INFORMATION GEOMETRY OF THE SPACE OF PROBABILITY MEASURES AND BARYCENTER MAPS

MITSUHIRO ITOH AND HIROYASU SATOH

Abstract. In this article, we present recent developments of information geometry, namely, geometry of the Fisher metric, dualistic structures and divergences on the space of probability measures, particularly the theory of geodesics of the Fisher metric. Moreover, we consider several facts concerning the barycenter of probability measures on the ideal boundary of a Hadamard manifold from a viewpoint of the information geometry.

1. Introduction

We present recent progress of information geometry, geometry with respect to the Fisher metric, the dual connection structure and divergences for a space of probability distributions. Especially we deal with geodesics with respect to the Fisher metric. Moreover, we discuss a barycenter map defined for probability measures which are defined on an ideal boundary of a Hadamard manifold.

Let $A_1, A_2, \cdots, A_n$ be points of the Euclidean space $E$ and $w = (w_1, \ldots, w_n)$ be an ordered $n$-tuple of non-negative numbers with $\sum_{i=1}^{n} w_i = 1$. Then, we call a point $P \in E$ satisfying $\overrightarrow{OP} = \sum_{i=1}^{n} w_i \overrightarrow{OA}_i$ the barycenter of $A_i$, $i = 1, 2, \cdots, n$ with weight $w$. Here $O$ is the origin of $E$. We remark that the barycenter $P$ is independent of the choice of a reference point $O$. Consider for example $n = 3$ and $w_i = 1/3$, $i = 1, 2, 3$. We may set $O = A_1$. Then the barycenter $P$ of points $A_1, A_2, A_3$ satisfies $\overrightarrow{A_1P} = \frac{1}{3}(\overrightarrow{A_1A_2} + \overrightarrow{A_1A_3})$. This means that $P$ coincides with the center of gravity of the triangle $\triangle A_1A_2A_3$. In fact, $\overrightarrow{A_1P}$ equals the vector, multiplied by $2/3$, of the geometric vector from $A_1$ to the midpoint of its opposite side $\frac{1}{2} \left( \overrightarrow{A_1A_2} + \overrightarrow{A_1A_3} \right)$. We can also define the barycenter by a critical point of the function $f : E \to \mathbb{R}; Q \mapsto \sum_{i=1}^{n} w_i |A_i - Q|^2$. In case of points being continuously distributed, for a non-negative function $w = w(x)$ satisfying $\int_{E} w(x) \, dx = 1$, we define the function $f : E \to \mathbb{R}; y \mapsto \int_{x \in E} d(y, x)^2 w(x) \, dx$ and call a critical point of $f$ a barycenter with weight $w = w(x)$, where $d(\cdot, \cdot)$ is the distance of $E$. Since the weight function $w = w(x)$ is regarded as a density of substance distributed on $E$ of unit total mass, or density function of a probability distribution, a barycenter can be defined for a probability measure $w(\cdot) \, dx$. The notion of barycenter contributes to the conjugacy theorem of maximal compact subgroups of a semi-simple Lie group, which is one of theorems necessary for the theory of symmetric spaces, and is a consequence of the Cartan fixed point theorem in which the barycenter is utilized([20, 11]). The convexity of distance function on a Riemannian manifold of

---

This article is published in Sugaku Expositions 34 (2021), 231–253, originally appeared in Japanese in Sugaku 69 4 (2017), 387–406.

2020 Mathematics Subject Classification. Primary 46E27; Secondary 58B20, 53B21.
non-positive curvature plays an important role in investigation of barycenter. In this article, we consider barycenters with respect to a convex function, namely the Busemann function, in place of the distance function.

On considering barycenters, we need to study a space of probability measures. Let \((M, g)\) be a Riemannian manifold and \(\mathcal{P}^+(M)\) be the space consisting of all probability measures having positive density function. We can define a Riemannian metric, called the Fisher metric on \(\mathcal{P}^+(M)\). The Fisher metric provides a positive definite inner product to each tangent space \(T_{\mu}\mathcal{P}^+(M)\). Riemannian geometry of the Fisher metric on a space of probability measures has been established by T. Friedrich[15]. In [15], explicit formulae of the Levi-Civita connection and geodesics of the Fisher metric are obtained in terms of density functions. From the representation form of geodesics, we find that all geodesics of the Fisher metric are periodic, whereas \(\mathcal{P}^+(M)\) is not geodesically complete. Moreover, it is shown that the Fisher metric is a metric of constant sectional curvature \(1/4\). We develop the information geometry of the Fisher metric based on T. Friedrich’s work and apply it to the geometry of barycenter maps. We can argue the theory of geodesics, one of the basic subjects and tools of Riemannian geometry, on the space \((\mathcal{P}^+(M), G)\).

We define in this article the geometric mean of two probability measures. This notion gives a good understanding about shortest geodesics, the exponential map, the distance function and thus it enables us to develop the theory of geodesics elaborately.

The Fisher metric is a natural generalization of the Fisher information matrix in mathematical statistics and information theory. As is commonly known, a family of connections \(\{\nabla^{(\alpha)}\}_{\alpha \in \mathbb{R}}\) in which a pair \((\nabla^{(\alpha)}, \nabla^{(-\alpha)})\) is dual with respect to the Fisher metric plays a significantly important role in the geometry of statistical manifolds. Here \(\nabla^{(\alpha)}\) is called the \(\alpha\)-connection (see [1]).

We can formulate the information geometry of \(\alpha\)-connections on a statistical manifolds more widely on the space \(\mathcal{P}^+(M)\), in particular we can develop on the space \(\mathcal{P}^+(M)\) the information geometry of the specific \(\alpha\)-connections, namely, \((+1)\)-connection \(\nabla^{(+1)}\) and \((-1)\)-connection \(\nabla^{(-1)}\), which are flat and dual each other (see Corollary 3.5, §3 and [22]). With respect to this fact, a family of straight lines \(t \mapsto \mu + t\tau\) permits an interpretation of a family of geodesics of \(m\)-connection, namely, the \((-1)\)-connection, which corresponds to an affine coordinate system in a statistical manifold. By this affine parametrization, we observe that the Kullback-Leibler divergence, one of the important entropies in information geometry, provides a potential by which the Fisher metric is considered as a Hesse metric.

On the other hand, Besson et al. ([7]) showed Mostow’s rigidity theorems for compact manifolds of negative curvature by using the geometric quantity called the volume entropy and by applying the notion of barycenter which is initiated by Douady-Earle ([10]) on the hyperbolic plane. In their study Besson et al. considered probability measures, including the Poisson kernel measure, which are defined on the ideal boundary \(\partial X \cong S^{n-1}\) of a simply connected negatively curved manifold \(X^n\). They formulate for those measures the barycenter whose test function is the Busemann function. They deal with probability measures without atom, not restricting themselves to probability measures of positive density function. Here a probability measure \(\mu\) has no atom, when for any Borel set \(A\) with \(\mu(A) > 0\) there exists a Borel set \(B \subset A\) satisfying \(0 < \mu(B) < \mu(A)\). Their consideration is restrictive within rank one symmetric spaces of non-compact type. In this article,
we develop, however, geometry of barycenter on a Hadamard manifold \( X \) more
generic than symmetric spaces of non-compact type. Let \( \partial X \) be the ideal boundary
of \( X \) which is defined as a quotient of all geodesic rays by asymptotic equivalence
relation and \( B_\theta \) be the normalized Busemann function. For a probability measure
\( \mu \) on \( \partial X \), we consider the averaged Busemann function with weight \( \mu \),
\[
\mathbb{B}_\mu : X \to \mathbb{R}; \ y \mapsto \int_{\theta \in \partial X} B_\theta(y) \, d\mu(\theta).
\]
We call a critical point \( x \in X \) of \( B_\mu \) the barycenter of \( \mu \). Under the guarantee of
existence and uniqueness of barycenter, we can define a map \( \operatorname{bar} : \mathcal{P}^+(\partial X) \to X \),
by assigning to \( \mu \) a point \( x \in X \) which is a barycenter of \( \mu \) and we call \( \operatorname{bar} \) the
barycenter map (see also [12, p.554]). We discuss the existence and uniqueness of
barycenter for arbitrary \( \mu \) in §5. One of important properties of the barycenter
map is the following: an isometry \( \varphi \) of \( X \) induces naturally a homeomorphism
\( \hat{\varphi} : \partial X \to \partial X \) which satisfies \( \varphi \circ \operatorname{bar} = \operatorname{bar} \circ \hat{\varphi} \).

Now we will explain the motivation and background of our investigations of
barycenters. An \( n \)-dimensional Hadamard manifold \((X,g)\) is diffeomorphic to \( \mathbb{R}^n \)
and hence to an open ball \( D^n \) with boundary \( S^{n-1} \). A Hadamard manifold \( X \)
admits also the ideal boundary \( \partial X \). Then, we are able to consider the Dirichlet
problem at infinity: given an \( f \in C^0(\partial X) \), find a solution \( u = u(x) \) on \( X \)
to \( \Delta u|_X = 0, u|_{\partial X} = f \), where \( \Delta \) is the Laplace-Beltrami operator. This is
a geometric extension of the classical Dirichlet problem on a given bounded re-
gion with boundary in \( \mathbb{R}^n \) to a Hadamard manifold with ideal boundary. Now,
we suppose that a solution \( u \) of the Dirichlet problem at infinity with bound-
ary condition is described in terms of the integral kernel \( P(x,\theta) \), called Poisson
kernel, as \( u(x) = \int_{\theta \in \partial X} P(x,\theta) f(\theta) \, d\theta \). Then, the fundamental solution
\( P(x,\theta) \) together with the measure \( d\theta \) gives a probability measure on \( \partial X \) with positive
density function (see §6 and refer to [41, 3, 12] for precise definition of Poisson ker-
nel). When the existence of the Poisson kernel is guaranteed, we can define a map
\( \Theta : X \to \mathcal{P}^+(\partial X) ; \ x \mapsto P(x,\theta) \, d\theta \), which we call the Poisson kernel map. As men-
tioned above, \( P^+(\partial X) \) carries the Fisher metric and hence, the map \( \Theta \) is regarded
as an embedding from a Hadamard manifold \((X,g)\) into an infinite dimensional
Riemannian manifold \((P^+(\partial X),G)\) of constant curvature. Then, we obtain

\textbf{Theorem 1.1} ([32, 23, 24]).

(i) Let \((X^n, g)\) be an \( n \)-dimensional Damek-Ricci space. Then, \((X^n, g)\) carries
a Poisson kernel. Moreover, its Poisson kernel map \( \Theta \) is a homothety,
i.e., satisfies \( \Theta^* G = C g \), \( C = Q/n \) and is harmonic (i.e., minimal).
Here \( Q > 0 \) denotes the volume entropy of \( X \) which means the exponential
volume growth of \( X \).

(ii) Conversely, if a Hadamard manifold \((X^n, g)\) admits a Poisson kernel and
its Poisson kernel map \( \Theta \) is homothetic and harmonic, then the Poisson
kernel of \( X \) is expressed in the form

\begin{equation}
(1.1) \quad P(x,\theta) = \exp(-Q B_\theta(x)),
\end{equation}

where \( B_\theta(x) \) is the normalized Busemann function.

In this article, we call the Poisson kernel represented in Theorem 1.1 (ii), the
Busemann–Poisson kernel. See §6 for details. A Hadamard manifold which carries
the Busemann–Poisson kernel satisfies visibility axiom and is asymptotically harmonic. We say that $X$ satisfies the visibility axiom, if there exists a geodesic in $X$ joining arbitrary distinct two ideal points (see §5). Moreover, we say that $X$ is asymptotically harmonic, if all horospheres, level hypersurfaces of the Busemann function $B_{\theta}$ for any $\theta \in \partial(X)$ have constant mean curvature $c$ ([34]). We remark that if $X$ is asymptotically harmonic and all horospheres have constant mean curvature $c$, then $c$ coincides $-Q$ ([31]). When $X$ carries the Busemann–Poisson kernel, the barycenter of the probability measure $P(x, \theta) d\theta = \exp(-Q B_{\theta}(x)) d\theta$ is just the point $x$. When we regard the barycenter map $\bar{\Theta} : X \rightarrow P^+(\partial X)$ as the projection of a fiber space structure, the Poisson kernel map $\Theta : X \rightarrow P^+(\partial X)$ provides a section of the projection $\bar{\Theta}$.

In connection with the Busemann-Poisson kernel we make a brief introduction of Damek-Ricci spaces. A Damek-Ricci space is a solvable Lie group carrying a left-invariant metric. The family of all Damek-Ricci spaces is a class of harmonic Hadamard manifolds including rank one symmetric spaces, i.e., the complex, quaternionic hyperbolic $n$-spaces $\mathbb{C}H^n$, $\mathbb{H}H^n$, $n \geq 1$, and the octonionic hyperbolic plane $\mathbb{O}H^2$. The real hyperbolic spaces are regarded as special ones among Damek-Ricci spaces. Refer to [6, 4, 33] for Damek-Ricci spaces. Here we call a Riemannian manifold harmonic, if the volume density function or mean curvature of any geodesic spheres depends only on radius and independent of the direction from the center ([45]). Since a horosphere is certain limit of geodesic spheres, harmonic manifolds are also asymptotically harmonic. We notice that Damek-Ricci spaces carry the Busemann–Poisson kernel ([4, 24]). Since a Hadamard $(X, g)$ carrying the Busemann–Poisson kernel is asymptotically harmonic, one has the problem whether such a manifold is harmonic or not and, moreover, one can raise the problem whether it is isometric or homothetic to a Damek-Ricci space. We take one further step and suppose that such a manifold $(X, g)$ is quasi-isometric to a Damek-Ricci space $S$. Then, under this assumption, we can raise a problem whether there exists an isometry or a homothety between $(X, g)$ and the space $S$. We see that from this assumption an isometry of $S$ induces a quasi-isometry of $(X, g)$ and hence a homeomorphism of the ideal boundary $\partial X$ of $X$, since any Damek-Ricci space is Gromov hyperbolic. With these backgrounds, we proceed to our argument of barycenter of probability measures on the ideal boundary.

The article is organized as follows. In §2, we treat differential geometry of the Fisher metric $G$, i.e., the Levi-Civita connection of the metric $G$ on the space of probability measures. In §3, we introduce an outline of the theory of $\alpha$-connections which admits duality with respect to the Fisher metric. In §4, we recall the basic properties and the important results of the ideal boundary of a Hadamard manifold and the normalized Busemann function that are needed for discussing barycenters in the later sections. We define the barycenter in §5 and consider its existence and uniqueness, and then its geometric properties. In the last section, we develop our argument of the fiber space structure of the barycenter map and isometricity of a barycentrically associated map that is a transformation of $X$ induced by a homeomorphism of $\partial X$ via the barycenter map.
2. The space of probability measures and the Fisher metric

2.1. The space of probability measures. Let $M$ be a compact, connected $C^\infty$-manifold and $\mathcal{B}(M)$ be the collection of all Borel sets on $M$. $\mathcal{B}(M)$ is the smallest $\sigma$-algebra which contains all open subsets of $M$.

A probability measure on the measurable space $(M, \mathcal{B}(M))$, or simply on $M$, is a real valued function $P : \mathcal{B}(M) \to \mathbb{R}$ satisfying the following:

(i) $P(A) \geq 0$ holds for any $A \in \mathcal{B}(M)$. In particular, $P(M) = 1$ and $P(\emptyset) = 0$.

(ii) For any countable sequence of sets $\{E_j \mid j = 1, 2, \cdots\}$ of $\mathcal{B}(M)$ satisfying $E_j \cap E_k = \emptyset$, $j \neq k$

\begin{equation}
(2.1) \quad P \left( \bigcup_{j=1}^{\infty} E_j \right) = \sum_{j=1}^{\infty} P(E_j).
\end{equation}

Let $\lambda$ be a volume form on $M$. We normalize $\lambda$ such that $\int_{x \in M} d\lambda(x) = 1$ and fix it as the standard probability measure on $M$.

Let $\mathcal{P}(M)$ be the space of all probability measures on $M$. Obviously $\lambda \in \mathcal{P}(M)$. Any probability measure $\mu$ which we consider in this article is supposed to be absolutely continuous with respect to $\lambda$, denoted as $\mu \ll \lambda$. Here, a probability measure $\mu \in \mathcal{P}(M)$ is said to be absolutely continuous with respect to $\mu_1 \in \mathcal{P}(M)$, provided $\mu(A) = 0$ holds for any $A \in \mathcal{B}(M)$ satisfying $\mu_1(A) = 0$. A measure $\mu$ is absolutely continuous with respect to a finite measure $\mu_1$ if and only if there exists a function $f \in L^1(M, \mu_1)$ such that $\mu = f\mu_1$, i.e.,

\[ \mu(A) = \int_{x \in A} f(x) d\mu_1(x) \quad (\forall A \in \mathcal{B}(M)). \]

We call such a function $f \in L^1(M, \mu_1)$ the Radon-Nikodym derivative of $\mu$ with respect to $\mu_1$, denoted by $d\mu/d\mu_1$. Probability measures which we consider in this article are measures whose Radon-Nikodym derivative with respect to $\lambda$ is everywhere positive and continuous. We denote the space of all such probability measures by $\mathcal{P}^+(M)$ and call it the space of probability measures:

\begin{equation}
(2.2) \quad \mathcal{P}^+(M) = \left\{ \mu \in \mathcal{P}(M) \mid \mu \ll \lambda, \frac{d\mu}{d\lambda} \in C^0(M), \frac{d\mu}{d\lambda}(x) > 0 \ (\forall x \in M) \right\}.
\end{equation}

Then, a natural embedding is defined from $\mathcal{P}^+(M)$ into the space of $L^2$-functions $L^2(M, \lambda)$:

\begin{equation}
(2.3) \quad \rho = \rho^{(1/2)} : \mathcal{P}^+(M) \hookrightarrow L^2(M, \lambda); \mu = f \lambda \mapsto 2 \sqrt{\frac{d\mu}{d\lambda}} = 2 \sqrt{f}.
\end{equation}

We define a topology of $\mathcal{P}^+(M)$ by this embedding. We remark that there exists a sequence $\{\mu_i\}$ of probability measures of $\mathcal{P}^+(M)$ which has not necessarily a limit in $\mathcal{P}^+(M)$, even if the sequence $\{\sqrt{d\mu_i/d\lambda}\}$ is convergent in $L^2(M, \lambda)$, We notice that geometry of infinite dimensional space of probability measures is discussed in [38].

Each probability measure in $\mathcal{P}^+(M)$ is regarded as a point in a space. A path joining $\mu = f \lambda, \mu_1 = f_1 \lambda$ in $\mathcal{P}^+(M)$ given by $\mu(t) = (1-t)\mu + t\mu_1 = ((1-t)f + tf_1) \lambda$ is inside $\mathcal{P}^+(M)$ for $0 \leq t \leq 1$. Differentiating this curve $\mu(t)$ with respect to $t$, we have

\[ \dot{\mu}(t) = \frac{d}{dt}((1-t)\mu + t\mu_1) = \mu_1 - \mu = (f_1 - f) \lambda \]
and then
\[
\int_M d\mu(t) = 0.
\]
From this consideration, we define a tangent space \( T_\mu \mathcal{P}^+(M) \) at \( \mu \in \mathcal{P}^+(M) \) as follows:
\[
(2.4) \quad T_\mu \mathcal{P}^+(M) := \left\{ \tau = h \lambda \mid h \in C^0(M) \cap L^2(M, \lambda), \int_{x \in M} h(x) d\lambda(x) = 0 \right\}.
\]

**Remark 2.1.** A path joining two points is a 1-simplex. In general, an \( n \)-simplex \( \Delta^n \) whose vertices of number \( n + 1 \) are probability measures of \( \mathcal{P}^+(M) \) is a proper subset of \( \mathcal{P}^+(M) \).

**Remark 2.2.** The right hand side of (2.4) is an infinite dimensional vector space whose definition is independent of the choice of \( \mu \). We denote therefore this space by \( \mathcal{V} \). Let \( \mu \in \mathcal{P}^+(M) \). Then the sum \( \mu + \tau \) of \( \mu \) with \( \tau \in \mathcal{V} \) belongs to \( \mathcal{P}^+(M) \), if the \( C^0 \)-norm \( \sup_{x \in M} |(d\tau/d\mu)(x)| \) of \( \tau \) with respect to \( \mu \) is small enough. This fact suggests that \( \mathcal{P}^+(M) \) is like an open subset in an affine space whose associated vector space is \( \mathcal{V} \).

We define a curve \( c : (a, b) \to \mathcal{P}^+(M) \) in \( \mathcal{P}^+(M) \) by
\[
c(t) = f(x, t) \lambda \quad (a < t < b).
\]
We assume that the density function \( f(x, t) \) parameterized in \( t \) is of \( C^2 \)-class with respect to \( t \) for each \( x \in M \). The velocity vector of the curve \( c(t) \) is given by
\[
\frac{dc}{dt}(t) = \frac{\partial f}{\partial t}(x, t) \lambda \in T_{c(t)} \mathcal{P}^+(M), \quad t \in (a, b).
\]

### 2.2. Fisher Metric

We give definition of the Fisher metric and state its properties.

**Definition 2.3** ([15, 22, 24, 28, 27, 32, 43]). Let \( \mu = f(x) \lambda \in \mathcal{P}^+(M) \). We define a positive definite inner product \( G_\mu \) on \( T_\mu \mathcal{P}^+(M) \) by
\[
(2.5) \quad G_\mu(\tau, \tau_1) = \int_{x \in M} \frac{d\tau}{d\mu}(x) \frac{d\tau_1}{d\mu}(x) \mu(x) = \int_{x \in M} \frac{h(x)}{f(x)} h_1(x) f(x) \lambda(x)
\]
for \( \tau = h(x) \lambda, \tau_1 = h_1(x) \lambda \in T_{c(\tau)} \mathcal{P}^+(M) \). We call a map \( \mu \mapsto G_\mu \) the Fisher metric on \( \mathcal{P}^+(M) \). We denote \( G_\mu \)-norm of \( \tau \) by \( |\tau|_\mu := \sqrt{G_\mu(\tau, \tau)} \).

See [38] also for the Fisher metric defined on an infinite dimensional space of probability measures.

Now, we give a push-forward measure which is basic in measure theory and also in probability theory. Let \( \Phi \) be a homeomorphism of \( M \). A push-forward of measures by \( \Phi \) is a map \( \Phi_* : \mathcal{P}^+(M) \to \mathcal{P}^+(M) \) which assigns to any \( \mu \in \mathcal{P}^+(M) \) a measure \( \Phi_* \mu \in \mathcal{P}^+(M) \). Here \( \Phi_* \mu \) is defined as follows: for any \( A \in \mathcal{B}(M) \), \( \Phi_* \mu(A) := \mu(\Phi^{-1}(A)) \), i.e., \( \Phi_* \mu \) is a probability measure which satisfies
\[
(2.6) \quad \int_{x \in M} h(x) d(\Phi_* \mu)(x) = \int_{x' \in \Phi^{-1}(M)} h(\Phi(x')) d\mu(x')
\]
for any measurable function \( h : M \to \mathbb{R} \). When \( \Phi : M \to M \) is a diffeomorphism of \( M \), we find \( \Phi_* \mu = (\Phi^{-1})_* \mu \), i.e., the push-forward \( \Phi_* \) is the pullback of measures by the inverse diffeomorphism \( \Phi^{-1} \). Notice that the differential map of the push-forward \( \Phi_* \) is given by \( (d\Phi_*)\mu(\tau) = \Phi_*(\tau) \).
Theorem 2.4 ([15]). The push-forward by a homeomorphism $\Phi : M \to M$ acts isometrically on $(P^+(M), G)$, i.e., for any $\tau, \tau_1 \in T_\mu P^+(M)$ and $\mu \in P^+(M)$

\[(2.7) \quad G_{\Phi,\mu} \left((d\Phi_\sharp)\mu(\tau), (d\Phi_\sharp)\mu(\tau_1)\right) = G_{\mu}(\tau, \tau_1).\]

Remark 2.5. It is known that for any $\mu \in P^+(M)$ there exists a homeomorphism $\Phi : M \to M$ satisfying $\Phi_\sharp \lambda = \mu$ (see [14, 37]), i.e., the group of homeomorphisms of $M$ acts on $P^+(M)$ transitively. Hence, we find that $(P^+(M), G)$ is a Riemannian homogeneous space.

The embedding $\rho : P^+(M) \hookrightarrow L^2(M, \lambda)$ defined at (2.3) satisfies the following.

Proposition 2.6. The pull-back of the $L^2$-inner product $\langle \cdot, \cdot \rangle_{L^2}$ on $L^2(M, \lambda)$ by $\rho$ coincides with the Fisher metric $G$, i.e., for $\tau, \tau_1 \in T_\mu P^+(M)$, $\mu \in P^+(M)$ we have

\[(2.8) \quad (d\rho_\mu \tau, d\rho_\mu \tau_1)_{L^2} = G_{\mu}(\tau, \tau_1).\]

Remark 2.7. Since the image of the embedding $\rho$ satisfies $\text{Im} \, \rho \subset S^\infty(2) = \{ h \in L^2(M, \lambda) \, | \, \|h\|_{L^2} = 2 \}$, we can realize the Riemannian geometry of $(P^+(M), G)$ as an extrinsic submanifold geometry inside the infinite dimensional sphere $S^\infty(2)$ of radius 2.

Remark 2.8. The Fisher metric $G$ is obtained also as the second derivative of the Kullback-Leibler divergence

\[D_{\text{KL}} : P^+(M) \times P^+(M) \to \mathbb{R}; \quad D_{\text{KL}}(\mu || \mu_1) = -\int_{x \in M} \log \left( \frac{d\mu_1}{d\mu} \right)(x) \, d\mu(x),\]

i.e., one observes $\frac{d^2}{dt^2} \bigg|_{t=0} D_{\text{KL}}(\mu || \mu + t\tau) = G_\mu(\tau, \tau)$. See [1] for the arguments for statistical manifolds. $D_{\text{KL}}$ is called a potential in Hesse geometry ([42]).

Since the Fisher metric $G$ is regarded as a Riemannian metric on $P^+(M)$, the Levi-Civita connection is uniquely determined for the metric $G$. By the method of constant vector fields employed by T. Friedrich [15], we can express the Levi-Civita connection explicitly.

For any vector $\tau \in \mathcal{V}$, we define a constant vector field on $P^+(M)$ as follows: $\tau_\mu := \tau \in T_\mu P^+(M)$ at $\mu \in P^+(M)$. The integral curve of the constant vector field $\tau$ starting from $\mu \in P^+(M)$ is given by $t \mapsto \mu + t\tau$.

Theorem 2.9 ([15]). Let $\tau, \tau_1 \in \mathcal{V}$ be vectors regarded as constant vector fields on $P^+(M)$. Then, the Levi-Civita connection $\nabla$ of the Fisher metric $G$ is represented by

\[(2.8) \quad (\nabla_\tau \tau_1)_\mu = -\frac{1}{2} \left( \frac{d\tau}{d\mu} \frac{d\tau_1}{d\mu} - G_\mu(\tau, \tau_1) \right) \mu.\]

We remark that the right hand side of (2.8) gives an element of $\mathcal{V}$.

On the Riemannian curvature tensor of the metric $G$ we have the following.

Theorem 2.10 ([15, 22]). The Riemannian curvature tensor of the Fisher metric $G$ is given as follows:

\[(2.9) \quad R_\mu(\tau_1, \tau_2)\tau = \frac{1}{4} \left( G_\mu(\tau, \tau_2)\tau_1 - G_\mu(\tau, \tau_1)\tau_2 \right),\]

which indicates that $(P^+(M), G)$ is a Riemannian manifold of constant sectional curvature $1/4$. 
The following theorem gives a description of geodesics associated to the Levi-Civita connection.

**Theorem 2.11 ([22, 27, 28]).** Let \( t \mapsto \mu(t) = f(x,t) \lambda \) be a geodesic on \( \mathcal{P}^+(M) \) with respect to the Levi-Civita connection \( \nabla \) with initial conditions \( \mu(0) = \mu \) and \( \dot{\mu}(0) = \tau \in T_\mu \mathcal{P}^+(M) \), \( |\tau|_\mu = 1 \). Then, \( \mu(t) \) is described by

\[
\mu(t) = \left( \cos \frac{t}{2} + \frac{\tau}{2} \frac{d\mu}{d\mu} \right)^2 \mu.
\]

As a consequence, it is shown that the geodesic \( \mu(t) \) is periodic of period \( 2\pi \), while \( (\mathcal{P}^+(M), G) \) is not geodesically complete. In fact, the value of \( \mu(t) \) at \( t = \pi \) is given from (2.10) by \( \mu(\pi) = (d\tau/d\mu)^2 \mu \), whereas there exists \( x_0 \in M \) such that \( d\tau/d\mu(x_0) = 0 \), since \( \tau \) satisfies \( \int_{x \in M} (d\tau/d\mu) d\mu(x) = \int_{x \in M} d\tau(x) = 0 \). Hence, the density function of a probability measure \( \mu(\pi) \) is never positive at least at \( x_0 \) and then \( \mu(\pi) \notin \mathcal{P}^+(M) \). We find thus that any geodesic \( \mu(t) \) can not reach \( \mu(\pi) \) inside \( \mathcal{P}^+(M) \).

Theorem 2.11 is shown by the lemmas given by Friedrich (see [15, 22]), while the detail of the proof is omitted. We give now definition of a specific function together with a certain map for intimate investigation of geodesics associated to the Fisher metric.

**Definition 2.12.** Let \( \mu, \mu_1 \in \mathcal{P}^+(M) \). Define a function \( \ell : \mathcal{P}^+(M) \times \mathcal{P}^+(M) \rightarrow [0, \pi) \) and a map \( \sigma : \mathcal{P}^+(M) \times \mathcal{P}^+(M) \rightarrow \mathcal{P}^+(M) \), respectively, as follows:

\[
(2.10) \quad \cos \frac{\ell(\mu, \mu_1)}{2} := \int_{x \in M} \sqrt{\frac{d\mu_1}{d\mu}(x) d\mu(x)},
\]

\[
(2.12) \quad \sigma(\mu, \mu_1) := \left( \cos \frac{\ell(\mu, \mu_1)}{2} \right)^{-1} \sqrt{\frac{d\mu_1}{d\mu} \mu}.
\]

We call \( \sigma(\mu, \mu_1) \) a normalized geometric mean of \( \mu \) and \( \mu_1 \).

Then, we have the following.

**Theorem 2.13 ([22, 27]).** For \( \mu, \mu_1 \), there exists a unique geodesic \( \mu(t) \) joining \( \mu \) and \( \mu_1 \) which satisfies \( \mu(0) = \mu \), \( \mu(\ell) = \mu_1 \) and \(|\mu(0)|_\mu = 1 \) and \( \mu([0, \ell]) \subset \mathcal{P}^+(M) \). Moreover, \( \ell = \ell(\mu, \mu_1) \).

In fact, let \( \ell = \ell(\mu, \mu_1) \). Then

\[
(2.13) \quad \mu(t) = \left( \cos \frac{t}{2} + \frac{1}{\tan(\ell/2)} \frac{d\tau}{d\mu} \right)^2 \mu, \quad \tau = \frac{1}{\tan(\ell/2)} \left( \sigma(\mu, \mu_1) - \mu \right)
\]

is the unique geodesic satisfying \( \mu(0) = \mu \), \( \mu(\ell) = \mu_1 \), and \( \tau \in T_\mu \mathcal{P}^+(M) \) fulfills \( |\tau|_\mu = 1 \) ([27]). Moreover, by using the normalized geometric mean \( \sigma(\mu, \mu_1) \) we have for any \( t \)

\[
\mu(t) = a(t) \mu + b(t) \mu_1 + c(t) \sigma(\mu, \mu_1).
\]

Here \( a(t), b(t), c(t) \) are the functions on \([0, \ell]\), respectively, defined as follows:

\[
a(t) = \frac{\sin^2 \frac{\ell - t}{2}}{\sin^2 \frac{\ell}{2}}, \quad b(t) = \frac{\sin^2 \frac{t}{2}}{\sin^2 \frac{\ell}{2}}, \quad c(t) = 2 \cos \frac{\ell}{2} \frac{\sin \frac{\ell - t}{2} \sin \frac{t}{2}}{\sin^2 \frac{\ell}{2}}.
\]

Since they satisfy \( a(t), b(t), c(t) \geq 0 \) and \( a(t) + b(t) + c(t) = 1 \), \( \mu(t) \) lies on the 2-simplex \( \Delta^2(\mu, \mu_1, \sigma(\mu, \mu_1)) \) and it is concluded that \( \mu(t) \in \mathcal{P}^+(M) \) for any \( t \in [0, \ell] \).
(see Figure 1). Moreover, we find easily that the function \( \ell(\mu, \mu_1) \) gives arc-length of the geodesic segment ([30]).

**Remark 2.14.** The normalized geometric mean \( \sigma(\mu, \mu_1) \) of two probability measures is characterized as the intersection of lines \( L_1, L_2 \) which are the tangent lines of the geodesic segment \( \mu(t), t \in [0, \ell] \) at the end points \( \mu, \mu_1 \), respectively.

Hence, we have

\[
\sigma(\mu, \mu_1) = \mu + \tan \frac{\ell}{2} = \mu_1 + \tan \frac{\ell}{2} \tau_1,
\]

where \( \ell = \ell(\mu, \mu_1) \) and \( \tau = \dot{\mu}(0), \tau_1 = -\dot{\mu}(\ell) \) (see Figure 1).

![Figure 1](image.png)

**Figure 1.** A geodesic segment \( \mu(t) \) lies on a plane which contains the endpoints of \( \mu(t) \) and their normalized geometric mean.

**Remark 2.15.** Ohara [36] considered a Hessian metric \( g \) with respect to a certain potential on a symmetric cone \( \Omega \), and defined a dualistic structure \( (\nabla^{(a)}, \nabla^{(-a)}) \), \(-1 \leq a \leq 1 \) on \( (\Omega, g) \). He showed that a midpoint of a geodesic segment with respect to \( \nabla^{(a)} \) (called an \( a \)-geodesic segment) is an \( a \)-power mean of endpoints. Here the \( a \)-power mean is an operator mean on \( \Omega \) generated by a function

\[
\sigma^{(a)}_1(t) = \left\{ (1 + t^a)/2 \right\}^{1/a}.
\]

Incidentally, on the Riemannian manifold \( (\mathcal{P}^+(M), G) \) a midpoint of a geodesic segment joining \( \mu, \mu_1 \in \mathcal{P}^+(M) \) is given by the normalized \( 1/2 \)-power mean.

**Theorem 2.16.** The Riemannian distance \( d \) between probability measures \( \mu \) and \( \mu_1 \) of \( \mathcal{P}^+(M) \) is given by \( d(\mu, \mu_1) = \ell(\mu, \mu_1) \). A shortest path joining \( \mu \) and \( \mu_1 \) is the geodesic segment \( \mu(t), t \in [0, \ell] \) which is given at Theorem 2.13.

**Remark 2.17.** The arc-length function \( \ell \) and the embedding \( \rho \) satisfy

\[
\cos \frac{\ell(\mu, \mu_1)}{2} = 1 - \frac{1}{8}\|\rho(\mu) - \rho(\mu_1)\|_{L^2}^2.
\]

We find from this relation that for each \( \mu \in \mathcal{P}^+(M) \) the subset \( W = \{ \mu_1 \in \mathcal{P}^+(M) \|\rho(\mu) - \rho(\mu_1)\|_{L^2} < \varepsilon \} \), \( \varepsilon < \pi/4 \) is a totally normal neighborhood of \( \mu \). Hence, we are able to define the exponential map \( \exp_\mu : T_\mu \mathcal{P}^+(M) \to \mathcal{P}^+(M) \) and moreover able to discuss Gauss’ lemma and certain minimizing properties of geodesics. By using these facts and properties we obtain Theorem 2.16. Moreover, we can show that the diameter of \( (\mathcal{P}^+(M), G) \) is equal to \( \pi \), while a proof of this fact will be elsewhere ([29]).
3. α–connections and Dually Flat Structures

In information geometry, a torsion-free affine connection $\nabla^{(\alpha)}$ parametrized by $\alpha \in \mathbb{R}$, called an $\alpha$-connection, plays an important role. The $(+1)$-connection and the $(-1)$-connection which are flat and called the $e$-connection and the $m$-connection, respectively, are particularly important. The connections $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$ are said to be dual, or adjoint to each other, with respect to the Fisher metric. Moreover, the connection $\nabla^{(0)}$ coincides with the Levi-Civita connection of the Fisher metric (see [1]). As we state as follows, the notion of $\alpha$-connection can be defined also on the space of probability measures $\mathcal{P}^+(M)$ (see [22, 30]).

**Definition 3.1.** Let $\alpha \in \mathbb{R}$. Define an embedding $\rho^{(\alpha)}: \mathcal{P}^+(M) \to L^k(M, \lambda)$ by

$$\mu = f(x)\lambda \mapsto \begin{cases} \frac{2}{1-\alpha} f(x)^{(1-\alpha)/2}, & \alpha \neq 1, \\ \log f(x), & \alpha = 1 \end{cases}$$

Here we set $k = 2/(1-\alpha)$ for $\alpha \neq 1$, and $k = \infty$ for $\alpha = 1$.

**Lemma 3.2.**

$$(3.1) \quad \langle d\rho^{(\alpha)}(\tau), d\rho^{(-\alpha)}(\tau_1) \rangle = G_\mu(\tau, \tau_1), \quad \tau, \tau_1 \in T_\mu \mathcal{P}(M).$$

Here $d\rho^{(\alpha)}$ is the differential map of $\rho^{(\alpha)}$ at $\mu$ and $\langle \cdot, \cdot \rangle$ is the natural pairing map $L^k(M, \lambda) \times L^k(M, \lambda) \to \mathbb{R}$ ($k^* := 2/(1+\alpha)$).

**Definition 3.3.** Define an $\alpha$-connection $\nabla^{(\alpha)}$ for $\alpha \in \mathbb{R}$ by

$$(3.2) \quad G_\mu(\nabla^{(\alpha)}\tau_1, \tau_2) := \int_{x \in M} \frac{\partial^2}{\partial t_1 \partial t_2} \left\{ \rho^{(\alpha)}(\mu(t, t_1, t_2)) \right\} \frac{\partial}{\partial t_2} \left\{ \rho^{(-\alpha)}(\mu(t, t_1, t_2)) \right\} \bigg|_{t=t_1=t_2=0} d\lambda(x).$$

Here $\mu(t_1, t_2) := \mu + t\tau + t_1\tau_1 + t_2\tau_2$.

In fact, the $\alpha$-connection is presented by

$$\nabla^{(\alpha)}\tau_1 = -\frac{1+\alpha}{2} \left( \frac{d\tau}{d\mu} \frac{d\tau_1}{d\mu} - G_\mu(\tau, \tau_1) \right) \mu,$$

from which we can assert that $\nabla^{(\alpha)}$ is torsion-free, namely $\nabla^{(\alpha)}$ is a symmetric connection and $\nabla^{(-1)}$ is a zero-connection and also that $\nabla^{(0)}$ is the Levi-Civita connection. Moreover, from (3.2), we obtain the following fact:

**Proposition 3.4.** $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$ are dual each other, i.e., fulfill

$$\tau \ G_\mu(\tau_1, \tau_2) = G_\mu(\nabla^{(\alpha)}\tau_1, \tau_2) + G_\mu(\tau_1, \nabla^{(-\alpha)}\tau_2).$$

Moreover, the Riemannian curvature tensor $R^{(\alpha)}$ of $\nabla^{(\alpha)}$ is expressed by

$$(3.3) \quad R^{(\alpha)}(\tau_1, \tau_2)\tau = \frac{1-\alpha^2}{4} \left\{ G_\mu(\tau, \tau_2)\tau_1 - G_\mu(\tau, \tau_1)\tau_2 \right\}.$$

Hence, from (3.3), we have the following.

**Corollary 3.5.** $\alpha$-connections on $\mathcal{P}^+(M)$ which are flat and dual each other are only the $e$-connection $\nabla^{(+1)}$ and the $m$-connection $\nabla^{(-1)}$. 
As shown in Theorem 2.11, geodesics on \( P^+(M) \) with respect to the Levi-Civita connection admit the expression formula. It is interesting to derive such a formula for a geodesic \( \mu(t) = f(t)\lambda, \ f(t) := f(x, t), \ x \in M \) with respect to the \( \alpha \)-connection. We follow the argument of the proof of Theorem 2.11 to obtain the following equation for a geodesic with respect to the \( \alpha \)-connection:

\[
\frac{\partial}{\partial t} \left( \frac{\dot{f}(t)}{f(t)} \right) + \frac{1 - \alpha}{2} \left( \frac{\dot{f}(t)}{f(t)} \right)^2 + \frac{1 + \alpha}{2} \int_{x \in M} \left( \frac{\dot{f}(t)}{f(t)} \right)^2 \, f(t) \, d\lambda = 0,
\]

\[\int_{x \in M} \dot{f}(t) \, d\lambda = 0, \quad \dot{f}(t) := \frac{\partial f}{\partial t}(x, t).\]

**Proposition 3.6.** Geodesics with respect to the \( m \)-connection \( \nabla^{(-1)} \) coincide with affine lines defined by \( t \mapsto \mu + t \tau \) in \( P^+(M) \). In fact, the density function given by \( f(x, t) = f(x) + t \, h(x), \ x \in M \) is a solution to (3.4), provided \( \alpha = -1 \).

**Proposition 3.7.** Let \( \mu(t) \) be a geodesic with respect to the \( c \)-connection \( \nabla^{(+1)} \) with initial conditions \( \mu(0) = f(x)\lambda, \dot{\mu}(0) = h(x)\lambda \). Then \( \mu(t) \) is represented in the form

\[
\mu(t) = \exp \left\{ \int_0^t \left( \int_0^s g_0(u) \, du \right) ds + t \, h(x) \right\} \mu(0).
\]

Here \( g_0(t) \) is a function of \( t \) given by \( g_0(t) = -G_{\mu(t)}(\dot{\mu}(t), \dot{\mu}(t)) \). In particular, if \( \mu(0) = \mu, \mu(\ell) = \mu_1, \ell > 0 \), then \( \mu(t) \) admits the following form

\[
\mu(t) = \left\{ \int_M \left( \frac{d\mu_1}{d\mu} \right)^{\ell/\ell} \, d\mu \right\}^{-1} \left( \frac{d\mu_1}{d\mu} \right)^{\ell/\ell} \mu
\]

(see [29]).

### 4. Riemannian manifolds of non-positive curvature

In this section, we will discuss information geometry of barycenters together with the barycenter map which is another main subject of this article. Let \( X \) be a Hadamard manifold and \( \partial X \) its ideal boundary. The barycenter map bar is a map from \( \partial X \) to \( X \), defined for \( \mu \in \partial X \) by assigning to \( \mu \) a critical point of \( \mu \)-average of the normalized Busemann function \( B_\theta : X \to \mathbb{R} \). Thus, one has bar : \( \partial X \to X \). Before giving the notion of barycenter and the barycenter map, we briefly explain Hadamard manifolds, their ideal boundary and the Busemann function.

Let \( (X, g) \) be a Hadamard manifold, i.e., a complete, simply connected Riemannian manifold of non-positive sectional curvature. In what follows, we assume that \( (X, g) \), simply \( X \), is an \( n \)-dimensional Hadamard manifold.

**Remark 4.1.** Euclidean spaces, real hyperbolic spaces, rank one symmetric spaces of non-compact type and Damek-Ricci spaces are examples of Hadamard manifolds.

Now we summarize geometric properties of Hadamard manifolds ([39]):

(i) there exists a unique shortest geodesic joining any two given points.

(ii) the distance function \( x \mapsto d(x, x_0), \ x_0 \in X \) is a convex function on \( X \).

Here a function \( f : X \to \mathbb{R} \) is said to be convex on \( X \) when for any geodesic \( \gamma \) the function \( t \mapsto f(\gamma(t)) \) is convex.
Now we will give definition of the ideal boundary of a Hadamard manifold. Any geodesic on \( X \) is assumed to be parametrized by arc-length.

**Definition 4.2.** Let \( \gamma, \gamma_1 : \mathbb{R} \to X \) be two geodesics on \( X \). When there exists a constant \( C > 0 \) such that

\[
d(\gamma(t), \gamma_1(t)) < C \quad (\forall t \geq 0),
\]

we say that \( \gamma \) and \( \gamma_1 \) are asymptotically equivalent and write as \( \gamma \sim \gamma_1 \).

The relation \( \sim \) gives rise to an equivalence relation on the space \( \text{Geo}(X) \) of all geodesic rays \( \gamma : [0, \infty) \to X \). We call the quotient space \( \text{Geo}(X)/\sim \) the ideal boundary of \( X \), denoted by \( \partial X \). An equivalence class represented by \( \gamma \in \text{Geo}(X) \) is called an asymptotic class and denoted by \([\gamma]\) or \(\gamma(\infty)\).

We denote by \( S_x X \) the set of all unit tangent vectors of \( X \) at \( x \) and define a map \( \beta_x : S_x X \to \partial X \) by \( \beta_x(v) := [\gamma] \), where \( \gamma(t) := \exp_x tv \). Here \( \exp_x \) is the exponential map of \( X \) at \( x \). Since \( X \) is of non-positive sectional curvature, \( \beta_x \) is bijective.

We can define a topology on \( X \cup \partial X \), called the cone topology by setting a fundamental system of neighborhoods ([5, 13]). The map \( \beta_x : S_x X \to \partial X \) is a homeomorphism with respect to the restriction of the cone topology.

Let \( d\theta \) be a probability measure given by the normalized standard volume element of \( S_x X \cong S^{n-1} \). By the aid of the push-forward by the homeomorphism \( \beta_x : S_x X \to \partial X \), \( d\theta \) induces a probability measure on \( \partial X \). We denote this probability measure by the same symbol \( d\theta \).

**Definition 4.3.** A Hadamard manifold \( X \) is said to satisfy the visibility axiom, if the following holds ([13]): for any \( \theta, \theta_1 \in \partial X \), \( \theta \neq \theta_1 \), there exists a geodesic \( \gamma : \mathbb{R} \to X \) satisfying \([\gamma] = \theta \) and \([\gamma^{-1}] = \theta_1 \), where \( \gamma^{-1} \) is the geodesic with inverse direction \( \gamma^{-1}(t) := \gamma(-t) \).

**Remark 4.4.** Rank one symmetric spaces of non-compact type, including real hyperbolic spaces, and Damek-Ricci spaces satisfy the visibility axiom.

We define the Busemann function that appears in defining barycenters. Let \( \gamma : \mathbb{R} \to X \) be a geodesic. Then, we define a function \( b_t : X \to \mathbb{R} \), \( t \geq 0 \) by

\[
b_t(x) := d(x, \gamma(t)) - t = d(x, \gamma(t)) - d(\gamma(0), \gamma(t)),
\]

where \( d \) is the distance function on \( X \). Since \( X \) is a Hadamard manifold, for any \( x \in X \) there exists a limit \( \lim_{t \to \infty} b_t(x) \), denoted by \( b_\infty(x) \). We call this correspondence \( x \mapsto b_\infty(x) \) the Busemann function and denote it by \( B_\gamma \).

A level set of the Busemann function \( B_\gamma \) passing through \( x_0 \in X \), \( \{x \in X \mid B_\gamma(x) \equiv B_\gamma(x_0)\} \) is called a horosphere centered at \( \gamma(\infty) \), which is considered as a limit surface of geodesic spheres centered at \( \gamma(t) \) with radius \( d(\gamma(t), x_0) \), \( t \to \infty \) ([31]).

**Example 4.5.** The real hyperbolic plane \( \mathbb{RH}^2 \) is represented by a unit disk model in the complex plane \( X = \mathbb{RH}^2 = \{z \in \mathbb{C} \mid |z| < 1\} \) and its ideal boundary by \( \partial X = \{z \in \mathbb{C} \mid |z| = 1\} \). Let \( \gamma = \gamma(t) \) be a geodesic satisfying \( \gamma(0) = 0 \), \( [\gamma] = e^{t\varphi} \). Then, the Busemann function is given by the form \( B_\gamma(z) = \log(|z - e^{t\varphi}|^2/(1 - |z|^2)) \) (see [21]).

In what follows, we choose a reference point \( x_0 \in X \) and fix it.

**Definition 4.6.** For any \( \theta \in \partial X \) let \( \gamma \) be a geodesic satisfying \( \gamma(0) = x_0 \) and \([\gamma] = \theta \). We denote by \( B_\theta \) the Busemann function \( B_\gamma \) associated with \( \gamma \) and call it the normalized Busemann function with base point \( x_0 \).
Basic properties of normalized Busemann functions are summarized as follows:

(i) $B_\theta(x_0) = 0$ for any $\theta \in \partial X$.
(ii) $B_\theta(\gamma(t)) = -t$, where $\gamma$ is a geodesic satisfying $\gamma(0) = x_0$ and $[\gamma] = \theta$.
(iii) $B_\theta$ is Lipschitz continuous. In fact, $|B_\theta(x) - B_\theta(y)| \leq d(x,y)$ for any $x, y \in X$.
(iv) $B_\theta$ is of $C^2$-class ([19]).
(v) $B_\theta$ admits the gradient vector field $\nabla B_\theta$ of $C^1$-class with unit norm $|\nabla B_\theta| \equiv 1$.
(vi) For any unit tangent vector $u \in S_x X$, there exists a unique $\theta \in \partial X$ such that $u = (\nabla B_\theta)_x$.
(vii) $B_\theta$ is convex.
(viii) the Hessian $\nabla dB_\theta$ is positive semi-definite

\begin{equation}
(\nabla dB_\theta)_x(u, u) \geq 0 \quad (\forall u \in T_x X, \forall x \in X).
\end{equation}

Moreover, the Hessian satisfies $\nabla dB_\theta(\nabla B_\theta, \cdot) = 0$.
(ix) $\gamma \sim \gamma_1 \iff B_{\gamma_1}(-) - B_{\gamma_1}(·) \equiv \text{const}$ ([39]).

Remark 4.7. From (iv) and (viii) $\Delta B_\theta := -\text{trace} \nabla dB_\theta \leq 0$ holds for any $\theta \in \partial X$.

Example 4.8. The Hessian $\nabla dB_\theta$ of the real hyperbolic space $\mathbb{RH}^n$ is given by the form

\begin{equation}
(\nabla dB_\theta)_x(u, v) = \langle u, v \rangle - \langle u, (\nabla B_\theta)_x \rangle \langle v, (\nabla B_\theta)_x \rangle, \quad u, v \in T_x X, \ x \in X
\end{equation}

(see [7]).

Let $S = S_{\theta, x}$ be the shape operator of a horosphere centered at $\theta \in \partial X$ which passes through $x \in X$. Since the gradient vector field $\nabla B_\theta$ is a unit normal vector field of the horosphere, we have $(\nabla dB_\theta)_x(u, v) = -\langle S_{\theta, x}(u), v \rangle$, where $u, v$ are vectors tangent to the horosphere (see [22]).

Let $\gamma$ be a geodesic with $\theta = [\gamma]$ and let $H_t = H_{\gamma(t), \theta}$ be the horosphere centered at $\theta$ passing through $\gamma(t)$. We find that the shape operator $S_t : T_{\gamma(t)} H_t \rightarrow T_{\gamma(t)} H_t$ of $H_t$ at $\gamma(t)$ satisfies the Riccati equation $S'_t + S_t^2 + R_t = 0$ along $\gamma(t)$, $-\infty < t < \infty$ ([25]), where $R_t$ is the Jacobi operator defined by the Riemannian curvature tensor.

Let $\phi : X \rightarrow X$ be an isometry of a Hadamard manifold $X$. Then, $\phi$ induces a transformation $\hat{\phi}$ of the ideal boundary $\partial X$ as follows:

\begin{equation}
\hat{\phi}(\theta) := [\phi \circ \gamma] \quad (\forall \theta \in \partial X),
\end{equation}

where $\gamma$ is a geodesic satisfying $\gamma(0) = x_0$, and $[\gamma] = \theta$. Then, we obtain the following.

Theorem 4.9 (Busemann cocycle formula [17]). For any $x \in X$

\begin{equation}
B_\theta(\phi(x)) = B_{\hat{\phi}^{-1}(\theta)}(x) + B_\theta(\phi(x_0)),
\end{equation}

where $\phi^{-1} : X \rightarrow X$ is the inverse map of $\phi$.

5. Probability Measures on the Ideal Boundary and Their Barycenter

In this section and the next section, we assume the following.

Hypothesis 5.1. $X$ satisfies the visibility axiom and the normalized Busemann function $\theta \mapsto B_\theta(x)$ is continuous as a function on $\partial X$ for any fixed $x \in X$. 
The real hyperbolic plane satisfies Hypothesis 5.1 (see Example 4.5). Damek-Ricci spaces also satisfy this hypothesis (see [23]).

With respect to this hypothesis we have the following.

Theorem 5.2 ([5]). A Hadamard manifold $X$ satisfies the visibility axiom if and only if for any $\theta \in \partial X$ and any geodesic $\gamma$ satisfying $[\gamma] \neq \theta$ it holds $\lim_{t \to \infty} B_\theta(\gamma(t)) = \infty$.

Definition 5.3. Let $\mu \in \mathcal{P}^+(\partial X)$ be a probability measure on $\partial X$. We define the $\mu$-averaged Busemann function $B_\mu : X \to \mathbb{R}$ by

$$B_\mu(x) := \int_{\theta \in \partial X} B_\theta(x) d\mu(\theta), \quad x \in X.$$ 

Under Hypothesis 5.1, it is shown that the averaged Busemann function $B_\mu : X \to \mathbb{R}$ admits a minimum (see Theorem 5.6). We call this minimal point, i.e., a critical point of $B_\mu$, a barycenter of $\mu$. Before discussing the existence and uniqueness of the minimum of $B_\mu$, we state some of the main properties of averaged Busemann function $B_\mu$.

Proposition 5.4. The averaged Busemann function $B_\mu$ has the following properties:

(i) $B_\mu$ is convex and $B_\mu(x_0) = 0$.

(ii) $B_\mu(\gamma(t)) \to \infty$ ($t \to \infty$), where $\gamma : \mathbb{R} \to X$ is an arbitrary geodesic.

(iii) $B_\mu$ is Lipschitz continuous, i.e., $|B_\mu(x) - B_\mu(y)| \leq d(x, y)$ for any $x, y \in X$.

(iv) For any $\mu \in \mathcal{P}^+(\partial X)$ and any $x \in X$, $|(\nabla B_\mu)_x| \leq 1$ holds.

(v) The Hessian $\nabla d B_\mu$ is represented by

$$\nabla d B_\mu_x(u, v) = \int_{\theta \in \partial X} (\nabla d B_\theta)_x(u, v) d\mu(\theta) \quad u, v \in T_x X, \quad x \in X$$

and hence is positive semi-definite, provided that $X$ has bounded Ricci curvature and $d \Delta B_\theta$ is uniformly bounded with respect to $\theta$.

Remark 5.5. If $X$ is a Hadamard manifold of volume entropy $Q \geq 0$ which is asymptotically harmonic, then one has $d \Delta B_\theta \equiv 0$ and by using the Riccati equation $-(n - 1)Q^2 \leq \text{Ric}_x \leq 0$, while the detailed argument is omitted. For the volume entropy and the asymptotical harmonicity refer to Definition 6.7 and Remark 6.9, §6, respectively.

Proof.

(i) Because $B_\mu$ is an average of a convex function, $B_\mu$ is also convex. It is obvious that $B_\mu(x_0) = 0$.

(ii) This will be shown later (see the proof of Theorem 5.6).

(iii) This is obvious, since $B_\theta$ is Lipschitz continuous and satisfies $|B_\theta(x) - B_\theta(y)| \leq d(x, y)$ for any $x, y \in X$.

(iv) First, we show that $B_\mu$ admits a gradient vector field. Let $v$ be an arbitrary tangent vector at $x \in X$ and take a geodesic $\sigma = \sigma(t)$ satisfying $\sigma(0) = x$ and $\dot{\sigma}(0) = v \in T_x X$. We remark that $v \in T_x X$ is not necessarily a unit vector. Since $\nabla B_\theta$ is uniformly bounded ($|\nabla B_\theta| \equiv 1$), we may interchange the order of integration of $B_\theta(\sigma(t))$ with respect to $\mu \in \mathcal{P}^+(\partial X)$ and differentiation with respect to $t$. Then,
the directional derivative to $v$ of $\mathbb{B}_\mu$ is given by
\begin{align}
v\mathbb{B}_\mu &= \frac{d}{dt} \bigg|_{t=0} \int_{\theta \in \partial X} B_\theta(\sigma(t)) d\mu(\theta) \\
&= \int_{\theta \in \partial X} \frac{\partial}{\partial t} \bigg|_{t=0} B_\theta(\sigma(t)) d\mu(\theta) = \int_{\theta \in \partial X} \langle (\nabla B_\theta)_x, v \rangle d\mu(\theta),
\end{align}
from which we find that $\mathbb{B}_\mu$ is of $C^1$-class. This implies that the gradient vector field $\nabla \mathbb{B}_\mu$ is well-defined on $X$ and satisfies
\begin{align}
\langle (\nabla \mathbb{B}_\mu)_x, v \rangle = \int_{\partial X} \langle (\nabla B_\theta)_x, v \rangle d\mu(\theta).
\end{align}
Letting $v = (\nabla \mathbb{B}_\mu)_x$ in (5.3), we have
\begin{align}
|\langle (\nabla \mathbb{B}_\mu)_x, v \rangle| \leq \int_{X} |\langle (\nabla B_\theta)_x, v \rangle| d\mu(\theta) = |\langle (\nabla \mathbb{B}_\mu)_x, v \rangle|,
\end{align}
from which (iv) is shown.

(v) In general, the Hessian of a function $f : X \rightarrow \mathbb{R}$ is given by
\[ \nabla df(v, v) = \frac{d^2}{dt^2} \bigg|_{t=0} f(\sigma(t)), \]
where $\sigma$ is a geodesic satisfying $\sigma(0) = x \in X$ and $\dot{\sigma}(0) = v$. Therefore, to define the Hessian of the $\mu$-averaged Busemann function, it suffices that the order of integration with respect to $\mu \in \mathcal{P}^+(\partial X)$ and differentiation with respect to $t$ in (5.5), the equation below which is obtained from (5.2), is interchangeable:
\begin{align}
\frac{d^2}{dt^2} \bigg|_{t=0} \mathbb{B}_\mu(\sigma(t)) = \frac{d}{dt} \bigg|_{t=0} \int_{\theta \in \partial X} \langle (\nabla B_\theta)_{\sigma(t)}, \dot{\sigma}(t) \rangle d\mu(\theta),
\end{align}
so that it suffices to show that
\[ (\nabla dB_\theta)_x(v, v) = \frac{d^2}{dt^2} \bigg|_{t=0} B_\theta(\sigma(t)) \]
is uniformly bounded as a function of $\theta \in \partial X$. From Bochner’s formula ([16, Proposition 4.15]), uniform boundedness of $|\nabla dB_\theta|$ is asserted under the assumption of the boundedness of the Ricci curvature of $X$ and the uniform boundedness of $d\Delta B_\theta$. Therefore, one obtains (5.1).

**Theorem 5.6.** Let $X$ be a Hadamard manifold satisfying Hypothesis 5.1. Then, an arbitrary probability measure $\mu \in \mathcal{P}^+(\partial X)$ admits a barycenter.

**Remark 5.7.** In [7], Besson et al. show the existence of barycenter. They assume that the Hadamard manifold $X$ is a rank one symmetric space of non-compact type, for which Hypothesis 5.1 is satisfied. They obtain Theorem 5.6 for general probability measures with no atom.

We will outline a proof of Theorem 5.6. For a given constant $C > 0$, we set $A_C := \{ y \in X \mid \mathbb{B}_\mu(y) \leq C \}$. It is seen that $x_0 \in A_C$, because $\mathbb{B}_\mu(x_0) = 0$. Hence, we find that $A_C$ is a non-empty closed subset of $X$.

Now we show that $A_C$ is bounded. Since $\mathbb{B}_\mu$ is convex, $A_C$ is a convex set.

Choose $\theta \in \partial X$ arbitrarily and fix it. Let $\gamma$ be a geodesic on $X$ satisfying $\gamma(0) = x_0$ and $|\gamma| = \theta$. Then we will show that the convex function $\mathbb{B}_\mu$ satisfies
\[ \lim_{t \rightarrow \infty} \mathbb{B}_\mu(\gamma(t)) = +\infty, \]
as follows.
Since $B_θ$ is convex and $B_θ(x_0) = 0$, it holds that for any geodesic $σ$ through $x_0$ at $t = 0$ ($σ(0) = x_0$),

\[ t_1 B_θ(σ(t)) ≥ t B_θ(σ(t_1)) \quad (0 ≤ t_1 ≤ t) \quad ∀θ ∈ \partial X. \]

Moreover, since the $μ$-averaged Busemann function $B_μ$ is also convex, from Proposition 5.4 (i), it holds similarly as (5.6)

\[ t_1 B_μ(σ(t)) ≥ t B_μ(σ(t_1)) \quad (0 ≤ t_1 ≤ t) \quad ∀μ ∈ \mathcal{P}^+(\partial X). \]

Next, we choose an arbitrary ideal point $θ_0 ∈ \partial X$ and fix it. Let $γ_0$ be a geodesic satisfying $γ_0(0) = x_0$ and $|γ_0| = θ_0$. For a positive number $t$, we set a subset $J_{θ_0}(t)$ of $\partial X$ by $J_{θ_0}(t) := \{ θ ∈ \partial X \mid B_θ(γ_0(t)) ≤ 0 \}$. From the assumption of the theorem, that is, from Hypothesis 5.1, $B_θ(x)$ is continuous as a function of $θ$ so that $J_{θ_0}(t)$ is a compact subset of $\partial X$. Obviously $θ_0 ∈ J_{θ_0}(t)$, since $B_{θ_0}(γ_0(t)) = -t$.

**Lemma 5.8.** There exists $t_1 ∈ (0, ∞)$ such that $μ(J_{θ_0}(t_1)) < 1$.

In fact, from (5.6), we have $J_{θ_0}(t) ⊂ J_{θ_0}(t_1)$, if $0 ≤ t_1 ≤ t$. We have from Theorem 5.2, i.e., from the condition equivalent to the visibility axiom (5.8)

\[ \lim_{t \to ∞} \mu(J_{θ_0}(t)) = \mu \left( \bigcap_{t \in [0, ∞)} J_{θ_0}(t) \right) = \mu(\{ θ_0 \}) = 0, \]

from which we obtain Lemma 5.8. We notice that if $t ≥ t_1$, then $μ(J_{θ_0}(t)) ≤ μ(J_{θ_0}(t_1)) < 1$ holds.

We take a compact subset $K ⊂ \partial X \setminus J_{θ_0}(t_1)$ which satisfies $μ(K) > 0$. In fact, we can choose such a subset $K$ because $μ(\partial X \setminus J_{θ_0}(t_1)) > 0$. Since $B_θ(γ(t)) > 0$ holds for any $θ$ in $\partial X \setminus J_{θ_0}(t_1)$, we have

\[
\int_{θ ∈ \partial X} B_θ(γ(t))dμ(θ) = \left( \int_{θ ∈ J_{θ_0}(t)} + \int_{θ ∈ \partial X \setminus J_{θ_0}(t)} \right) B_θ(γ(t))dμ(θ) ≥ \int_{θ ∈ J_{θ_0}(t)} B_θ(γ(t))dμ(θ) + \int_{θ ∈ K} B_θ(γ(t))dμ(θ).
\]

Now we fix $t_1 > 0$ in our lemma. Since the subset $K$ is compact, we can retake the constant $C > 0$ such that

\[ B_θ(γ(t_1)) ≥ C > 0 \quad (∀θ ∈ K). \]

From (5.6), (5.9) and (5.10), we have

\[ \int_{θ ∈ \partial X} B_θ(γ(t))dμ(θ) ≥ \frac{t}{t_1} \int_{θ ∈ J_{θ_0}(t)} B_θ(γ(t))dμ(θ) + C \frac{t}{t_1} μ(K). \]

To estimate the first term of the right hand side of (5.11), we set $D := \sup \{|B_θ(γ(t_1))| \mid θ ∈ \partial X\}$. Since $\partial X$ is compact, the continuous function $θ → B_θ(γ(t_1))$ is bounded as a function of $θ$. Hence, $D < ∞$ is asserted. From (5.11), we get the following:

\[ B_μ(γ(t)) = \int_{θ ∈ \partial X} B_θ(γ(t))dμ(θ) ≥ \frac{t}{t_1} (Cμ(K) - Dμ(J_{θ_0}(t))). \]

Therefore, we obtain from (5.8) $\lim_{t \to ∞} B_μ(γ(t)) = +∞$.

Now, we suppose that $A_C$ be not bounded. Then, we can take a sequence $\{ y_i \}$ of points in $A_C$ such that $d_i = d(y_i, x_0) → +∞$, $i → +∞$ ($d_i < d_j$, $i < j$). Therefore, there exists a sequence $\{ v_i \}$ of unit tangent vectors such that $y_i = \exp_{x_0} d_i v_i$. Take
an appropriate subsequence of \( \{v_i\} \), denoted by the same letter \( \{v_i\} \) for brevity and set \( \lim_i v_i = v_\infty \). A geodesic \( \gamma_\infty(t) \) given by \( \gamma_\infty(t) = \exp_{x_0} tv_\infty \) satisfies \( \mathbb{B}_\mu(\gamma_\infty(t)) = \lim_{t \to \infty} \mathbb{B}_\mu(\exp_{x_0} tv_i), \ t > 0 \). Since for any \( t > 0 \) there exists \( i_0 > 0 \) such that \( d_i \geq t \) for any \( i \geq i_0 \), we have from (5.7)
\[
(5.13) \quad \mathbb{B}_\mu(\exp_{x_0} tv_i) \leq \frac{t}{d_i} \mathbb{B}_\mu(\exp_{x_0} d_i v_i) = \frac{t}{d_i} \mathbb{B}_\mu(y_i) \leq \frac{t}{d_i} C \leq C,
\]
a contradiction, since \( \mathbb{B}_\mu(\gamma_\infty(t)) \to \infty (t \to \infty) \). Hence, \( A_C \) is bounded. Since \( A_C \) is bounded and closed, \( \mathbb{B}_\mu \) has a minimum value on \( A_C \), i.e., \( \mu \) has a barycenter.

With respect to the uniqueness of barycenter, one has the following.

**Theorem 5.9.** Let \( X \) be a Hadamard manifold satisfying Hypothesis 5.1. If for any \( \mu \in \mathcal{P}^+(\partial X) \) the Hessian of \( \mu \)-averaged Busemann function is positive definite everywhere on \( X \), then there exists uniquely a barycenter of \( \mu \).

For the positive definiteness of the Hessian of \( \mu \)-averaged Busemann function, we have the following.

**Theorem 5.10.** Assume that every averaged Busemann function admits its Hessian in the form (5.1) in (v), Proposition 5.4.

(i) If the Ricci curvature of \( X \) is negative everywhere, i.e., \( \text{Ric}_x < 0, \forall x \in X \), then for any \( \mu \in \mathcal{P}^+(\partial X) \) the Hessian \( \nabla \mathbb{B}_\mu \) of the \( \mu \)-averaged Busemann function is positive definite everywhere on \( X \).

(ii) If there exists a certain probability measure \( \mu_0 \in \mathcal{P}^+(\partial X) \) such that the Hessian \( \nabla \mathbb{B}_{\mu_0} \) of \( \mathbb{B}_{\mu_0} \) is positive definite on \( X \), the Hessian \( \nabla \mathbb{B}_\mu \) of the \( \mu \)-averaged Busemann function \( \mathbb{B}_\mu \) is also positive definite for any \( \mu \in \mathcal{P}^+(\partial X) \).

**Proof.** We will give an outline of a proof of (i). Assume that there exists a \( \mu \in \mathcal{P}^+(\partial X) \) such that for \( x \in X \) and \( u(\neq 0) \in T_x X \), \( (\nabla \mathbb{B}_{\mu})_x(u, u) = 0 \) holds. Since the Hessian \( \nabla \mathbb{B}_\mu \) in (5.1) is positive semi-definite for any \( \theta \in \partial X \), we have then \( (\nabla \mathbb{B}_\mu)_x(u, u) = 0 \). Hence we see \( \langle S_{\theta,x} u, u \rangle = 0 \). Here \( S_{\theta,x} \) is the shape operator of the horosphere \( H_{x,\theta} \) centered at \( \theta \) passing through \( x \in X \). Since \( S_{\theta,x} \) is negative semi-definite, we have \( \langle S_{\theta,x} u, v \rangle = 0 \) for any \( v \in T_x X \). In fact, if we assume that there exists a tangent vector \( v \) such that \( \langle S_{\theta,x} u, v \rangle \neq 0 \), \( \langle S_{\theta,x} (u + tv), (u + tv) \rangle \) takes a positive value for some \( t \neq 0 \), which is a contradiction, since \( S_{\theta,x} \) is negative semi-definite. Hence, we have \( S_{\theta,x} u = 0 \). From the Ricci equation for \( S_{\theta} \), we can show that the sectional curvature of any 2-plane which contains \( u \) vanishes so that we obtain \( \text{Ric}_x(u) = 0 \). This is a contradiction, since the Ricci curvature is assumed to be negative. Hence we obtain (i).

To show (ii) we set \( C = \min \{ (d\mu_i/d\mu_0)(\theta) | \theta \in \partial X \} \). Then \( C > 0 \) from the compactness of \( \partial X \). From the positive semi-definiteness of the Hessian \( \nabla \mathbb{B}_\mu \) we find that \( (\nabla \mathbb{B}_{\mu})_x(u, u) \geq C(\nabla \mathbb{B}_{\mu_0})_x(u, u), \forall u \in T_x X, \forall x \in X \), from which (ii) is obtained.

**Corollary 5.11.** Under Hypothesis 5.1, namely, the assumption of Theorem 5.6 and either (i) or (ii) of Theorem 5.10, an arbitrary probability measure \( \mu \in \mathcal{P}^+(\partial X) \) has a unique barycenter.

**Remark 5.12.** The barycenter of the standard probability measure \( d\theta \in \mathcal{P}^+(\partial X) \) is the base point \( x_0 \in X \) (see [27]).
6. Barycenter maps and barycenterically associated maps

When there exists a unique barycenter $x = \text{bar}(\mu) \in X$ for any $\mu \in \mathcal{P}^+(\partial X)$, we can define a map $\text{bar} : \mathcal{P}^+(\partial X) \rightarrow X; \mu \mapsto x$ and call it barycenter map.

We notice that the differential map $(d\text{bar})_\mu : T_\mu \mathcal{P}(\partial X) \rightarrow T_x X$, $x = \text{bar}(\mu)$ of the barycenter map bar is well defined and is surjective ([22]).

**Definition 6.1.** Let $\text{bar}^{-1}(x)$ be the inverse image of the map bar at a point $x$. For any $\mu \in \text{bar}^{-1}(x)$ we define a map $\nu^\mu_x : T_x X \rightarrow T_\mu \mathcal{P}^+(\partial X)$ by

$$\nu^\mu_x(u)(\theta) := (dB_\theta)_x(u) d\mu(\theta) \quad (u \in T_x X, \ \theta \in \partial X).$$

(see [28]).

Notice that the image of the map $\nu^\mu_x$ is included in the tangent space $T_\mu \mathcal{P}^+(\partial X)$ at $\mu$, since $x$ is the barycenter of $\mu$ if and only if

$$(d\mathbb{B}_\mu)_x(u) = \int_{\theta \in \partial X} (dB_\theta)_x(u) d\mu(\theta) = 0$$

holds for any $u \in T_x X$. Moreover, we find that $\nu^\mu_x$ is injective ([22]).

Let $\mu(t)$ be a curve in $\text{bar}^{-1}(x)$, $|t| < \varepsilon$ satisfying $\mu(0) = \mu$ and $\dot{\mu}(0) = \tau \in T_\mu \mathcal{P}^+(\partial X)$. Then, $\tau$ fulfills $\int_{\theta \in \partial X} (dB_\theta)_x(u)d\tau(\theta) = 0$ for any $u \in T_x X$, i.e., $G_\mu(\tau, \nu^\mu_x(u)) = 0$ with respect to the Fisher metric $G$, from which we find that tangent vectors of $\text{bar}^{-1}(x)$ are orthogonal to the image of $\nu^\mu_x$. The following proposition indicates that $\text{bar} : \mathcal{P}^+(\partial X) \rightarrow X$ admits a fiber space structure in a tangent space level.

**Proposition 6.2.** Let $x \in X$. For each $\mu \in \text{bar}^{-1}(x)$, $T_\mu \mathcal{P}^+(\partial X)$ is decomposed into the $G_\mu$-orthogonally direct sum:

$$T_\mu \mathcal{P}^+(\partial X) = T_\mu \text{bar}^{-1}(x) \oplus \text{Im}{\nu^\mu_x}.$$  

**Remark 6.3.** $T_\mu \text{bar}^{-1}(x) \subset T_\mu \mathcal{P}^+(\partial X)$ is the vertical subspace of the fiber space structure. On the other hand, $\text{Im}{\nu^\mu_x}$ is the horizontal subspace, normal to the fibers, which means that the barycenter map bar satisfies an infinitesimal trivialization. This proposition suggests that the barycenter map bar itself satisfies a local trivialization.

Now we are going to see geometric properties of a barycenter map.

**Proposition 6.4.** Let $\phi$ be an isometry of $X$ and $\hat{\phi} : \partial X \rightarrow \partial X$ be a homeomorphism of $\partial X$ induced by $\phi$ (see (4.4)). Let $\hat{\phi}_\mu : \mathcal{P}^+(\partial X) \rightarrow \mathcal{P}^+(\partial X)$ be the push-forward of $\hat{\phi} : \partial X \rightarrow \partial X$. Then, we have for any $\mu \in \mathcal{P}^+(\partial X)$ $\text{bar}(\hat{\phi}_\mu \mu) = \phi(\text{bar}(\mu))$, i.e., $\text{bar} \circ \hat{\phi}_\mu = \phi \circ \text{bar}$ holds.

In fact, by integrating both side of the Busemann cocycle formula (4.5) with respect to $\mu$, we obtain the averaged-Busemann cocycle formula

$$\mathbb{B}_\mu(\phi^{-1}x) = \mathbb{B}_{\hat{\phi}_\mu}(x) + \mathbb{B}_\mu(\phi^{-1}x_0), \quad \forall \mu \in \mathcal{P}^+(\partial X),$$

from which our proposition is obtained.

Poisson kernel probability measures are measures significantly important in considering the barycenter maps. As mentioned in §1, the Poisson kernel $P(x, \theta)$ is the fundamental solution of the Dirichlet problem at infinity. We define the Poisson kernel based on the argument of Shoen-Yau given in [41] as follows:
Definition 6.5 ([41]). Let \( x_0 \in X \) be a fixed point and let \( P(x, \theta) \) be a function on \( X \times \partial X \). Then, \( P(x, \theta) \) is called the Poisson kernel, normalized at a base point \( x_0 \), when it satisfies the following conditions:

1. \( \Delta P(\cdot, \theta) = 0 \quad \forall \theta \in \partial X. \)
2. \( P(x_0, \theta) = 1 \quad \forall \theta \in \partial X. \)
3. \( P(x, \theta) d\theta \in \mathcal{P}^+(\partial X) \quad \forall x \in X. \)
4. For any fixed \( \theta \in \partial X \), the function on \( X, x \mapsto P(x, \theta) \) is extended to a continuous function on \( X \cup \partial X \setminus \{\theta\} \) and satisfies \( \lim_{x \to \theta'} P(x, \theta) = 0 \) for any \( \theta' \) satisfying \( \theta' \neq \theta \).

For any \( f \in C^0(\partial X) \), the function \( u(x) = \int_{\theta \in \partial X} P(x, \theta) f(\theta) d\theta \) is a solution to \( \Delta u = 0 \) and satisfies the boundary condition at infinity \( u|_{\partial X} = f \) (see [41, 23]).

Remark 6.6. Under the negative sectional curvature condition \(-b^2 \leq K \leq -a^2 < 0\), the existence and uniqueness of the Poisson kernel is guaranteed (see [41, 3]). Refer also to [2, 44] for the Dirichlet problem at infinity.

Definition 6.7. If the Poisson kernel \( P(x, \theta) \) is represented in the form \( P(x, \theta) = \exp(-Q_B \theta(x)) \), we call it the Busemann–Poisson kernel, where \( Q = Q(X) := \lim_{r \to \infty} \frac{1}{r} \log \text{Vol}(B(x, r)) \) is a constant, called the volume entropy of \( X \) which represents exponential volume growth of geodesic spheres.

Remark 6.8. The family of all Damek-Ricci spaces, which is a class of Hadamard manifolds, includes rank one symmetric spaces of non-compact type. Any Damek-Ricci space carries the Busemann–Poisson kernel. For this see [23].

Remark 6.9. It is necessary for a Hadamard manifold \( X \) to carry the Busemann–Poisson kernel that \( X \) is asymptotically harmonic, i.e., any horosphere has a constant mean curvature \(-Q\). Moreover, from Definition 6.5, (iii) and Theorem 5.2, the Hadamard manifold \( X \) also satisfies the visibility axiom.

Proposition 6.10. If the Poisson kernel \( P(x, \theta) \) is the Busemann–Poisson kernel, in particular, then the barycenter of the probability measure \( \mu_x := P(x, \theta) d\theta \in \mathcal{P}^+(\partial X) \) defined by \( P(x, \theta) \), parametrized by \( x \in X \) is the point \( x \). Hence, \( \text{bar} : \mathcal{P}^+(\partial X) \to X \) is surjective (see [7, 22]).

Theorem 6.11 ([22]). For the Busemann–Poisson kernel probability measure \( \mu_x \), \( x \in X \), the Hessian of \( \mu_x \)-averaged Busemann function \( (\nabla d_{\mu_x})(y), y \in X \) is positive definite.

In fact, we have

\[
(\nabla d_{\mu_x})(u, u) = Q G_{\mu_x}(v^{\mu_x}(u), v^{\mu_x}(u)), \quad u \in T_x X, \quad x \in X.
\]

From an argument similar to the proof of Theorem 5.10 (ii), we have for any \( y \in X \)

\[
(\nabla d_{\mu_y})(u, u) > 0, \quad u \in T_x X, \quad u \neq 0, \quad x \in X.
\]

From the above theorem, if a Hadamard manifold \( X \) carrying the Busemann–Poisson kernel satisfies the assumption of Theorem 5.6, namely Hypothesis 5.1, then, the function \( \theta \mapsto B_\theta(\cdot) \) is continuous on \( \partial X \) so that from Theorem 5.10 (ii), any \( \mu \in \mathcal{P}^+(\partial X) \) has a unique barycenter.

Remark 6.12. Let \( \mu, \mu_1 \in \text{bar}^{-1}(x) \). Then the path \( \mu(t) := (1-t)\mu + t\mu_1 \) belongs to \( \text{bar}^{-1}(x) \) for any \( t \in [0, 1] \). The curve \( t \mapsto \mu(t), t \in [0, 1] \) yields a geodesic with respect to the \( m \)-connection \( \nabla^{(-1)} \).
Proposition 6.13. [22] Let \( t \mapsto \mu(t) \in \mathcal{P}^+(\partial X) \) be a geodesic with respect to the Levi-Civita connection. Then, for any \( x \in X \) and \( \mu(t) \in \text{bar}^{-1}(x) \) if and only if the following are fulfilled:

(i) \( \mu(0) \in \text{bar}^{-1}(x) \).

(ii) \( \dot{\mu}(0) \in T_{\mu(0)}\text{bar}^{-1}(x) \).

(iii) \( h_{\mu(0)}(\dot{\mu}(0), \dot{\mu}(0)) = 0 \), where \( h \) is the second fundamental form of the submanifold \( \text{bar}^{-1}(x) \),

\[
\begin{align*}
h_{\mu} : T_{\mu}\text{bar}^{-1}(x) \times T_{\mu}\text{bar}^{-1}(x) &\to N_{\mu} = \text{Im}\nu_{\mu}^x; \\
(\alpha, \beta) &\mapsto (\nabla_{\alpha}\beta)^\perp.
\end{align*}
\]

Here \((\nabla_{\alpha}\beta)^\perp\) is the normal component of a vector \(\nabla_{\alpha}\beta\).

Remark 6.14. The barycenter of the standard probability measure \( \lambda = \text{d}\theta \) is the base point \( x_0 \) (Remark 5.12). By the identification \( \partial X \cong S_{x_0}X \), we set \( \tau = q(v)\text{d}\theta \in T_{\partial X}\mathcal{P}^+(\partial X) \) for an arbitrary unit tangent vector \( v \in S_{x_0}X \) by \( q(v) = v^i v^i \), \( i, j = 1, \ldots, n, i \neq j \). Here \( \{v^i, 1 \leq i \leq n \} \) is the components of \( v \) with respect to a certain orthonormal basis \( \{e_i\} \) of \( T_{x_0}X \); \( v = \sum_{i=1}^n v^i e_i \). Since \( \tau \) satisfies (ii) and (iii) of Proposition 6.13, the geodesic with initial velocity vector \( \tau \) belongs to \( \text{bar}^{-1}(x_0) \).

Theorem 6.15. Let \( \mu, \mu_1 \in \text{bar}^{-1}(x), x \in X \). A geodesic with respect to the Levi-Civita connection joining \( \mu \) and \( \mu_1 \) is contained in \( \text{bar}^{-1}(x) \) if and only if \( \sigma(\mu, \mu_1) \in \text{bar}^{-1}(x) \) holds. Here \( \sigma(\mu, \mu_1) \) is the normalized geometric mean (see Definition 2.12 in §2.2).

Theorem 6.15 is immediate from Theorem 2.13, which expresses the representation formula for a geodesic in terms of the normalized geometric mean. See [29] for details.

Let \( \Phi : \partial X \to \partial X \) be a homeomorphism and \( \Phi_2 : \mathcal{P}^+(\partial X) \to \mathcal{P}^+(\partial X) \) be the push-forward by \( \Phi \). We recall that \( \Phi_2 \) is isometric with respect to the Fisher metric \( G \) (see Theorem 2.4 in §2.2).

Definition 6.16 ([22]). Let \( \phi : X \to X \) be a bijective map. We call \( \phi \) a map barycentrically associated to \( \Phi \), when \( \phi \) satisfies \( \text{bar} \circ \Phi_2 = \phi \circ \text{bar} \).

Let \( \phi \) be an isometry of \( X \) and \( \tilde{\phi} : \partial X \to \partial X \) be the homeomorphism induced by \( \phi \). Then, \( \phi \) is a map barycentrically associated to \( \tilde{\phi} \) (see Proposition 6.4).

Definition 6.17. We define a map \( \Theta : X \to \mathcal{P}^+(\partial X) \) by \( \Theta(x) := \mu_x(= P(x, \theta)\text{d}\theta) \), \( x \in X \). We call this map the Poisson kernel map.

In what follows, we assume that a Hadamard manifold \( X \) satisfies Hypothesis 5.1 and carries the Busemann-Poisson kernel. Then, we recall that \( \text{bar} \circ \mu_x = x \) (see Proposition 6.10). Hence, the map \( \Theta \) satisfies \( \text{bar} \circ \Theta = \text{id}_X \), i.e., \( \Theta \) is a section of the projection \( \text{bar} : \mathcal{P}^+(\partial X) \to X \). Its differential map satisfies \((d\Theta)_x = -Q\nu_x^\perp\) for any \( x \in X \).

We can show that any isometry \( \varphi \) of \( X \) satisfies \( \Theta \circ \varphi = \tilde{\varphi} \circ \Theta \). In relation with this fact, we have the following theorem.

Theorem 6.18 ([26, 27, 28]). Let \( \Phi : \partial X \to \partial X \) be a homeomorphism of \( \partial X \) and \( \phi : X \to X \) be a bijective \( C^1 \)-map which is barycentrically associated to \( \Phi \). If the maps \( \Phi \) and \( \phi \) admit a relation \( \Theta \circ \phi = \Phi_2 \circ \Theta \), then \( \phi \) turns out to be an isometry of \( X \). Moreover, the transformation \( \tilde{\phi} \) of \( \partial X \) canonically induced by \( \phi \) coincides with \( \Phi \).
Remark 6.19. If we assume singly $\Theta \circ \phi = \Phi \circ \Theta$, it follows then that $\mu_{\phi(x)} = \Theta(\phi(x)) = \Phi(\mu_x)$ and hence $\phi(x) = \bar{\mu}(\mu_{\phi(x)}) = \bar{\mu}(\Phi_x) \mu_x$ from which we find that the barycenter of $\Phi_x \mu_x$ is $\phi(x)$.

Remark 6.20. In general, we define for a homeomorphism $\Phi : \partial X \to \partial X$ a map $\phi : X \to X$ by $\phi(x) := \bar{\mu}(\Phi_x) \mu_x$, $x \in X$. It is shown that when $X$ is a real hyperbolic space, the map $\phi$ is of $C^1$-class and satisfies

$$\text{Jac} \phi_x := \sqrt{\det((d\phi_x)^*) (d\phi_x)} \leq 1.$$ 

Moreover, equality in the above holds if and only if the differential map $d\phi_x : T_x X \to T_{\phi(x)} X$ is a linear isometry (see [7, 8, 40, 35]).

We can weaken the assumption of Theorem 6.18 in the following way.

**Theorem 6.21 ([27, 22]).** Let $\Phi : \partial X \to \partial X$ be a homeomorphism of $\partial X$ and $\varphi : X \to X$ be a map of $C^1$-class which is barycentrically associated to $\Phi$. Suppose that there exists a section $\Sigma : X \to P^+(\partial X)$ of the fiber space structure induced by the barycenter map $\bar{\mu} : P^+(\partial X) \to X$, i.e., a map $\Sigma$ satisfying $\bar{\mu} \circ \Sigma = \text{id}_X$, such that each of the following two diagrams commutes:

\[
\begin{array}{ccc}
P^+(\partial X) & \xrightarrow{\Phi^+} & P^+(\partial X) \\
\Sigma & \uparrow & \Sigma \\
X & \xrightarrow{\varphi} & X
\end{array}
\]

\[
\begin{array}{ccc}
T_{\mu_x} P^+(\partial X) & \xrightarrow{(d\Phi_x)^* \mu_x} & T_{\mu_{\varphi(x)}} P^+(\partial X) \\
\nu_{\mu_x}^\varphi & \uparrow & \nu_{\mu_{\varphi(x)}}^\varphi \\
T_x X & \xrightarrow{d\varphi_x} & T_{\varphi(x)} X
\end{array}
\]

Here $\mu_x := \Sigma(x) \in P^+(\partial X)$. Then, $\varphi$ is an isometry of $X$ and the transformation $\hat{\varphi}$ of $\partial X$ induced by $\varphi$ coincides with $\Phi$.

This theorem is shown by using geometric properties of the normalized Busemann function. Refer to [27] for details. Theorem 6.18 is also obtained as a corollary of this theorem.

**Acknowledgments**

The authors would like to appreciate Professor Naoyuki Koike since he offered an opportunity for writing a manuscript. The authors also wish to thank the referees for their careful refereeing and many useful and valuable comments.

**References**

[1] S. Amari and H. Nagaoka, Methods of Information Geometry, Trans. Math. Monogr. 191, Amer. Math. Soc., Providence RI; Oxford University Press, Oxford, 2000.
[2] M. T. Anderson, The Dirichlet problem at infinity for manifolds of negative curvature, J. Differential Geom. 18 (1983), 701–721.
[3] M. T. Anderson and R. Schoen, Positive harmonic functions on complete manifolds of negative curvature, Ann. of Math. (2) 121 (1985), 429–461.
[4] J.-P. Anker, E. Damek and C. Yacoub, Spherical analysis on harmonic $\mathbb{AN}$ groups, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 23 (1996), 643–679.
[5] W. Ballmann, M. Gromov and V. Schroeder, Manifolds of Nonpositive Curvature, Progr. Math. 61, Birkhäuser, Boston, MA, 1985.
[6] J. Berndt, F. Tricerri and L. Vanhecke, Generalized Heisenberg Groups and Damek-Ricci Harmonic spaces, Lecture Notes in Math. 1598, Springer-Verlag, Berlin, 1995.
[7] G. Besson, G. Courtois and S. Gallot, Entropies et rigidités des espaces localement symétriques de courbure strictement négative (French), Geom. Funct. Anal. 5 (1995), 731–799.
[8] G. Besson, G. Courtois and S. Gallot, Minimal entropy and Mostow’s rigidity theorems, Ergodic Theory Dynam. Systems 16 (1996), 623–649.
[9] P. S. Bullen, Handbook of Means and Their Inequalities, Math. Appl. 560, Kluwer, Dordrecht, 2003.
[10] A. Douady and C. J. Earle, Conformally natural extension of homeomorphisms of the circle, Acta Math. 157 (1986), 23–48.
[11] P. Eberlein, Geometry of Nonpositively Curved Manifolds, Chicago Lectures in Math., University of Chicago Press, Chicago, IL, 1996.
[12] P. Eberlein, Geodesic flows in manifolds of nonpositive curvature, Smooth Ergodic Theory and Its Applications, Seattle, WA, 1999, (eds. A. Katok, R. de la Llave, Y. Pesin and H. Weiss), Proc. Sympos. Pure Math. 69, Amer. Math. Soc., Providence, RI, 2001, 525–571.
[13] P. Eberlein and B. O’Neill, Visibility manifolds, Pacific J. Math. 46 (1973), 45–109.
[14] A. Fathi, Structure of the group of homeomorphisms preserving a good measure on a compact manifold, Ann. Sci. Éc. Norm. Sup. (4) 13 (1980), 45–93.
[15] T. Friedrich, Die Fisher-Information und symplektische Strukturen (German), Math.Nachr. 153 (1991), 273–296.
[16] S. Gallot, D. Hulin and J. Lafontaine, Riemannian Geometry, 2nd ed., Universitext, Springer-Verlag, Berlin, 1990.
[17] Y. Guivarc’h, L. Ji and J. C. Taylor, Compactifications of Symmetric Spaces, Progr. Math. 156, Birkhäuser, Boston, MA. 1998.
[18] P. Halmos, Measure theory, Grad. Texts in Math. 18, Springer, 1976.
[19] E. Heintze and H.-C. Im Hof, Geometry of horospheres, J. Differential Geom. 12 (1977), 481–491.
[20] S. Helgason, Differential Geometry and Symmetric Spaces, Pure Appl. Math. 12, Academic Press, New York, 1962.
[21] S. Helgason, Groups and Geometric Analysis, Pure Appl. Math. 113, Academic Press, New York, 1984.
[22] M. Itoh, Fisher Information Geometry of Barycenter Maps, Lecture Note, Tokyo University of Science, 2015.
[23] M. Itoh and H. Satoh, Information geometry of Poisson kernel on Damek-Ricci spaces, Tokyo J. Math. 33 (2010), 129–144.
[24] M. Itoh and H. Satoh, Fisher information geometry, Poisson kernels and asymptotical harmonicity, Diff. Geom. Appl. 29 (2011), S107–S115.
[25] M. Itoh and H. Satoh, Horospheres and hyperbolic spaces, Kyushu J. Math. 67 (2013), 309–326.
[26] M. Itoh and H. Satoh, Information geometry of barycenter map, Real and Complex Submanifolds, Springer Proc. Math. Stat. 106, Springer, 2014, 79–88.
[27] M. Itoh and H. Satoh, Geometry of Fisher information metric and the barycenter map, Entropy, 17 (2015), 1814–1849.
[28] M. Itoh and H. Satoh, Information geometry of Busemann-barycenter for probability measures, Internat. J. Math. 26 (2015), 1541007.
[29] M. Itoh and H. Satoh, Geometric mean of probability measures and geodesics of Fisher metric, to appear in Math. Nachr.
[30] M. Itoh and H. Satoh, α-connections and duality for space of probability measures, in preparation.
[31] M. Itoh, H. Satoh and Y.J. Suh, —it Horospheres and hyperbolicity of Hadamard manifolds, Diff. Geom. Appl. 35, suppl. (2014), 50–68.
[32] M. Itoh and Y. Shishido, Fisher information metric and Poisson kernels, Diff. Geom. Appl. 26 (2008), 347–356.
[33] S. Kashiwakura, Symmetry of Damek-Ricci spaces and curvature negativity (Japanese), Master thesis, University of Tsukuba, 2009.
[34] F. Ledrappier, —it Harmonic measures and Bowen-Margulis measures, Israel J. Math. 71 (1990), 275–287.
[35] T. Nakagawa, The rigidity theorem for real hyperbolic manifolds (Japanese), Master thesis, University of Tsukuba, 2003.
[36] A. Ohara, Geodesics for dual connections and means on symmetric cones, Integral Equations Operator Theory 50 (2004), 537–548.
[37] J. Oxtoby and S. Ulam, Measure preserving homeomorphisms and metrical transitivity, Ann. of Math. (2) 42 (1941), 874–920.

[38] G. Pistone and C. Sempi, An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one, Ann. Statist. 23 (1995), 1543–1561.

[39] T. Sakai Riemannian geometry, Transl. Math. Monogr. 149, Amer. Math. Soc., Providence, RI, 1996.

[40] J. Saunderson, Mostow’s rigidity theorem, Ph.D. thesis, University of Melbourne, 2008.

[41] R. Schoen and S.-T. Yau, Lectures on Differential Geometry, Conf. Proc. Lecture Notes Geom. Topology, I, Int. Press, Cambridge, MA, 1994.

[42] H. Shima, Geometry of Hessian Structures, World Scientific, 2007.

[43] Y. Shishido, —it Differential geometry of the Fisher information metric and the space of probability measures, Ph.D. thesis, University of Tsukuba, 2007.

[44] D. Sullivan. The Dirichlet problem at infinity for a negatively curved manifold, J. Differential Geom. 18 (1983), 723–732.

[45] Z. I. Szabó, The Lichnerowicz conjecture on harmonic manifolds, J. Differential Geom. 31 (1990), 1–28.

University of Tsukuba, Institute of Mathematics, 1-1-1 Ten-no-dai, Tsukuba City, 305-8571, Japan
Email address: itohm@math.tsukuba.ac.jp

Nippon Institute of Technology, Liberal Arts and Sciences, 4-1 Gakuen-dai, Miyashiro-machi, Saitama Prefecture, 345-8501 Japan
Email address: hiroyasu@nit.ac.jp