Entropy **2014**, *16*, 2472-2487; doi:10.3390/e16052472

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

F-Geometry and Amari's $\alpha-$Geometry on a Statistical Manifold

Harsha K. V. * and Subrahamanian Moosath K S *

Indian Institute of Space Science and Technology, Department of Space, Government of India, Valiamala P.O, Thiruvananthapuram-695547, Kerala, India

* Authors to whom correspondence should be addressed; E-Mails: harsha.11@iist.ac.in (K.V.H.); smoosath@iist.ac.in (K.S.S.M.); Tel.: +91-95-6736-0425 (K.V.H.); +91-94-9574-3148 (K.S.S.M.).

Received: 13 December 2013; in revised form: 21 April 2014 / Accepted: 25 April 2014 / Published: 6 May 2014

Abstract: In this paper, we introduce a geometry called F-geometry on a statistical manifold $\mathcal{S}$ using an embedding F of $\mathcal{S}$ into the space $\mathbb{R}_{\mathbb{X}}$ of random variables. Amari's $\alpha-$geometry is a special case of $F-$geometry. Then using the embedding F and a positive smooth function G, we introduce $(F, G)-$metric and $(F, G)-$connections that enable one to consider weighted Fisher information metric and weighted connections. The necessary and sufficient condition for two $(F, G)-$connections to be dual with respect to the $(F, G)-$metric is obtained. Then we show that Amari's $0-$connection is the only self dual $F-$connection with respect to the Fisher information metric. Invariance properties of the geometric structures are discussed, which proved that Amari's $\alpha-$connections are the only $F-$connections that are invariant under smooth one-to-one transformations of the random variables.

Keywords: embedding; Amari's $\alpha-$connections; $F-$metric; $F-$connections; $(F, G)-$metric; $(F, G)-$connections; invariance

1. Introduction

Geometric study of statistical estimation has opened up an interesting new area called the Information Geometry. Information geometry achieved a remarkable progress through the works of Amari [1,2], and his colleagues [3,4]. In the last few years, many authors have considerably contributed in this area [5–9]. Information geometry has a wide variety of applications in other areas of engineering and science, such as neural networks, machine learning, biology, mathematical finance, control system theory, quantum systems, statistical mechanics, *etc*.

A statistical manifold of probability distributions is equipped with a Riemannian metric and a pair of dual affine connections [2,4,9]. It was Rao [10] who introduced the idea of using Fisher information as a Riemannian metric in the manifold of probability distributions. Chentsov [11] introduced a family of affine connections on a statistical manifold defined on finite sets. Amari [2] introduced a family of affine connections called $\alpha-$connections using a one parameter family of functions, the $\alpha-$embeddings. These $\alpha-$connections are equivalent to those defined by Chentsov. The Fisher information metric and these affine connections are characterized by invariance with respect to the sufficient statistic [4,12] and play a vital role in the theory of statistical estimation. Zhang [13] generalized Amari's $\alpha-$representation and using this general representation together with a convex function he defined a family of divergence functions from the point of view of representational and referential duality. The Riemannian metric and dual connections are defined using these divergence functions.

In this paper, Amari's idea of using $\alpha-$embeddings to define geometric structures is extended to a general embedding. This paper is organized as follows. In Section 2, we define an affine connection called $F-$connection and a Riemannian metric called $F-$metric using a general embedding F of a statistical manifold $\mathcal{S}$ into the space of random variables. We show that $F-$metric is the Fisher information metric and Amari's $\alpha-$geometry is a special case of $F-$geometry. Further, we introduce $(F, G)-$metric and $(F, G)-$connections using the embedding F and a positive smooth function G.

In Section 3, a necessary and sufficient condition for two $(F, G)-$connections to be dual with respect to the $(F, G)-$metric is derived and we prove that Amari's $0-$connection is the only self dual $F-$connection with respect to the Fisher information metric. Then we prove that the set of all positive finite measures on X, for a finite X, has an $F-$affine manifold structure for any embedding F. In Section 4, invariance properties of the geometric structures are discussed. We prove that the Fisher information metric and Amari's $\alpha-$connections are invariant under both the transformation of the parameter and the transformation of the random variable. Further we show that Amari's $\alpha-$connections are the only $F-$connections that are invariant under both the transformation of the parameter and the transformation of the random variable.

Let $(X, \mathcal{B})$ be a measurable space, where X is a non-empty subset of $\mathbb{R}$ and $\mathcal{B}$ is the σ-field of subsets of X. Let $\mathbb{R}_X$ be the space of all real valued measurable functions defined on $(X, \mathcal{B})$. Consider an $n-$dimensional statistical manifold $\mathcal{S} = \{p(x;\xi) \ / \ \xi = [\xi^1, ..., \xi^n] \in \mathbb{E} \subseteq \mathbb{R}^n\}$, with coordinates $\xi = [\xi^1, ..., \xi^n]$, defined on X. $\mathcal{S}$ is a subset of $\mathcal{P}(X)$, the set of all probability measures on X given by

$$\mathcal{P}(X) := \{p : X \longrightarrow \mathbb{R} \ / \ p(x) > 0 \ (\forall\, x \in X); \int_X p(x)dx = 1\}. \tag{1}$$

The tangent space to $\mathcal{S}$ at a point p_ξ is given by

$$T_\xi(\mathcal{S}) = \{\sum_{i=1}^n \alpha^i \partial_i \ / \ \alpha^i \in \mathbb{R}\} \quad \text{where} \ \ \partial_i = \frac{\partial}{\partial \xi^i}. \tag{2}$$

Define $\ell(x;\xi) = \log p(x;\xi)$ and consider the partial derivatives $\{\frac{\partial \ell}{\partial \xi^i} = \partial_i \ell \ ; i = 1, ..., n\}$ which are called scores. For the statistical manifold $\mathcal{S}$, $\partial_i \ell$'s are linearly independent functions in x for a fixed ξ. Let $T^1_\xi(\mathcal{S})$ be the n-dimensional vector space spanned by n functions $\{\partial_i \ell \ ; \ i = 1, ..., n\}$ in x. So

$$T^1_\xi(\mathcal{S}) = \{\sum_{i=1}^n A^i \partial_i \ell \ / \ A^i \in \mathbb{R}\}. \tag{3}$$

Then there is a natural isomorphism between these two vector spaces $T_\xi(\mathcal{S})$ and $T^1_\xi(\mathcal{S})$ given by

$$\partial_i \in T_\xi(\mathcal{S}) \longleftrightarrow \partial_i \ell(x;\xi) \in T^1_\xi(\mathcal{S}). \tag{4}$$

Obviously, a tangent vector $A = \sum_{i=1}^n A^i \partial_i \in T_\xi(\mathcal{S})$ corresponds to a random variable $A(x) = \sum_{i=1}^n A^i \partial_i \ell(x;\xi) \in T^1_\xi(\mathcal{S})$ having the same components A^i. Note that $T_\xi(\mathcal{S})$ is the differentiation operator representation of the tangent space, while $T^1_\xi(\mathcal{S})$ is the random variable representation of the same tangent space. The space $T^1_\xi(\mathcal{S})$ is called the 1-representation of the tangent space.

Let A and B be two tangent vectors in $T_\xi(\mathcal{S})$ and $A(x)$ and $B(x)$ be the $1-$representations of A and B respectively. We can define an inner product on each tangent space $T_\xi(\mathcal{S})$ by

$$g_\xi(A,B) =< A, B >_\xi = E_\xi[A(x)B(x)] = \int A(x)B(x)p(x;\xi)dx. \tag{5}$$

Especially the inner product of the basis vectors ∂_i and ∂_j is

$$g_{ij}(\xi) = < \partial_i, \partial_j >_\xi = E_\xi[\partial_i \ell \; \partial_j \ell] = \int \partial_i \ell(x;\xi)\partial_j \ell(x;\xi)p(x;\xi)dx. \tag{6}$$

Note that $g =<,>$ defines a Riemannian metric on $\mathcal{S}$ called the **Fisher information metric**.

On the Riemannian manifold $(\mathcal{S}, g =<,>)$, define n^3 functions Γ_{ijk} by

$$\Gamma_{ijk}(\xi) = E_\xi[(\partial_i\partial_j \ell(x;\xi))(\partial_k \ell(x;\xi))]. \tag{7}$$

These functions Γ_{ijk} uniquely determine an affine connection ∇ on $\mathcal{S}$ by

$$\Gamma_{ijk}(\xi) =< \nabla_{\partial_i}\partial_j, \partial_k >_\xi \; . \tag{8}$$

∇ is called the $1-$connection or the exponential connection.

Amari [2] defined a one parameter family of functions called the $\alpha-$embeddings given by

$$L_\alpha(p) = \begin{cases} \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}} & \alpha \neq 1 \\ \log p & \alpha = 1 \end{cases} \tag{9}$$

Using these, we can define n^3 functions Γ^α_{ijk} by

$$\Gamma^\alpha_{ijk} = \int \partial_i\partial_j L_\alpha(p(x;\xi))\partial_k L_{-\alpha}(p(x;\xi))dx \tag{10}$$

These Γ^α_{ijk} uniquely determine affine connections ∇^α on the statistical manifold $\mathcal{S}$ by

$$\Gamma^\alpha_{ijk} = < \nabla^\alpha_{\partial_i}\partial_j, \partial_k > \tag{11}$$

which are called $\alpha-$**connections**.

2. $F-$Geometry of a Statistical Manifold

On a statistical manifold $\mathcal{S}$, the Fisher information metric and exponential connection are defined using the log embedding. In a similar way, $\alpha-$connections are defined using a one parameter family of functions, the $\alpha-$embeddings. In general, we can give other geometric structures on $\mathcal{S}$ using different embeddings of the manifold $\mathcal{S}$ into the space of random variables $\mathbb{R}_X$.
Let $F : (0, \infty) \longrightarrow \mathbb{R}$ be an injective function that is at least twice differentiable. Thus we have $F'(u) \neq 0, \ \forall \ u \in (0, \infty)$. F is an embedding of $\mathcal{S}$ into $\mathbb{R}_X$ that takes each $p(x;\xi) \longmapsto F(p(x;\xi))$. Denote $F(p(x;\xi))$ by $F(x;\xi)$ and $\partial_i F$ can be written as

$$\partial_i F(x;\xi) = p(x;\xi)F'(p(x;\xi))\partial_i \ell(p(x;\xi)). \tag{12}$$

It is clear that $\partial_i F(x;\xi); \ \ i = 1, ..., n$ are linearly independent functions in x for fixed ξ since $\partial_i \ell(p(x;\xi)); \ \ i = 1, .., n$ are linearly independent. Let $T_{F(p_\xi)}F(\mathcal{S})$ be the n-dimensional vector space spanned by n functions $\partial_i F; \ i = 1, ..., n$ in x for fixed ξ. So

$$T_{F(p_\xi)}F(\mathcal{S}) = \{\sum_{i=1}^{n} A^i \partial_i F \ / \ A^i \in \mathbb{R}\} \tag{13}$$

Let the tangent space $T_{F(p_\xi)}(F(\mathcal{S}))$ to $F(\mathcal{S})$ at the point $F(p_\xi)$ be denoted by $T_\xi^F(\mathcal{S})$. There is a natural isomorphism between the two vector spaces $T_\xi(\mathcal{S})$ and $T_\xi^F(\mathcal{S})$ given by

$$\partial_i \in T_\xi(\mathcal{S}) \longleftrightarrow \partial_i F(x;\xi) \in T_\xi^F(\mathcal{S}). \tag{14}$$

$T_\xi^F(\mathcal{S})$ is called the $F-$representation of the tangent space $T_\xi(\mathcal{S})$.
For any $A = \sum_{i=1}^n A^i \partial_i \in T_\xi(\mathcal{S})$, the corresponding $A(x) = \sum_{i=1}^n A^i \partial_i F \in T_\xi^F(\mathcal{S})$ is called the $F-$representation of the tangent vector A and is denoted by $A^F(x)$. Note that $T_\xi^F(\mathcal{S}) \subseteq T_{F(p_\xi)}(\mathbb{R}_X)$. Since $\mathbb{R}_X$ is a vector space, its tangent space $T_{F(p_\xi)}(\mathbb{R}_X)$ can be identified with $\mathbb{R}_X$. So $T_\xi^F(\mathcal{S}) \subseteq \mathbb{R}_X$.

Definition 2.1. *$F-$expectation of a random variable f with respect to the distribution $p(x;\xi)$ is defined as*

$$E_\xi^F(f) = \int f(x) \frac{1}{p(F'(p))^2} dx. \tag{15}$$

We can use this $F-$expectation to define an inner product in $\mathbb{R}_X$ by

$$< f, g >_\xi^F = \ E_\xi^F[f(x)g(x)], \tag{16}$$

which induces an inner product on $T_\xi(\mathcal{S})$ by

$$< A, B >_\xi^F = \ E_\xi^F[A^F(x)B^F(x)] \ ; \ A, B \in T_\xi(\mathcal{S}). \tag{17}$$

Proposition 2.2. *The induced metric $<, >^F$ on $\mathcal{S}$ is the Fisher information metric $g =<, >$ on $\mathcal{S}$.*

Proof. For any basis vectors $\partial_i, \partial_j \in T_\xi(\mathcal{S})$

$$\begin{aligned}
< \partial_i, \partial_j >_\xi^F &= E_\xi^F[\partial_i F\ \partial_j F] \\
&= \int \partial_i F\ \partial_j F \frac{1}{p(F'(p))^2} dx \\
&= \int (p\ F'(p)\ \partial_i \ell)\ (p\ F'(p)\ \partial_j \ell)\ \frac{1}{p(F'(p))^2} dx \\
&= \int \partial_i \ell\ \partial_j \ell\ p(x;\xi)\ dx \\
&= E_\xi[\partial_i \ell\ \partial_j \ell] \\
&= g_{ij}(\xi) \\
&= < \partial_i, \partial_j >_\xi .
\end{aligned} \tag{18}$$

So the metric $<,>^F$ on $\mathcal{S}$ induced by the embedding F of $\mathcal{S}$ into $\mathbb{R}_X$ is the Fisher information metric $g =<,>$ on $\mathcal{S}$. □

We can induce a connection on $\mathcal{S}$ using the embedding F.
Let $\pi^F_{|p_\xi} : \mathbb{R}_X \longrightarrow T^F_\xi(\mathcal{S})$ be the projection map.

Definition 2.3. *The connection induced by the embedding* F *on* $\mathcal{S}$*, the* $F-$*connection, is defined as*

$$\begin{aligned}
\nabla^F_{\partial_i} \partial_j &= \pi^F_{|p_\xi}(\partial_i \partial_j F) \\
&= \sum_n \sum_m g^{mn} < \partial_i \partial_j F, \partial_m F >_\xi^F \partial_n .
\end{aligned} \tag{19}$$

where $[g^{mn}(\xi)]$ *is the inverse of the Fisher information matrix* $G(\xi) = [g_{mn}(\xi)]$*. Note that the* $F-$*connections are symmetric.*

Lemma 2.4. *The* $F-$*connection and its components can be written in terms of scores as*

$$\nabla^F_{\partial_i} \partial_j = \sum_n \sum_m g^{mn} E_\xi \left[(\partial_i \partial_j \ell + (1 + \frac{pF''(p)}{F'(p)}) \partial_i \ell\ \partial_j \ell)(\partial_m \ell) \right] \partial_n \tag{20}$$

and

$$\Gamma^F_{ijk}(\xi) = E_\xi \left[(\partial_i \partial_j \ell + (1 + \frac{pF''(p)}{F'(p)}) \partial_i \ell\ \partial_j \ell)(\partial_k \ell) \right] \tag{21}$$

Proof. From Equation (12), we have

$$\partial_i \partial_j F = pF'(p) \partial_i \partial_j \ell + [pF'(p) + p^2 F''(p)]\ \partial_i \ell\ \partial_j \ell. \tag{22}$$

Therefore

$$\begin{aligned}
< \partial_i \partial_j F, \partial_m F >_\xi^F &= \int \partial_i \partial_j F\ \partial_m F \frac{1}{p(F'(p))^2} dx \\
&= \int \left(pF'(p) \partial_i \partial_j \ell + [pF'(p) + p^2 F''(p)]\ \partial_i \ell\ \partial_j \ell \right)\ \frac{\partial_m \ell}{F'(p)} dx \\
&= \int \left(\partial_i \partial_j \ell\ \partial_m \ell + (1 + \frac{pF''(p)}{F'(p)}) \partial_i \ell\ \partial_j \ell\ \partial_m \ell \right) p dx \\
&= E_\xi \left[(\partial_i \partial_j \ell + (1 + \frac{pF''(p)}{F'(p)}) \partial_i \ell\ \partial_j \ell)(\partial_m \ell) \right] .
\end{aligned} \tag{23}$$

Hence we can write

$$\begin{aligned}
\nabla^F_{\partial_i}\partial_j &= \pi^F_{|p_\xi}(\partial_i\partial_j F) \\
&= \sum_n \sum_m g^{mn} E_\xi \left[(\partial_i\partial_j \ell + (1 + \frac{pF''(p)}{F'(p)})\partial_i \ell \; \partial_j \ell)(\partial_m \ell) \right] \partial_n .
\end{aligned} \tag{24}$$

Then we have the Christoffel symbols of the $F-$connection

$$\Gamma^n_{ij} = \sum_m g^{mn} E_\xi \left[(\partial_i\partial_j \ell + (1 + \frac{pF''(p)}{F'(p)})\partial_i \ell \; \partial_j \ell)(\partial_m \ell) \right] \tag{25}$$

and components of the $F-$connection are given by

$$\Gamma^F_{ijk}(\xi) = < \nabla^F_{\partial_i}\partial_j, \partial_k >_\xi = E_\xi \left[(\partial_i\partial_j \ell + (1 + \frac{pF''(p)}{F'(p)})\partial_i \ell \; \partial_j \ell)(\partial_k \ell) \right]. \tag{26}$$

□

Theorem 2.5. *Amari's $\alpha-$geometry is a special case of the $F-$geometry.*

Proof. Let $F(p) = L_\alpha(p)$, $L_\alpha(p)$ is the $\alpha-$embedding of Amari.
The components Γ^α_{ijk} of the $\alpha-$connection are given by

$$\begin{aligned}
\Gamma^\alpha_{ijk}(\xi) &= < \nabla^\alpha_{\partial_i}\partial_j, \partial_k >_\xi \\
&= E_\xi \left[(\partial_i\partial_j \ell + \frac{1-\alpha}{2}\partial_i \ell \; \partial_j \ell)(\partial_k \ell) \right].
\end{aligned} \tag{27}$$

From Equation (26), when $F(p) = L_\alpha(p)$
we have

$$F'(p) = L'_\alpha(p) = p^{-\left(\frac{1+\alpha}{2}\right)} \tag{28}$$

$$F''(p) = L''_\alpha(p) = -\frac{1+\alpha}{2} p^{-\left(\frac{3+\alpha}{2}\right)}. \tag{29}$$

Then we get

$$1 + \frac{pF''(p)}{F'(p)} = 1 + \frac{pL''_\alpha(p)}{L'_\alpha(p)} = \frac{1-\alpha}{2} \tag{30}$$

Hence

$$\begin{aligned}
\Gamma^F_{ijk}(\xi) = < \nabla^F_{\partial_i}\partial_j, \partial_k >_\xi &= E_\xi \left[(\partial_i\partial_j \ell + (1 + \frac{pF''(p)}{F'(p)})\partial_i \ell \; \partial_j \ell)(\partial_k \ell) \right] \\
&= E_\xi \left[(\partial_i\partial_j \ell + \frac{1-\alpha}{2}\partial_i \ell \; \partial_j \ell)(\partial_k \ell) \right] \\
&= \Gamma^\alpha_{ijk}(\xi)
\end{aligned} \tag{31}$$

which are the components of the $\alpha-$connection. Hence $F-$connection reduces to $\alpha-$connection. Thus we obtain that $\alpha-$geometry is a special case of $F-$geometry. □

Remark 2.6. *Burbea [14] introduced the concept of weighted Fisher information metric using a positive continuous function. We use this idea to define weighted* $F-$*metric and weighted* $F-$*connections. Let* $G : (0, \infty) \longrightarrow \mathbb{R}$ *be a positive smooth function and* F *be an embedding, define* $(F, G)-$*expectation of a random variable with respect to the distribution* p_ξ *as*

$$E_\xi^{F,G}(f) = \int f(x) \frac{G(p)}{p(F'(p))^2} dx. \tag{32}$$

Define $(F, G)-$*metric* $<, >_\xi^{F,G}$ *in* $T_{p_\xi}(\mathcal{S})$ *by*

$$\begin{aligned} < \partial_i, \partial_j >_\xi^{F,G} &= E_\xi^{F,G}[\partial_i F \; \partial_j F] \\ &= \int \partial_i F \; \partial_j F \frac{G(p)}{p(F'(p))^2} dx \\ &= \int \partial_i \ell \; \partial_j \ell \; G(p) \; p \; dx \\ &= E_\xi[G(p) \; \partial_i \ell \; \partial_j \ell]. \end{aligned} \tag{33}$$

Define $(F, G)-$*connection as*

$$\begin{aligned} \Gamma_{ijk}^{F,G} &= < \nabla_{\partial_i}^{F,G} \partial_j, \partial_k >_\xi \\ &= E_\xi \left[\left((\partial_i \partial_j \ell + (1 + \frac{pF''(p)}{F'(p)}) \partial_i \ell \; \partial_j \ell)(\partial_k \ell) \right) (G(p)) \right]. \end{aligned} \tag{34}$$

When $G(p) = 1$, $(F, G)-$*connection reduces to the* $F-$*connection and the metric* $<, >^{F,G}$ *reduces to the Fisher information metric. This is a more general way of defining Riemannian metrics and affine connections on a statistical manifold.*

3. Dual Affine Connections

Definition 3.1. *Let* M *be a Riemannian manifold with a Riemannian metric* g*. Two affine connections,* ∇ *and* ∇^* *on the tangent bundle are said to be* ***dual connections*** *with respect to the metric* g *if*

$$Zg(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z^* Y) \tag{35}$$

holds for any vector fields X, Y, Z *on* M.

Theorem 3.2. *Let* F, H *be two embeddings of statistical manifold* $\mathcal{S}$ *into the space* $\mathbb{R}_\mathbb{X}$ *of random variables. Let* G *be a positive smooth function on* $(0, \infty)$*. Then the* $(F, G)-$*connection* $\nabla^{F,G}$ *and the* $(H, G)-$*connection* $\nabla^{H,G}$ *are dual connections with respect to the* $(F, G)-$*metric iff the functions* F *and* H *satisfy*

$$H'(p) = \frac{G(p)}{pF'(p)}. \tag{36}$$

We call such an embedding H *as a* $G-$*dual embedding of* F.
The components of the dual connection $\nabla^{H,G}$ *can be written as*

$$\Gamma_{ijk}^{H,G} = \int \left(\partial_i \partial_j \ell + (\frac{pG'(p)}{G(p)} - \frac{pF''(p)}{F'(p)}) \partial_i \ell \; \partial_j \ell \right) \partial_k \ell \; G(p) p \; dx. \tag{37}$$

Proof. $\nabla^{F,G}$ and $\nabla^{H,G}$ are dual connections with respect to the $G-$metric means,

$$\partial_k < \partial_i, \partial_j >^{F,G} = < \nabla^{F,G}_{\partial_k}\partial_i, \partial_j >^{F,G} + < \partial_i, \nabla^{H,G}_{\partial_k}\partial_j >^{F,G} . \tag{38}$$

for any basis vectors $\partial_i, \partial_j, \partial_k \in T_\xi(\mathcal{S})$.

$$\begin{aligned} \partial_k < \partial_i, \partial_j >^{F,G} &= \int \partial_k\partial_j\ell\, \partial_i\ell\, pG(p)dx + \int \partial_k\partial_i\ell\, \partial_j\ell\, pG(p)dx \\ &+ \int (1 + \frac{pG'(p)}{G(p)})\partial_i\ell\, \partial_j\ell\, \partial_k\ell\, pG(p)dx. \end{aligned} \tag{39}$$

$$\begin{aligned} < \nabla^{F,G}_{\partial_k}\partial_i, \partial_j >^{F,G} + < \partial_i, \nabla^{H,G}_{\partial_k}\partial_j >^{F,G} &= \int \partial_k\partial_i\ell\, \partial_j\ell\, pG(p)dx \\ &+ \int 1 + \frac{pF''(p)}{F'(p)}\partial_i\ell\, \partial_j\ell\, \partial_k\ell\, pG(p)dx \\ &+ \int 1 + \frac{pH''(p)}{H'(p)}\partial_i\ell\, \partial_j\ell\, \partial_k\ell\, pG(p)dx \\ &+ \int \partial_k\partial_j\ell\, \partial_i\ell\, pG(p)dx \end{aligned} \tag{40}$$

Then the condition (38) holds iff

$$\begin{aligned} &\int [2 + \frac{pF''(p)}{F'(p)} + \frac{pH''(p)}{H'(p)}]\partial_i\ell\, \partial_j\ell\, \partial_k\ell\, pG(p)dx = \\ &\qquad \int [1 + \frac{pG'(p)}{G(p)}]\partial_i\ell\, \partial_j\ell\, \partial_k\ell\, pG(p)dx \end{aligned} \tag{41}$$

$$\iff [2 + \frac{pF''(p)}{F'(p)} + \frac{pH''(p)}{H'(p)}] = 1 + \frac{pG'(p)}{G(p)}. \tag{42}$$

$$\iff 1 + \frac{pH''(p)}{H'(p)} = \frac{pG'(p)}{G(p)} - \frac{pF''(p)}{F'(p)} \tag{43}$$

$$\iff \frac{H''(p)}{H'(p)} = \frac{G'(p)}{G(p)} - \frac{F''(p)}{F'(p)} - \frac{1}{p} \iff H'(p) = \frac{G(p)}{pF'(p)}. \tag{44}$$

Hence $\nabla^{F,G}$ and $\nabla^{H,G}$ are dual connections with respect to the $(F, G)-$metric iff Equation (36) holds. From Equation (43), we can rewrite the components of dual connection $\nabla^{H,G}$ as

$$\Gamma^{H,G}_{ijk} = \int \left(\partial_i\partial_j\ell + (\frac{pG'(p)}{G(p)} - \frac{pF''(p)}{F'(p)})\partial_i\ell\, \partial_j\ell\right) \partial_k\ell\, G(p)p\, dx. \tag{45}$$

□

Corollary 3.3. *Amari's $0-$connection is the only self dual $F-$connection with respect to the Fisher information metric.*

Proof. From Theorem 3.2, for $G(p) = 1$ the $F-$connection ∇^F and the $H-$connection ∇^H are dual connections with respect to the Fisher information metric iff the functions F and H satisfy

$$H'(p) = \frac{1}{pF'(p)} \tag{46}$$

Thus the $F-$connection ∇^F is self dual iff the embedding F satisfies the condition

$$F'(p) = \frac{1}{pF'(p)} \iff F'(p) = p^{-(\frac{1}{2})} \iff F(p) = 2p^{\frac{1}{2}} = L_0(p). \tag{47}$$

That is, Amari's $0-$connection is the only self dual $F-$connection with respect to the Fisher information metric. □

So far, we have considered the statistical manifold $\mathcal{S}$ as a subset of $\mathcal{P}(X)$, the set of all probability measures on X. Now we relax the condition $\int p(x)dx = 1$, and consider $\mathcal{S}$ as a subset of $\tilde{\mathcal{P}}(X)$, which is defined by

$$\tilde{\mathcal{P}}(X) := \{p : X \longrightarrow \mathbb{R} \ / \ p(x) > 0 \ (\forall \ x \in X); \int\limits_X p(x)dx < \infty\}. \tag{48}$$

Definition 3.4. *Let M be a Riemannian manifold with a Riemannian metric g. Let ∇ be an affine connection on M. If there exists a coordinate system $[\theta^i]$ of M such that $\nabla_{\partial_i}\partial_j = 0$ then we say that ∇ is flat, or alternatively M is flat with respect to ∇, and we call such a coordinate system $[\theta^i]$ an affine coordinate system for ∇.*

Definition 3.5. *Let $\mathcal{S} = \{p(x;\xi) \ / \ \xi = [\xi^1, ..., \xi^n] \in \mathbb{E} \subseteq \mathbb{R}^n\}$ be an $n-$dimensional statistical manifold. If for some coordinate system $[\theta^i]$; $i = 1, ..., n$*

$$\partial_i\partial_j F(p(x;\theta)) = 0 \tag{49}$$

then we can see from Equation (19) that $[\theta^i]$ is an $F-$affine coordinate system and that $\mathcal{S} = \{p_\theta\}$ is $F-$flat. We call such $\mathcal{S}$ as an $F-$affine manifold.
The condition (49) is equivalent to the existence of the functions $C, F_1, .., F_n$ on X such that

$$F(p(x;\theta)) = C(x) + \sum_{i=1}^{n} \theta^i F_i(x) \tag{50}$$

Theorem 3.6. *For any embedding F, $\tilde{\mathcal{P}}(X)$ is an $F-$affine manifold for finite X.*

Proof. Let $X = \{x_1, ..., x_n\}$ be a finite set constituted by n elements. Let $F_i : X \longrightarrow \mathbb{R}$ be the functions defined by $F_i(x_j) = \delta_{ij}$ for $i, j = 1, .., n$. Let us define n coordinates $[\theta^i]$ by

$$\theta^i = F(p(x_i)) \tag{51}$$

Then we get $F(p(x)) = \sum_{i=1}^n \theta^i F_i(x)$. Therefore $\tilde{\mathcal{P}}(X)$ is an $F-$affine manifold for any embedding $F(p)$. □

Remark 3.7. *Zhang [13] introduced ρ-representation, which is a generalization of α-representation of Amari. Zhang's geometry is defined using this ρ-representation together with a convex function. Zhang also defined the ρ-affine family of density functions and discussed its dually flat structure. The $F-$geometry defined using a general F-representation is different from the Zhang's geometry. The metric defined in the F-embedding approach is the Fisher information metric and the Riemannian metric defined using the ρ-representation is different from the Fisher information metric. The F-connections defined are not in general dually flat and are different from the dual connections defined by Zhang.*

Remark 3.8. *On a statistical manifold $\mathcal{S}$, we introduced a dualistic structure (g, ∇^F, ∇^H), where g is the Fisher information metric and ∇^F, ∇^H are the dual connections with respect to the Fisher information metric. Since F-connections are symmetric, the manifold $\mathcal{S}$ is flat with respect to ∇^F iff $\mathcal{S}$ is flat with respect to ∇^H. Thus if $\mathcal{S}$ is flat with respect to ∇^F, then $(\mathcal{S}, g, \nabla^F, \nabla^H)$ is a dually flat space. The dually flat spaces are important in statistical estimation [4].*

4. Invariance of the Geometric Structures

For the statistical manifold $\mathcal{S} = \{p(x;\xi) \mid \xi \in \mathbb{E} \subseteq \mathbb{R}^n\}$, the parameters are merely labels attached to each point $p \in \mathcal{S}$, hence the intrinsic geometric properties should be independent of these labels. Consequently, it is natural to consider the invariance properties of the geometric structures under suitable transformations of the variables in a statistical manifold. Here we can consider two kinds of invariance of the geometric structures; covariance under re-parametrization of the parameter of the manifold and invariance under the transformations of the random variable [15]. Now let us investigate the invariance properties of the F-geometric structures defined in Section 2.

4.1. Covariance under Re-Parametrization

Let $[\theta^i]$ and $[\eta_j]$ be two coordinate systems on S, which are related by an invertible transformation $\eta = \eta(\theta)$. Let us denote $\partial_i = \frac{\partial}{\partial \theta^i}$ and $\partial^j = \frac{\partial}{\partial \eta_j}$. Let the coordinate expressions of the metric g be given by $g_{ij} =< \partial_i, \partial_j >$ and $\tilde{g}_{ij} =< \partial^i, \partial^j >$. Let the components of the connection ∇ with respect to the coordinates $[\theta^i]$ and $[\eta_j]$ be given by Γ_{ijk}, $\tilde{\Gamma}_{ijk}$ respectively.

Then the covariance of the metric g and the connection ∇ under the re-parametrization means,

$$\tilde{g}_{ij} = \sum_m \sum_n \frac{\partial \theta^m}{\partial \eta_i} \frac{\partial \theta^n}{\partial \eta_j} g_{mn} \tag{52}$$

$$\tilde{\Gamma}_{ijk} = \sum_{m,n,h} \frac{\partial \theta^m}{\partial \eta_i} \frac{\partial \theta^n}{\partial \eta_j} \frac{\partial \theta^h}{\partial \eta_k} \Gamma_{mnh} + \sum_{m,h} \frac{\partial \theta^h}{\partial \eta_k} \frac{\partial^2 \theta^m}{\partial \eta_i \partial \eta_j} g_{mh} \tag{53}$$

Lemma 4.1. *The Fisher information metric g is covariant under re-parametrization.*

Proof. The components of the Fisher information metric with respect to the coordinate system $[\theta^i]$ are given by

$$g_{ij}(\theta) = < \partial_i, \partial_j >_\theta = \int \partial_i p(x;\theta) \partial_j p(x;\theta) \frac{1}{p(x;\theta)} dx. \tag{54}$$

Let $\tilde{p}(x;\eta) = p(x;\theta(\eta))$. Then the components of the Fisher information metric with respect to the coordinate system $[\eta_j]$ are given by

$$\tilde{g}_{ij}(\eta) = < \partial^i, \partial^j >_\eta = \int \partial^i \tilde{p}(x;\eta) \partial^j \tilde{p}(x;\eta) \frac{1}{\tilde{p}(x;\eta)} dx. \tag{55}$$

Since

$$\partial^i \tilde{p}(x;\eta) = \sum_m \frac{\partial \theta^m}{\partial \eta_i} \frac{\partial p(x;\theta(\eta))}{\partial \theta^m} \tag{56}$$

we can write

$$\begin{aligned}
\tilde{g}_{ij}(\eta) &= \int \partial^i \tilde{p}(x;\eta)\partial^j \tilde{p}(x;\eta)\frac{1}{\tilde{p}(x;\eta)}dx \\
&= \int \sum_m \frac{\partial\theta^m}{\partial\eta_i}\frac{\partial p(x;\theta)}{\partial\theta^m}\sum_n \frac{\partial\theta^n}{\partial\eta_j}\frac{\partial p(x;\theta)}{\partial\theta^n}\frac{1}{p(x;\theta)}dx \\
&= \sum_m\sum_n \frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j}\int \partial_m p(x;\theta)\partial_n p(x;\theta)\frac{1}{p(x;\theta)}dx. \\
&= \left[\sum_m\sum_n \frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j} g_{mn}(\theta)\right]_{\theta=\theta(\eta)}
\end{aligned} \tag{57}$$

□

Lemma 4.2. *The $F-$connection ∇^F is covariant under re-parametrization.*

Proof. Let the components of ∇^F with respect to the coordinates $[\theta^i]$ and $[\eta_j]$ be given by Γ_{ijk}, $\tilde{\Gamma}_{ijk}$ respectively.
Let $\tilde{p}(x;\eta) = p(x;\theta(\eta))$. Let us denote $\log p(x;\theta)$ by $\ell(x;\theta)$ and $\log\tilde{p}(x;\eta)$ by $\tilde{\ell}(x;\eta)$.
The components of the $F-$connection ∇^F with respect to the coordinate system $[\theta^i]$ are given by

$$\Gamma_{ijk} = \int \left(\partial_i\partial_j\ell(x;\theta) + (1 + \frac{pF''(p)}{F'(p)})\partial_i\ell(x;\theta)\ \partial_j\ell(x;\theta)\right)\partial_k\ell(x;\theta)p(x;\theta)dx \tag{58}$$

The components of ∇^F with respect to the coordinate system $[\eta_j]$ are given by

$$\tilde{\Gamma}_{ijk} = \int \left(\partial^i\partial^j\tilde{\ell}(x;\eta) + (1 + \frac{\tilde{p}F''(\tilde{p})}{F'(\tilde{p})})\partial^i\tilde{\ell}(x;\eta)\ \partial^j\tilde{\ell}(x;\eta)\right)\partial^k\tilde{\ell}(x;\eta)\tilde{p}(x;\eta)dx \tag{59}$$

We can write

$$\partial^i\tilde{\ell}(x;\eta) = \sum_m \frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^m} \tag{60}$$

Then

$$\partial^i\partial^j\tilde{\ell}(x;\eta) = \sum_{m,n}\frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j}\frac{\partial^2\ell(x;\theta(\eta))}{\partial\theta^m\partial\theta^n} + \sum_m \frac{\partial^2\theta^m}{\partial\eta_i\partial\eta_j}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^m} \tag{61}$$

$$\partial^i\tilde{\ell}(x;\eta)\ \partial^j\tilde{\ell}(x;\eta) = \sum_{m,n}\frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^m}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^n} \tag{62}$$

$$\partial^k\tilde{\ell}(x;\eta) = \sum_h \frac{\partial\theta^h}{\partial\eta_k}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^h} \tag{63}$$

Hence we get

$$\begin{aligned}
\tilde{\Gamma}_{ijk} &= \int\sum_{m,n,h}\frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j}\frac{\partial\theta^h}{\partial\eta_k}\frac{\partial^2\ell(x;\theta(\eta))}{\partial\theta^m\partial\theta^n}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^h}p(x;\theta(\eta))dx + \\
&\quad \int\sum_{m,h}\frac{\partial^2\theta^m}{\partial\eta_i\partial\eta_j}\frac{\partial\theta^h}{\partial\eta_k}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^m}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^h}p(x;\theta(\eta))dx + \\
&\quad \int(1 + \frac{pF''(p)}{F'(p)})\sum_{m,n,h}\frac{\partial\theta^m}{\partial\eta_i}\frac{\partial\theta^n}{\partial\eta_j}\frac{\partial\theta^h}{\partial\eta_k}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^m}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^n}\frac{\partial\ell(x;\theta(\eta))}{\partial\theta^h}p(x;\theta(\eta))dx
\end{aligned} \tag{64}$$

$$
\begin{aligned}
&= \sum_{m,n,h} \frac{\partial \theta^m}{\partial \eta_i}\frac{\partial \theta^n}{\partial \eta_j}\frac{\partial \theta^h}{\partial \eta_k}\int \frac{\partial^2 \ell(x;\theta(\eta))}{\partial \theta^m \partial \theta^n}\frac{\partial \ell(x;\theta(\eta))}{\partial \theta^h} p(x;\theta(\eta))dx + \\
&\quad \sum_{m,h} \frac{\partial^2 \theta^m}{\partial \eta_i \partial \eta_j}\frac{\partial \theta^h}{\partial \eta_k}\int \frac{\partial \ell(x;\theta(\eta))}{\partial \theta^m}\frac{\partial \ell(x;\theta(\eta))}{\partial \theta^h} p(x;\theta(\eta))dx + \\
&\quad \sum_{m,n,h} \frac{\partial \theta^m}{\partial \eta_i}\frac{\partial \theta^n}{\partial \eta_j}\frac{\partial \theta^h}{\partial \eta_k}\int (1+\frac{pF''(p)}{F'(p)})\frac{\partial \ell(x;\theta(\eta))}{\partial \theta^m}\frac{\partial \ell(x;\theta(\eta))}{\partial \theta^n}\frac{\partial \ell(x;\theta(\eta))}{\partial \theta^h} p(x;\theta(\eta))dx \\
&= \sum_{m,n,h} \frac{\partial \theta^m}{\partial \eta_i}\frac{\partial \theta^n}{\partial \eta_j}\frac{\partial \theta^h}{\partial \eta_k}\Gamma_{mnh} + \sum_{m,h}\frac{\partial \theta^h}{\partial \eta_k}\frac{\partial^2 \theta^m}{\partial \eta_i \partial \eta_j} g_{mh}
\end{aligned}
$$

□

Hence we showed that $F-$connections are covariant under re-parametrization of the parameter. The covariance under re-parametrization actually means that the metric and connections are coordinate independent. Hence we obtained that the $F-$geometry is coordinate independent.

4.2. Invariance Under the Transformation of the Random Variable

Amari and Nagaoka [4] defined the invariance of Riemannian metric and connections on a statistical manifold under a transformation of the random variable as follows,

Definition 4.3. *Let $\mathcal{S} = \{p(x;\xi) \mid \xi \in \mathbb{E} \subseteq \mathbb{R}^n\}$ be a statistical manifold defined on a sample space X. Let x, y be random variables defined on sample spaces X, Y respectively and ϕ be a transformation of x to y. Assume that this transformation induces a model $\mathcal{S}' = \{q(y;\xi) \mid \xi \in \mathbb{E} \subseteq \mathbb{R}^n\}$ on Y. Let $\lambda : \mathcal{S} \longrightarrow \mathcal{S}'$ be a diffeomorphism defined as*

$$\lambda(p_\xi) = q_\xi \tag{65}$$

Let $g = <>, g' = <>'$ be two Riemannian metrics defined on $\mathcal{S}$ and $\mathcal{S}'$ respectively. Let ∇, ∇' be two affine connections on $\mathcal{S}$ and $\mathcal{S}'$ respectively. Then the invariance properties are given by

$$
\begin{aligned}
< X, Y >_p &= < \lambda_*(X), \lambda_*(Y) >'_{\lambda(p)} \;\; \forall\, X, Y \in T_p(\mathcal{S}) && (66) \\
\lambda_*(\nabla_X Y) &= \nabla'_{\lambda_*(X)}\lambda_*(Y) && (67)
\end{aligned}
$$

where λ_* is the push forward map associated with the map λ, which is defined by

$$\lambda_*(X)_{\lambda(p)} = (d\lambda)_p(X) \tag{68}$$

Now we discuss the invariance properties of the $F-$geometry under suitable transformations of the random variable. Let us restrict ourselves to the case of smooth one-to-one transformations of the random variable that are in fact statistically interesting. Amari and Nagaoka [4] mentioned a transformation, the sufficient statistic of the parameter of the statistical model, which is widely used in statistical estimation. In fact the one-to-one transformations of the random variable are trivial examples of sufficient statistic.

Consider a statistical manifold $\mathcal{S} = \{p(x;\xi) \mid \xi \in \mathbb{E} \subseteq \mathbb{R}^n\}$ defined on a sample space X. Let ϕ be a smooth one-to-one transformation of the random variable x to y. Then the density function $q(y;\xi)$ of the induced model $\mathcal{S}'$ takes the form

$$q(y:\xi) = p(w(y);\xi)w'(y) \tag{69}$$

where w is a function such that $x = w(y)$ and $\phi'(x) = \frac{1}{w'(\phi(x))}$.
Let us denote $\log q(y;\xi)$ by $\ell(q_y)$ and $\log p(x;\xi)$ by $\ell(p_x)$.

Lemma 4.4. *The Fisher information metric and Amari's α-connections are invariant under smooth one-to-one transformations of the random variable.*

Proof. Let ϕ be a smooth one-to-one transformation of the random variable x to y.
From Equation (69)

$$\begin{aligned} p(x;\xi) &= q(\phi(x);\xi)\phi'(x) && (70) \\ \partial_i \ell(q_y) &= \partial_i \ell(p_{w(y)}) && (71) \\ \partial_i \ell(q_{\phi(x)}) &= \partial_i \ell(p_x) && (72) \end{aligned}$$

The Fisher information metric g' on the induced manifold $\mathcal{S}'$ is given by

$$\begin{aligned} g'_{ij}(q_\xi) &= \int_Y \partial_i \ell(q_y)\, \partial_j \ell(q_y)\, q(y;\xi) dy \\ &= \int_X \partial_i \ell(q_{\phi(x)})\, \partial_j \ell(q_{\phi(x)})\, q(\phi(x);\xi)\, \phi'(x) dx \\ &= \int_X \partial_i \ell(p_x)\, \partial_j \ell(p_x)\, p(x;\xi) dx \\ &= g_{ij}(p_\xi) \end{aligned} \qquad (73)$$

which is the Fisher information metric on $\mathcal{S}$.
The components of Amari's $\alpha-$connections on the induced manifold $\mathcal{S}'$ are given by

$$\begin{aligned} \acute{\Gamma}^{\alpha}_{ijk}(q_\xi) &= \int_Y \partial_i\partial_j \ell(q_y)\, \partial_k \ell(q_y)\, q(y;\xi) dy\ + \\ &\quad \int_Y \frac{1-\alpha}{2} \partial_i \ell(q_y)\, \partial_j \ell(q_y)\, \partial_k \ell(q_y)\, q(y;\xi) dy \\ &= \int_X \partial_i\partial_j \ell(q_{\phi(x)})\, \partial_k \ell(q_{\phi(x)})\, q(\phi(x);\xi)\phi'(x) dx\ + \\ &\quad \int_X \frac{1-\alpha}{2} \partial_i \ell(q_{\phi(x)})\, \partial_j \ell(q_{\phi(x)})\, \partial_k \ell(q_{\phi(x)})\, q(\phi(x);\xi)\phi'(x) dx \\ &= \int_X \partial_i\partial_j \ell(p_x)\, \partial_k \ell(p_x)\, p(x;\xi) dx\ + \\ &\quad \int_X \frac{1-\alpha}{2} \partial_i \ell(p_x)\, \partial_j \ell(p_x)\, \partial_k \ell(p_x)\, p(x;\xi) dx \\ &= \Gamma^{\alpha}_{ijk}(p_\xi) \end{aligned} \qquad (74)$$

which are the components of Amari's $\alpha-$connections on the manifold $\mathcal{S}$. Thus we obtained that the Fisher information metric and Amari's α-connections are invariant under smooth one-to-one transformations of the random variable. □

Now we prove that α-connections are the only $F-$connections that are invariant under smooth one-to-one transformations of the random variable.

Theorem 4.5. *Amari's α-connections are the only $F-$connections that are invariant under smooth one-to-one transformations of the random variable.*

Proof. Let ϕ be a smooth one-to-one transformation of the random variable x to y.
The components of the $F-$connection of the induced manifold $\mathcal{S}'$ are

$$\begin{aligned} \acute{\Gamma}^F_{ijk}(q_\xi) &= \int_Y \left(\partial_i\partial_j\ell(q_y) + (1+\frac{qF''(q)}{F'(q)})\partial_i\ell(q_y)\ \partial_j\ell(q_y)\right)\partial_k\ell(q_y)\ q(y;\xi)dy \\ &= \int_X \partial_i\partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx + \\ &\quad \int_X (1+\frac{q(\phi(x);\xi)F''(q(\phi(x);\xi))}{F'(q(\phi(x);\xi))})\partial_i\ell(p_x)\ \partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx. \end{aligned} \tag{75}$$

and the components of the $F-$connection of the manifold $\mathcal{S}$ are

$$\begin{aligned} \Gamma^F_{ijk}(p_\xi) &= \int_X \partial_i\partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx + \\ &\quad \int_X (1+\frac{p(\phi(x);\xi)F''(p(x;\xi))}{F'(p(x;\xi))})\partial_i\ell(p_x)\ \partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx. \end{aligned} \tag{76}$$

Then by equating the components $\acute{\Gamma}^F_{ijk}(q_\xi)$, $\Gamma^F_{ijk}(p_\xi)$ of the $F-$connection, we get

$$\begin{aligned} &\int \frac{q(\phi(x);\xi)F''(q(\phi(x);\xi))}{F'(q(\phi(x);\xi))}\partial_i\ell(p_x)\ \partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx = \\ &\quad \int \frac{p(x;\xi)F''(p(x;\xi))}{F'(p(x;\xi))}\partial_i\ell(p_x)\ \partial_j\ell(p_x)\ \partial_k\ell(p_x)\ p(x;\xi)dx \end{aligned} \tag{77}$$

Then it follows that the condition for $F-$connection to be invariant under the transformation ϕ is given by

$$\frac{pF''(p)}{F'(p)} = k, \tag{78}$$

where k is a real constant.
Hence it follows from the Euler's homogeneous function theorem that the function F' is a positive homogeneous function in p of degree k. So

$$F'(\lambda p) = \lambda^k F'(p) \ \text{ for } \lambda > 0. \tag{79}$$

Since F' is a positive homogeneous function in the single variable p, without loss of generality we can take,

$$F'(p) = p^k. \tag{80}$$

Therefore

$$F(p) = \begin{cases} \frac{p^{k+1}}{k+1} & k \neq -1 \\ \log p & k = -1 \end{cases} \tag{81}$$

Let

$$k = \frac{-(1+\alpha)}{2},\ \alpha \in \mathbb{R}. \tag{82}$$

we get

$$F(p) = \begin{cases} \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}} & \alpha \neq 1 \\ \log p & \alpha = 1 \end{cases} \tag{83}$$

which is nothing but Amari's $\alpha-$embeddings $L_\alpha(p)$. Hence we obtain that Amari's $\alpha-$connections are the only $F-$connections that are invariant under smooth one-to-one transformations of the random variable. □

Remark 4.6. *In Section 2, we defined (F, G)-connections using a general embedding function F and a positive smooth function G. We can show that (F, G)-connection is invariant under smooth one-to-one transformation of the random variable when $G(p) = c$, where c is a real constant and $F(p) = L_\alpha(p)$ (proof is similar to that of Theorem 4.5). The notion of $(F, G)-$metric and $(F, G)-$connection provides a more general way of introducing geometric structures on a manifold. We were able to show that the Fisher information metric (up to a constant) and Amari's $\alpha-$connections are the only metric and connections belonging to this class that are invariant under both the transformation of the parameter and the one-to-one transformation of the random variable.*

5. Conclusions

The Fisher information metric and Amari's $\alpha-$connections are widely used in the theory of information geometry and have an important role in the theory of statistical estimation. Amari's $\alpha-$connections are defined using a one parameter family of functions, the $\alpha-$embeddings. We generalized this idea to introduce geometric structures on a statistical manifold $\mathcal{S}$. We considered a general embedding function F of $\mathcal{S}$ into $\mathbb{R}_X$ and obtained a geometric structure on $\mathcal{S}$ called the $F-$geometry. Amari's $\alpha-$geometry is a special case of $F-$geometry. A more general way of defining Riemannian metrics and affine connections on a statistical manifold $\mathcal{S}$ is given using a positive continuous function G and the embedding F.

Amari's $\alpha-$geometry is the only $F-$geometry that is invariant under both the transformation of the parameter and the random variable or equivalently under the sufficient statistic. We can relax the condition of invariance under the sufficient statistic and can consider other statistically significant transformations as well, which then gives an $F-$geometry other than $\alpha-$geometry that is invariant under these statistically significant transformations. We believe that the idea of $F-$geometry can be used in the further development of the geometric theory of q-exponential families. We look forward to studying these problems in detail later.

Acknowledgments

We are extremely thankful to Shun-ichi Amari for reading this article and encouraging our learning process. We would like to thank the reviewer who mentioned the references [13,16] that are of great importance in our future work.

Author Contributions

The authors contributed equally to the presented mathematical framework and the writing of the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Amari, S. Differential geometry of curved exponential families-curvature and information loss. *Ann. Statist.* **1982**, *10*, 357–385.
2. Amari, S. *Differential-Geometrical Methods in Statistics*; Lecture Notes in Statistics, Volume 28; Springer-Verlag: New York, NY, USA, 1985.
3. Amari, S.; Kumon, M. Differential geometry of Edgeworth expansions in curved exponential family. *Ann. Inst. Statist. Math.* **1983**, *35*, 1–24.
4. Amari, S.; Nagaoka, H. *Methods of Information Geometry, Translations of Mathematical Monographs*; Oxford University Press: Oxford, UK, 2000.
5. Barndorff-Nielsen, O.E.; Cox, D.R.; Reid, N. The role of differential geometry in statistical theory. *Internat. Statist. Rev.* **1986**, *54*, 83–96.
6. Dawid, A.P. A Discussion to Efron's paper. *Ann. Statist.* **1975**, *3*, 1231–1234.
7. Efron, B. Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.* **1975**, *3*, 1189–1242.
8. Efron, B. The geometry of exponential families. *Ann. Statist.* **1978**, *6*, 362–376.
9. Murray, M.K.; Rice, R.W. *Differential Geometry and Statistics*; Chapman & Hall: London, UK, 1995.
10. Rao, C.R. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta. Math. Soc.* **1945**, *37*, 81–91.
11. Chentsov, N.N. *Statistical Decision Rules and Optimal Inference*; Transted in English, Translation of the Mathematical Monographs; American Mathematical Society: Providence, RI, USA, 1982.
12. Corcuera, J.M.; Giummole, F. A characterization of monotone and regular divergences. *Ann. Inst. Statist. Math.* **1998**, *50*, 433–450.
13. Zhang, J. Divergence function, duality and convex analysis. *Neur. Comput.* **2004**, *16*, 159–195.
14. Burbea, J. Informative geometry of probability spaces. *Expo Math.* **1986**, *4*, 347–378.
15. Wagenaar, D.A. Information Geometry for Neural Networks. Available online: http://www.danielwagenaar.net/res/papers/98-Wage2.pdf (accessed on 13 December 2013).
16. Amari, S.; Ohara, A.; Matsuzoe, H. Geometry of deformed exponential families: Invariant, dually flat and conformal geometries. *Physica A* **2012**, *391*, 4308–4319.