A dually flat structure on the space of escort distributions

Atsumi Ohara
Department of Systems Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531 Japan
E-mail: ohara@sys.es.osaka-u.ac.jp

Hiroshi Matsuzoe
Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555 Japan
E-mail: matsuzoe@nitech.ac.jp

Shun-ichi Amari
Riken Brain Science Institute Wako-shi Hirosawa 2-1, Saitama 351-0198 Japan
E-mail: amari@brain.riken.jp

Abstract. This note studies geometrical structure of the manifold of escort probability distributions and proves that the resultant geometry is dually flat in the sense of information geometry. We use a conformal transformation that flattens the alpha-geometry of the space of the discrete probability distributions in order to realize escort probabilities in the framework of affine differential geometry. Dual pairs of potential functions and affine coordinate systems on the manifold are derived, and the associated canonical divergence is shown to be conformal to the alpha-divergence.

1. Introduction

Escort probabilities are naturally induced from the studies of multifractals [3] and non-extensive entropy (Tsallis entropy) [20, 21] to play an important but mysterious role. The purpose of this note is to investigate the escort probability from viewpoints of information geometry [1, 2] and affine differential geometry [12].

Recently, it is reported that $\alpha$-geometry, which is an information geometric structure with constant curvature, has a close relation with the Tsallis entropy [13, 14]. The main feature of the $\alpha$-geometry consists of a Riemannian metric together with a one-parameter family of dual affine connections, called the $\alpha$-connections.

We show that the manifold of escort probability distributions is dually flat by considering conformal transformations that flatten the $\alpha$-geometry on the manifold of usual probability distributions. On the resultant manifold, escort probabilities consist of an affine coordinate system. See also [17] for another type of flattening a curved dual manifold by a conformal transformation.
This gives us a clear geometrical interpretation of the escort probability, and simultaneously, produces its new obscure links to conformality and projectivity. Due to these two geometrical concepts, however, the obtained dually flat structure inherits several properties of the \(\alpha\)-geometry. Further, we show a pair of potential functions and derive dual affine coordinate systems for the structure.

Section 2 is devoted to preliminaries for the \(\alpha\)-geometry in the light of the affine differential geometry. In section 3, we consider the conformal transformations and discuss properties of the obtained dually flat structure.

In the sequel, all the affine connections are assumed torsion-free. The exponent \(q\) is fixed to be \((1 - \alpha)/2\) for \(-1 < \alpha < 1\).

2. Preliminaries

We briefly introduce the \(\alpha\)-geometry via the affine differential geometry. See for details [13, 14]. Let \(S^n\) denote the \(n\)-dimensional probability simplex, i.e.,

\[
S^n := \left\{ p = (p_i) \left| p_i > 0, \sum_{i=1}^{n+1} p_i = 1 \right. \right\},
\]

and \(p_i, i = 1, \ldots, n + 1\) denote probabilities of \(n + 1\) states. We introduce the \(\alpha\)-geometric structure on \(S^n\). Let \(\{\partial_i\}, i = 1, \ldots, n\) be natural basis tangent vector fields on \(S^n\) defined by

\[
\partial_i := \frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_{n+1}}, \quad i = 1, \ldots, n,
\]

where \(p_{n+1} = 1 - \sum_{i=1}^{n} p_i\). Now we define a Riemannian metric \(g\) on \(S^n\) by the Fisher information matrix,

\[
g_{ij}(p) := g(\partial_i, \partial_j) = \frac{1}{p_i} \delta_{ij} + \frac{1}{p_{n+1}}
\]

\[
= \sum_{k=1}^{n+1} p_k (\partial_i \log p_k) (\partial_j \log p_k), \quad i, j = 1, \ldots, n.
\]

Further, define an affine connection \(\nabla^{(\alpha)}\) called the \(\alpha\)-connection, which is represented in its coefficients by

\[
\Gamma_{ij}^{(\alpha)k}(p) = \frac{1 + \alpha}{2} \left( \frac{1}{p_k} \delta_{ij} + p_k g_{ij} \right), \quad i, j, k = 1, \ldots, n,
\]

where \(\delta_{ij}\) is equal to one if \(i = j = k\) and zero otherwise. Then we have the \(\alpha\)-covariant derivative \(\nabla^{(\alpha)}\), which gives

\[
\nabla^{(\alpha)}_{\partial_i} \partial_j = \sum_{k=1}^{n} \Gamma_{ij}^{(\alpha)k} \partial_k,
\]

when it is applied to the vector fields \(\partial_i\) and \(\partial_j\).

There are two specific features for the \(\alpha\)-geometry on \(S^n\) defined in such a way. First, the triple \((S^n, g, \nabla^{(\alpha)})\) is a statistical manifold [8] (See the appendix for its definition), i.e., we can confirm that the following relation holds:

\[
X g(Y, Z) = g(\nabla^{(\alpha)}_X Y, Z) + g(Y, \nabla^{(-\alpha)}_X Z), \quad X, Y, Z \in \mathcal{X}(S^n),
\]

where \(\mathcal{X}(S^n)\) denotes the set of all tangent vector fields on \(S^n\). Two statistical manifolds \((S^n, g, \nabla^{(\alpha)}\) and \((S^n, g, \nabla^{(-\alpha)})\) are said mutually dual.
The other is that \((S^n, g, \nabla^{(\alpha)})\) is a manifold with constant curvature \(\kappa = (1 - \alpha^2)/4\), i.e.,

\[
R^{(\alpha)}(X, Y)Z = \kappa\{g(Y, Z)X - g(X, Z)Y\},
\]

where \(R^{(\alpha)}\) is the curvature tensor with respect to \(\nabla^{(\alpha)}\). From this property the well-known nonadditive formula of the Tsallis entropy can be derived \[13\].

In \[14\] we have discussed the \(\alpha\)-geometry on \(S^n\) from a viewpoint of the affine differential geometry \[11\]. Consider the immersion \(f\) of \(S^n\) into \(\mathbb{R}^{n+1}_+\) by

\[
f : p = (p_i) \mapsto x = (x^i) = (L^{(\alpha)}(p_i)), \quad i = 1, \ldots, n + 1,
\]

where \((x^i)\) is the canonical flat coordinate system of \(\mathbb{R}^{n+1}_+\) and the function \(L^{(\alpha)}\) is defined by

\[
L^{(\alpha)}(t) := \frac{2}{1 - \alpha} t^{(1 - \alpha)/2} = \frac{1}{q}.\]

Note that \(f(S^n)\) is a level hypersurface in the ambient space \(\mathbb{R}^{n+1}_+\) represented by \(\Psi(x) = 2/(1 + \alpha)\), where

\[
\Psi(x) := \frac{2}{\alpha + 1} \sum_{i=1}^{n+1} \left(1 - \alpha \frac{2}{\alpha} x^i\right)^{2/(1-\alpha)} = \frac{1}{1 - q} \sum_{i=1}^{n+1} (qx^i)^{1/q}.
\]

We choose a transversal vector \(\xi\) on the level hypersurface by

\[
\xi := \sum_{i=1}^{n+1} \xi^i \frac{\partial}{\partial x^i}, \quad \xi^i = -q(1 - q)x^i = -\kappa x^i.
\]

Then we can confirm that the affine immersion \((f, \xi)\) realizes the \(\alpha\)-geometry on \(S^n\) \[14\]. Hence, it would be possible to develop theory of the \(\alpha\)-geometry and Tsallis statistics with ideas of the affine differential geometry \[11\].

Further, the escort probability \[3\] naturally appears in this setup. The escort probability \(P = (P_i)\) associated with \(p = (p_i)^q\) is the normalized version of \((p_i)^q\), and is defined by

\[
P_i(p) := \frac{(p_i)^q}{\sum_{i=1}^{n+1} (p_i)^q} = \frac{x^i}{Z_q}, \quad Z_q(p) := \sum_{i=1}^{n+1} x^i(p), \quad x(p) \in f(S^n).
\]

Hence, the simplex \(E^n\) in the ambient space \(\mathbb{R}^{n+1}_+\), i.e.,

\[
E^n := \left\{ x = (x^i) \left| \sum_{i=1}^{n+1} x^i = 1, \ x^i > 0 \right. \right\}
\]

represents the set of escort distributions \(P\).

Note that the element \(x^* = (x^*_i)\) in the dual space of \(\mathbb{R}^{n+1}\) defined by

\[
x^*_i(p) := L^{(-\alpha)}(p_i) = \frac{1}{1 - q} (p_i)^{1-q}, \quad i = 1, \ldots, n + 1,
\]

meets

\[
x^*_i(p) = \frac{\partial \Psi}{\partial x^i}(x(p)).
\]

Hence, it satisfies \[14\]

\[
-\sum_{i=1}^{n+1} \xi^i(p)x^*_i(p) = 1, \quad \sum_{i=1}^{n+1} x^*_i(p)X^i = 0,
\]

for an arbitrary vector \(X = \sum_{i=1}^{n+1} X^i \partial/\partial x^i\) at \(x(p)\) tangent to \(f(S^n)\). Thus, \(-x^*(p)\) can be interpreted as the conormal map \[12\].
3. A conformally and projectively flat geometric structure and escort probabilities

In this section we consider a conformal and projective transformation [5, 6, 7, 10] of the \(\alpha\)-geometry to introduce a dually flat one. This flattening of the \(\alpha\)-geometry conserves some of its properties. The escort probabilities \(P_i\) are found to represent one of mutually dual affine coordinate systems in the induced geometry.

Let us define a function \(\lambda\) on \(S^n\) by

\[
\lambda(p) := \frac{1}{Z_q} = \frac{1}{\sum_{i=1}^{n+1} L^{(\alpha)}(p_i)}
\]

which depends on \(\alpha\). Then, from (9) \(E^n\) is regarded as the image of \(S^n\) for another immersion \(\tilde{f} := \lambda f\), i.e.,

\[
\tilde{f} : S^n \ni (p_i) \mapsto (\tilde{P}_i) \in E^n,
\]

and \((P_i)_{i=1}^n\) is interpreted as a coordinate system of \(S^n\). It would be a natural way to introduce geometric structure on \(E^n\) (and hence on \(S^n\)) via the affine immersion \((\tilde{f}, \tilde{\xi})\) by taking a suitable transversal vector \(\tilde{\xi}\), similarly to the case of the \(\alpha\)-geometry mentioned above. Since \(E^n\) is a part of a hyperplane in \(\mathbb{R}^{n+1}\), the canonical affine connection of \(\mathbb{R}^{n+1}\) induces a flat connection, denoted by \(D^{(E)}\), on \(E^n\). However, for the same reason, we cannot define a Riemannian metric in this way\(^1\) regardless of any choice of the transversal vector \(\tilde{\xi}\).

The idea we adopt here is to define a Riemannian metric by utilizing a property of \((\alpha)\)-conformal flatness. Based on the results proved by Kurose [5, 6], we conclude that the manifold \((S^n, g, \nabla^{(\alpha)})\) is \pm 1-conformally flat (See the appendix for its definition) because it is a statistical manifold with constant curvature.

Actually, let \(\nabla^\ast\) be the flat connection\(^2\) on \(S^n\) pulled back by \(\tilde{f}\) from \((E^n, D^{(E)})\). Then, we can prove that \(\nabla^{(\alpha)}\) and \(\nabla^\ast\) are projectively equivalent, i.e., it holds that

\[
\nabla^\ast Y = \nabla^{(\alpha)} Y + d(\ln \lambda)(Y) X + d(\ln \lambda)(X) Y, \quad X, Y \in \mathcal{X}(S^n).
\]  

Hence, if we define another Riemannian metric \(h\) on \(S^n\) by

\[
h(X, Y) := \lambda g(X, Y), \quad X, Y \in \mathcal{X}(S^n),
\]

then, \((S^n, g, \nabla^{(\alpha)})\) is \(-1\)-conformally equivalent to \((S^n, h, \nabla^\ast)\) having a flat connection \(\nabla^\ast\). Further, the manifold \((S^n, h, \nabla^\ast)\) can be proved to be a statistical manifold.

Using the conormal map \(-x^\ast(p)\), we can define the \(\alpha\)-divergence as a contrast function\(^3\) on \((S^n, g, \nabla^{(\alpha)})\) as follows [6]:

\[
D^{(\alpha)}(p, r) = -\sum_{i=1}^{n+1} x_i^\ast(r)(x_i(p) - x_i(r))
\]

\[= \langle -x^\ast(r), x(p) - x(r) \rangle = \frac{1}{\kappa} - \langle x^\ast(r), x(p) \rangle.\]

The statistical manifolds \((S^n, g, \nabla^{(-\alpha)})\) and \((S^n, g, \nabla^{(\alpha)})\) are dual in the sense of (5). Further, it is known [2] that there exists the unique dual affine flat connection \(\nabla^\ast\) on \(S^n\) with respect to \((h, \nabla^\ast)\). Then, \((S^n, g, \nabla^{(-\alpha)})\) is proved \(1\)-conformally equivalent to \((S^n, h, \nabla^\ast)\) [6]. This fact

\(^1\) In the affine differential geometry, a Riemannian metric is realized as the affine fundamental form of an affine immersion [12].

\(^2\) For the sake of notational consistency with the existing literature, e.g., [1, 2], we first define \(\nabla^\ast\), and later \(\nabla\) as the dual of \(\nabla^\ast\).

\(^3\) A divergence that is compatible with statistical manifold structure in the sense of Eguchi [4].
directly implies [6] that a contrast function \( \rho \) of \((S^n, h, \nabla)\) is given by scaling \( D^{(-\alpha)} \) that is a contrast function of \((S^n, g, \nabla^{(-\alpha)})\), i.e.,

\[
\rho(p, r) = \lambda(r) D^{(-\alpha)}(p, r) = \frac{1}{Z_q(r)} D^{(-\alpha)}(p, r) = \frac{1}{Z_q(r)} \langle -x(r), x^*(p) - x^*(r) \rangle = \langle -P(r), x^*(p) - x^*(r) \rangle.
\]

(13)

We shall call \( \rho \) a conformal divergence.

Since \((S^n, h, \nabla, \nabla^*)\) is a dually flat structure, it is equipped with mutually dual affine coordinate systems \((\theta^i)\) and \((\eta_i)\), \(i = 1, \ldots, n\), a potential function \(\psi(\theta)\) and its conjugate \(\psi^*(\eta)\) [1, 2] satisfying

\[
\eta_i = \frac{\partial \psi}{\partial \theta^i}, \quad \theta^i = \frac{\partial \psi^*}{\partial \eta_i}.
\]

Further, a contrast function is written in the form of the canonical divergence [2], i.e.,

\[
\rho(p, r) = \psi(\theta(p)) + \psi^*(\eta(r)) - \sum_{i=1}^{n} \theta^i(p) \eta_i(r).
\]

(14)

From these facts, we see that the following holds:

**Proposition 1** For the dually flat structure \((S^n, h, \nabla, \nabla^*)\) defined via \(\pm 1\)-conformal transformation from \((S^n, g, \nabla^{(\alpha)}, \nabla^{(-\alpha)})\), the associated potential functions and dually flat affine coordinate systems are represented as follows:

\[
\eta_i(p) = P_i(p), \quad \theta^i(p) = x^*_i(p) - x^*_{n+1}(p), \quad i = 1, \ldots, n
\]

\[
\psi^*(\eta(p)) = \frac{1}{\kappa} (\lambda(p) - q), \quad \psi(\theta(p)) = -\ln_q(p_{n+1}),
\]

where \(\ln_q(t) := (t^{1-q} - 1)/(1 - q)\), and \(\kappa = (1 - \alpha^2)/4 = q(1 - q)\) is the scalar curvature. Here, the coordinate systems \((\theta^i)\) and \((\eta_i)\) are \(\nabla\)- and \(\nabla^*\)-affine, respectively.

**Proof** We have only to check that the canonical divergence coincides with a contrast function of \((S^n, h, \nabla)\). Substitute the potential functions and dual affine coordinates in the statement to the right-hand side of (14), then we see that it coincides with \(\rho(p, r)\) in (13). Q.E.D.

**Corollary 1** The escort probabilities \(P_i, i = 1, \ldots, n\) are canonical affine coordinates of the flat affine connection \(\nabla^*\) on \(S^n\).

There are three remarks. First, since it holds that

\[
\ln_q(p_i) = x^*_i(p) - \frac{1}{1 - q},
\]

we can alternatively represent

\[
\theta^i(p) = \ln_q(p_i) - \ln_q(p_{n+1}).
\]

Further, note that the conformal factor \(\lambda\) is alternatively represented as follows:

\[
\lambda(p) = \frac{1}{Z_q(p)} = \frac{q}{\sum_{i=1}^{n+1} (p_i)^q} = \frac{q}{(\exp_q(S_q(p)))^{1-q}} = \frac{\kappa \ln_q \left( \frac{1}{\exp_q(S_q)} \right)}{q},
\]

(15)
where
\[ S_q(p) := \frac{\sum_{i=1}^{n+1}(p_i)^q - 1}{1 - q}, \quad \exp_q(t) := [1 + (1 - q)t]^{1/(1-q)}. \]
Substituting (15) into the expression of \( \psi^* \), we have
\[ \psi^* = \ln_q \left( \frac{1}{\exp_q(S_q)} \right). \]
We easily see that the expressions of \( \theta^i \) and \( \psi^* \) here and those of \( \eta_i \) and \( \psi^* \) in the proposition recover the standard ones [1, 2] when \( q \to 1 \). Note that \( -\psi^* \) coincides with the entropy studied in [9, 18, 22] and referred to as the normalized Tsallis entropy. The conformal factor (or scaling factor) \( \lambda \) often appears in the study of the \( q \)-analysis.

Next, similarly to the above conformal transformation of \((S^n, g, \nabla^{(\alpha)})\), we can define another one for \((S^n, g, \nabla^{(-\alpha)})\) with a conformal factor
\[ \lambda'(p) := \frac{1}{\sum_{i=1}^{n+1} L^{(\alpha)}(p_i)}, \]
and construct another dually flat structure \((S^n, h' = \lambda' g, \nabla', \nabla'^*)\). Hence, the relations among them are shown in the Figure 1.

Finally, because of the projective equivalence (11), a submanifold in \( S^n \) is \( \nabla^{(\alpha)} \)-autoparallel if and only if it is \( \nabla \)-autoparallel. In particular, distributions constrained with the normalized \( q \)-expectations (escort averages) [21] consist of a simultaneously \( \nabla^{(\alpha)} \)- and \( \nabla \)-autoparallel submanifold in \( S^n \).

4. Concluding remarks
We have considered \( \pm 1 \)-conformal transformations of the \( \alpha \)-geometry and obtained dually flat structure \((S^n, h, \nabla, \nabla^*)\). Further the potential functions and dually flat coordinate systems associated with the structure have been derived. We see that the escort probability naturally appears to play an important role.

From a viewpoint of contrast functions, the geometric structure compatible to the Kullback-Leibler divergence is \((S^n, g, \nabla^{(1)}, \nabla^{(-1)})\), where \( g \) is the Fisher information and \( \nabla^{(\pm 1)} \) are respectively the \( e \)-connection and the \( m \)-connection. Similarly, the \( \alpha \)-divergence (or the Tsallis relative entropy), and the conformal divergence \( \rho \) in this note correspond to \((S^n, g, \nabla^{(\alpha)}, \nabla^{(-\alpha)})\) and \((S^n, h, \nabla, \nabla^*)\), respectively. They are summarized in Figure 2. The physical meaning or essence underlying these transformations would be interesting and significant.

We hope that arguments in this note might be helpful to more profound investigations and understandings of this area [15, 16, 19].
is known that (for arbitrary $\alpha$)

For details of this appendix see [8, 5, 6, 7, 10]. For a torsion-free affine connection $\nabla$ on a manifold $\mathcal{M}$, the triple $(\mathcal{M}, g, \nabla)$ is called a statistical manifold if it admits another torsion-free connection $\nabla^*$ satisfying

$$X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z) \quad (A.1)$$

for arbitrary $X, Y$ and $Z$ in $\mathcal{X}(\mathcal{M})$, where $\mathcal{X}(\mathcal{M})$ is the set of all tangent vector fields on $\mathcal{M}$. It is known that $(\mathcal{M}, g, \nabla)$ is a statistical manifold if and only if $\nabla g$ is symmetric. We call $\nabla$ and $\nabla^*$ duals of each other with respect to $g$, and $(\mathcal{M}, g, \nabla^*)$ is said the dual statistical manifold of $(\mathcal{M}, g, \nabla)$. The triple of a Riemannian metric and a pair of dual connections $(g, \nabla, \nabla^*)$ satisfying (A.1) is called a dualistic structure on $\mathcal{M}$.

For $\alpha \in \mathbb{R}$, statistical manifolds $(\mathcal{M}, g, \nabla)$ and $(\mathcal{M}, g', \nabla')$ are said to be $\alpha$-conformally equivalent if there exists a function $\phi$ on $\mathcal{M}$ such that

$$g'(X, Y) = e^\phi g(X, Y),
\quad g(\nabla^*_X Y, Z) = g(\nabla^*_X Y, Z) - \frac{1 + \alpha}{2} d\phi(Z) g(X, Y) + \frac{1 - \alpha}{2} \{d\phi(X) g(Y, Z) + d\phi(Y) g(X, Z)\}.$$ 

Statistical manifolds $(\mathcal{M}, g, \nabla)$ and $(\mathcal{M}, g', \nabla')$ are $\alpha$-conformally equivalent if and only if $(\mathcal{M}, g, \nabla^*)$ and $(\mathcal{M}, g, \nabla')$ are $-\alpha$-conformally equivalent.

A statistical manifold $(\mathcal{M}, g, \nabla)$ is called $\alpha$-conformally flat if it is locally $\alpha$-conformally equivalent to a flat statistical manifold. Note that $-1$-conformal equivalence implies projective equivalence. A statistical manifold of dimension greater than three has constant curvature if and only if it is $\pm 1$-conformally flat. If $(\mathcal{M}, g, \nabla)$ and $(\mathcal{M}, g', \nabla')$ are $1$-conformally equivalent, a contrast function $\rho'$ of $(\mathcal{M}, g', \nabla')$ is given by

$$\rho'(p, q) = e^{\phi(q)} \rho(p, q),$$

where $\rho$ is a contrast function of $(\mathcal{M}, g, \nabla)$.

Acknowledgments
We would thank the anonymous referee who has pointed out several errors in the manuscript and important literature.

Appendix: Statistical manifold and $\alpha$-conformally equivalence
For details of this appendix see [8, 5, 6, 7, 10]. For a torsion-free affine connection $\nabla$ and a pseudo Riemannian metric $g$ on a manifold $\mathcal{M}$, the triple $(\mathcal{M}, g, \nabla)$ is called a statistical manifold if it admits another torsion-free connection $\nabla^*$ satisfying

$$\alpha$$-divergence

$$\text{KL divergence} \quad (S^n, g, \nabla^{(i)}, \nabla^{(-1)}) \quad \longleftrightarrow \quad \text{\alpha-divergence} \quad (S^n, g, \nabla^{(\alpha)}, \nabla^{(-\alpha)}) \quad \longleftrightarrow \quad \text{conformal divergence} \quad (S^n, h, \nabla, \nabla^*) \quad (S^n, h', \nabla', \nabla'^*)$$

dually flat

constant curvature $\kappa$

dually flat

Figure 2. transformations of dualistic structures

References
[1] Amari S-I 1985 Differential-Geometrical Methods in Statistics, (Berlin Springer-Verlag)
[2] Amari S-I and Nagaoka H 2000 Methods of Information Geometry, (Rhode Island: AMS&Oxford).
[3] Beck C and Schlögl F 1993 Thermodynamics of Chaotic Systems (Cambridge: Cambridge University Press).
[4] Eguchi S 1992 Geometry of minimum contrast Hiroshima Math. J. 23 631–47
[5] Kurose T 1990 Dual Connections and Affine Geometry Math. Z. 203 115-21
[6] Kurose T 1994 On the Divergence of 1-conformally Flat Statistical Manifolds, Tôhoku Math. J. 46 427-33
[7] Kurose T 2002 Conformal-projective geometry of statistical manifolds. Interdiscip. Inform. Sci. 8 89–100
[8] Lauritzen S L 1987 Statistical manifolds S. Amari et al. eds. *Differential geometry in statistical inference* 163–216 (Hayward: Institute of Mathematical Statistics)

[9] Landsberg P T and Vedral V 1998 Distributions and channel capacities in generalized statistical mechanics *Phys. Lett. A* 247 211–7

[10] Matsuzoe H 1999 Geometry of contrast functions and conformal geometry *Hiroshima Math. J.* 29 175–91

[11] Matsuzoe H, Takeuchi J and Amari S-I 2006 Equiaffine structures on statistical manifolds and Bayesian statistics *Differential Geom. Appl.* 24 567–78

[12] Nomizu K and Sasaki T 1993 *Affine Differential Geometry* (Cambridge: Cambridge University Press)

[13] Ohara A 2007 Geometry of distributions associated with Tsallis statistics and properties of relative entropy minimization *Phys. Lett. A* 370 184–93

[14] Ohara A 2009 Geometric study for the Legendre duality of generalized entropies and its application to the porous medium equation *Euro. Phys. J. B* 70 15–28

[15] Ojima I 2004 Temperature as order parameter of broken scale invariance *Publ. RIMS (Kyoto Univ.)* 40 731-56

[16] Ojima I 2009 Meaning of Non-Extensive Entropies in Micro-Macro Duality *to be included in the present proceedings*

[17] Okamoto I, Amari S and Takeuchi K 1991 Asymptotic theory of sequential estimation procedures for curved exponential families *Annals of Statistics* 19 961-81

[18] Rajagopal A K and Abe S 1999 Implications of Form Invariance to the Structure of Nonextensive Entropies *Phys. Rev. Lett.* 83 1711-4

[19] Tanaka M 2009 Meaning of an escort distribution and $\tau$-transformation *to be included in the present proceedings*

[20] Tsallis C 1988 Possible generalization of Boltzmann-Gibbs statistics *J. Stat. Phys.* 52, 479-87

[21] Tsallis C 2009 *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Berlin/Heidelberg: Springer)

[22] Wada T and Scarfone A M 2005 Connections between Tsallis’ formalisms employing the standard linear average energy and ones employing the normalized $q$-average energy *Phys. Lett. A* 335 351–62