$$

(5)

and equality “$=$” in (5) holds provided at least $\cos(\frac{\ell}{2}) + \sin(\frac{\ell}{2}) \frac{d\tau}{d\mu}(\theta) > 0$ for any $\theta$.

For these see also p. 279, [Fr]. The integration in RHS of (5) is the $f$-divergence $D_f(\mu||\mu_1) = \int f(\frac{d\mu_1}{d\mu}) d\mu$, $f(u) = \sqrt{u}$ in statistical models (refer to p. 56, [AN]).

The formula (4), an improvement of the formula given by T. Friedrich (refer to p.279, [Fr]), can then assert

**Corollary 3.5.** Let $\gamma(t) = \exp_{\mu} t \tau$ be a geodesic of $\gamma(0) = \mu$ and $\gamma'(0) = \tau$. Then $\gamma$ is entirely contained in the fibre $\text{bar}^{-1}(x)$ over $x = \text{bar}(\mu)$ if and only if $\tau$ satisfies at $\mu$

$$G_\mu(\nabla_{\tau} \tau, N_\mu(U)) = 0, \forall U \in T_\mu X,$$

(6)

The condition (6) implies that the tangent vector $\tau$ is a totally geodesic vector with respect to the second fundamental form $H$, i.e., $H(\tau, \tau) = 0$ at $\mu$, since the image $\text{Im} N_\mu$ of the linear map $N_\mu$ distributes a normal bundle of $\text{bar}^{-1}(x)$ at any measure. Here, $H_\mu(\tau, \tau_1) := (\nabla_{\tau} \tau_1)_{\perp}$ at $\mu$.

**Example 3.1.** Let $o$ be the base point for $\partial X$, dim $X \geq 2$ such that $\partial X \cong S_{\theta} X$ and $\text{bar}(\mu) = o$ for the canonical measure $\mu = d\theta \in \mathcal{P}^+$. Identify $(d\theta_o)$, with $-\sum, \theta^i e_i, \theta^i \in \mathbb{R}$, with respect to an orthonormal basis $\{e_i\}$ of $T_o X$. Define $\tau = \frac{1}{\sqrt{c}} \theta^j \partial^j d\theta$, $i \neq j$ a vector tangent to $\mathcal{P}^+(c$ is a constant normalizing $\tau$ as a unit). Then $\tau \in T_\mu \text{bar}^{-1}(o)$ is seen and $\gamma(t) = \exp_{\mu} t \tau$ is a geodesic which is, from Corollary 3.5, contained in $\text{bar}^{-1}(o)$ at least for $t$, provided the density function is positive. In fact, the $\tau$ satisfies (6).
§4. Barycentrically associated maps and isometries of $X$

A Riemannian isometry $\varphi$ of $X$ transforms every geodesic into a geodesic and hence induces naturally a map $\hat{\varphi} : \partial X \to \partial X$, a homeomorphism with respect to the cone topology. Further the normalized Busemann function admits a cocycle formula ([GJT]);

$$B_\theta(\varphi x) = B_{\hat{\varphi}^{-1}\theta}(x) + B_\theta(\varphi o), \forall (x, \theta) \in X \times \partial X$$ (7)

($o$ is the normalization point of $B_\theta$).

**Proposition 4.1** (Equivariant action formula).

\[ \text{bar} \circ \hat{\varphi} = \varphi \circ \text{bar}, \quad \text{namely} \]

\[ \text{bar}(\hat{\varphi}_* \mu) = \varphi(\text{bar}(\mu)) \quad \forall \mu \in \mathcal{P}^+, \]

where $\Phi_* : \mathcal{P}^+ \to \mathcal{P}^+$ is the push-forward of a homeomorphism $\Phi$ of $\partial X$;

\[ \int_{\theta \in \partial X} h(\theta) \, d[\Phi_* \mu](\theta) = \int_{\theta \in \partial X} (h \circ \Phi)(\theta) \, d\mu(\theta) \] (9)

for any function $h = h(\theta)$ on $\partial X$ (see p.4, [V]).

So, we consider the situation converse of Proposition 4.1 as

**Definition 4.1.** Let $\Phi : \partial X \to \partial X$ be a homeomorphism of $\partial X$. Then, a map $\varphi : X \to X$ is called barycentrically associated to $\Phi$, when $\varphi$ satisfies the relation $\text{bar} \circ \Phi_* = \varphi \circ \text{bar}$ in the diagram

\[ \mathcal{P}^+(\partial X, d\theta) \xrightarrow{\Phi_*} \mathcal{P}^+(\partial X, d\theta) \]

\[ \downarrow \text{bar} \quad \downarrow \text{bar} \]

\[ X \xrightarrow{\varphi} X \] (10)

So an isometry $\varphi$ is a map barycentrically associated to $\Phi = \hat{\varphi}$.

Let $\text{bar} : \mathcal{P}^+ \to X$ be the barycenter map. Then, with respect to a homeomorphism $\Phi : \partial X \to \partial X$ and a map $\varphi : X \to X$ we obtain the following ([ItSat’14],[ItSat’14-2])

**Theorem 4.2.** Assume that a pair $(\Phi, \varphi)$ with $\varphi \in C^1$ satisfies; (a) $\text{bar}(\Phi_* \mu) = \varphi(\text{bar}(\mu)), \forall \mu \in \mathcal{P}^+$, and (b) $\Theta(\varphi(x)) = \Phi_* (\Theta(x)), \forall x \in X$;

\[ \mathcal{P}^+(\partial X, d\theta) \xrightarrow{\Phi_*} \mathcal{P}^+(\partial X, d\theta) \]

\[ \uparrow \Theta \quad \uparrow \Theta \]

\[ X \xrightarrow{\varphi} X \] (11)

Then $\varphi$ must be a Riemannian isometry of $X$.

Here, $\Theta : X \to \mathcal{P}^+ ; y \mapsto P(y, \theta)d\theta$ is a map associated with a Busemann-Poisson kernel $P(x, \theta) = \exp\{-q B_\theta(x)\}$.

**Remark 4.1.** If $X$ admits a Busemann-Poisson kernel, then $\Theta$ gives a cross section of the fibre space $\mathcal{P}^+ \to X$, since $\text{bar}(\mu_x) = x$ for $\mu_x = P(x, \theta)d\theta$, and moreover every $\mu \in \mathcal{P}^+$ admits a unique barycenter from Theorem 2.2, since it holds

\[ \int_{\partial X} (\nabla dB_\theta)_x(U, V) d\mu_x(\theta) = q \int_{\partial X} (dB_\theta)_x(U)(dB_\theta)_x(U) d\mu_x(\theta), \quad U, V \in T_x X \]
that is

\[(\nabla d B)_{x}(U, V) = q \ G_{\mu_x}(N_{\mu_x}(U), N_{\mu_x}(V))\]

\((q > 0 \text{ is the volume entropy of } X) \text{ and at any } y \in X\)

\[(\nabla d B)_{y}(U, U) \geq C(\nabla d B)_{x}(U, U)\]

for some constant \(C > 0\), depending on \(x, y\). From these the \(\mu_x\)-average Hessian \(\nabla d B\) turns out to be positive definite everywhere.

With respect to the conditions (a) and (b) of Theorem 4.2 we have

**Theorem 4.3.** Let \(X\) be an Hadamard manifold satisfying assumptions (i) and (ii) of Theorem 2.1 and admit a Busemann-Poisson kernel. Let \(\Phi : \partial X \to \partial X\) be a homeomorphism. If a map \(\varphi : X \to X\) is \(C^1\) with surjective differential \(d\varphi_x, \forall x \in X\), then (b) implies (a).

§5. **Damek-Ricci spaces and motivation**

A Damek-Ricci space is a solvable Lie group, an \(\mathbb{R}\)-extension of a generalized Heisenberg group and carries a left invariant Riemannian metric and further provides a space on which harmonic analysis is developed ([ADY],[DamR]). For precise definition and differential geometry of Damek-Ricci space refer to [BTV]. Damek-Ricci spaces are Hadamard manifolds whose typical examples are rank one symmetric spaces of non-compact type, complex hyperbolic, quaternionic hyperbolic and Cayley hyperbolic spaces as strictly negative curved ones, except for real hyperbolic spaces ([D],[L]). Any Damek-Ricci space satisfies the axiom of visibility and has \(\theta\)-continuous Busemann function (refer to [ItSat’10]) . Moreover, it admits a Busemann-Poisson kernel (see [ItSat’10]) so that it satisfies (i) and (ii) of Theorem 2.1. Most important implication of Damek-Ricci spaces is that they provides counterexample of Lichnerowicz conjecture of non-compact version (refer to [BTV]).

So, relating to this, our motivation is to characterize Damek-Ricci spaces from a viewpoint of geometry, since only a Lie group characterization of Damek-Ricci space is known from Heber’s theorem ([Heb]). A Damek-Ricci space turns out recently to be Gromov-hyperbolic, whereas it admits zero sectional curvature (see [ItSat’14-2] for this and refer to [CDP],[Bourd],[K] for the Gromov hyperbolicity).

Thus, we pose the following. Let \(X_o\) be a Damek-Ricci space and \(X\) an Hadamard manifold quasi-isometric to \(X_o\). Assume that \(X\) admits a Busemann-Poisson kernel. Then, is \(X\) isometric, or homothetic to \(X_o\) as a Riemannian manifold ? At least from this assumption we have that any Riemannian isometry of \(X_o\) induces a homeomorphism of \(\partial X\) of \(X\) (for the detail, see [ItSat’14-2]).

**References**


