International Journal of Theoretical Physics, Vol. 43, Nos. 7/8, August 2004 (© 2004)

Generalized Relative Entropies as Contrast Functionals on Density Matrices

Anna Jenčová[1]

We use a class of generalized relative entropies on density matrices to obtain one-parameter families of torsion-free affine connections.

KEY WORDS: generalized relative entropies; information geometry; affine connections.

1. INTRODUCTION

The aim of quantum information geometry is to introduce the quantum counterparts of the basic structures of the classical theory, namely Riemannian metrics and families of affine connections. It is an important feature of the classical information manifolds, that if invariancy with respect to bijective transformations of the sample space is required, then these structures are unique (up to a multiplication factor): the Fisher metric and the family of Chentsov-Amari α-connections (Amari, 1985; Chentsov, 1982).

Let $\mathcal{F} = \{p(\cdot, \theta) | \theta \in \Theta\}$ be a manifold of classical probability densities with respect to a common measure P. To define the affine connections, Amari (1985) used a family of functions

$$f_\alpha(x) = \begin{cases} \dfrac{2}{1-\alpha} x^{\frac{1-\alpha}{2}} & \alpha \neq 1 \\ \log(x) & \alpha = 1 \end{cases} \tag{1}$$

Let $l_\alpha(x, \theta) = f_\alpha(p(x, \theta))$. The coefficients of the Fisher information metric tensor and the α-connections are given by

$$g_{ij}(\theta) = \int \partial_i l_\alpha(x, \theta) \partial_j l_{-\alpha}(x, \theta)\, dP, \quad \forall \alpha$$

[1] Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, SK-814 73 Bratislava, Slovakia; e-mail: jenca@mat.savba.sk.

$$\Gamma^{\alpha}_{ijk}(\theta) = \int \partial_i \partial_j l_{\alpha}(x, \theta) \partial_k l_{-\alpha}(x, \theta)\, dP$$

These connection are torsion-free and the α and $-\alpha$ connections are dual with respect to the Fisher metric, in the sense that if $\nabla^{\pm\alpha}$ are the covariant derivatives and X, Y, Z are vector fields, then

$$Xg(Y, Z) = g\left(\nabla^{\alpha}_X Y, Z\right) + g\left(Y, \nabla^{-\alpha}_X Z\right)$$

There are more equivalent ways to introduce the Fisher metric and the affine connections. In the present paper, we follow the approach of Eguchi (1983), who used contrast functionals, see also Amari (1985).

A functional ρ over $\mathcal{F} \times \mathcal{F}$ is called a contrast functional if

(i) $\rho(\theta_1, \theta_2) \geq 0$ for all $\theta_1, \theta_2 \in \Theta$
(ii) $\rho(\theta_1, \theta_2) = 0$ if and only if $\theta_1 = \theta_2$

The Riemannian metric and Christoffel symbols of the affine connections are defined by

$$g^{\rho}_{ij}(\theta) = -\frac{\partial^2}{\partial\theta_i \partial\theta'_j} \rho(\theta, \theta')|_{\theta=\theta'} \tag{2}$$

$$\Gamma^{\rho}_{ijk}(\theta) = -\frac{\partial^3}{\partial\theta_i \theta_j \theta'_k} \rho(\theta, \theta')|_{\theta=\theta'} \tag{3}$$

Let $f : (0, \infty) \to \mathbb{R}$ be a convex function satisfying $f(1) = 0$, then

$$\rho_f(\theta_1, \theta_2) = E_{\theta_1}\left[f\left(\frac{p(X, \theta_2)}{p(X, \theta_1)} \right) \right]$$

defines a contrast functional. It was shown that in this case, $g^{\rho}_{ij} = f''(1) g_{ij}$, where g_{ij} denotes the coefficients of the Fisher metric and the corresponding affine connection coincides with the α-connection with $\alpha = 2f'''(1) + 3f''(1)$.

As one would expect, the situation is different in noncommutative case. Here, the equivalent of the Fisher metric would be a Riemannian metric, which is monotone with respect to completely positive trace preserving maps. For manifolds of $n \times n$ density matrices, it was proved by Chentsov and Morozova (1990) that such metric is not unique. Later, Petz (1996) characterized the class of monotone metrics in terms of operator monotone functions. Nagaoka (1994) defines the affine α-connection for $\alpha = -1$ (the mixture connection) using the natural flat affine structure on density matrices. The exponential connection is defined as its dual with respect to the given monotone metric. This approach was generalized in Jenčová (2001a), for all α. Unlike the classical case, the dual connections are not torsion free in general. In Jenčová (2001b), it was shown that the dual connection to the α-connection is torsion-free only for a special monotone metric λ^{α}.

Lesniewski and Ruskai (1999) used a class of generalized relative entropies, defined in Petz (1986), as contrast functionals on (non-normalized) density matrices. It was shown that each monotone metric can be obtained in the form (2) for a certain convex subset of relative entropies. The aim of the following paper is to use this subset to obtain a class of torsion free α-connections, such that α and -α-connections are dual. We question the coincidence with the Fisher metric and classical affine α-connections on commutative submanifolds, use the language of statistical manifolds by Lauritzen (Amari *et al.*, 1987), to give formulas for the Riemannian curvature tensor. We also treat some important examples.

2. GENERALIZED RELATIVE ENTROPIES AND MONOTONE METRICS

Let $\mathcal{D}$ denote the set of $n \times n$ complex Hermitian matrices and let $\mathcal{D}^+$ be the subset of positive definite matrices. As an open subset in $\mathcal{D}$, $\mathcal{D}^+$ inherits a natural affine parametrization and has the structure of a differentiable manifold. Let T_ρ be the tangent space at ρ and let λ be the monotone Riemannian metric. Then λ is of the form (Petz, 1996)

$$\lambda_\rho(X, Y) = \operatorname{Tr} X J_\rho(Y), \quad J_\rho^{-1} = f(L_\rho R_\rho^{-1}) R_\rho$$

where $f : (0, \infty) \to \mathbb{R}$ is an operator monotone function satisfying $f(t) = t f(t^{-1})$ and a normalization condition $f(1) = 1$, L_ρ and R_ρ are the left and right multiplication operator, respectively.

Let $\mathcal{G}$ be the set of operator convex functions $g : (0, \infty) \to \mathbb{R}$, satisfying $g(1) = 0$ and $g''(1) = 1$. It is known that each operator convex function with $g(1) = 0$ can be written in the form

$$g(w) = a(w-1) + b(w-1)^2 + c\frac{(w-1)^2}{w} + \int_0^\infty (w-1)^2 \frac{1+s}{w+s}\, d\mu(s) \quad (4)$$

where $b, c \geq 0$ and μ is a positive finite measure on $(0, \infty)$. The value of $a \in \mathbb{R}$ does not influence any of the following structures and therefore two functions in $\mathcal{G}$ that differ only in a will be treated as equal.

Let $\mathcal{P}$ be the set of positive finite measures μ on $[0, \infty]$, such that $\int_{[0,\infty]} d\mu = \frac{1}{2}$. Then (4) establishes a one-to-one correspondence between $\mathcal{G}$ and $\mathcal{P}$, with $c = \mu(\{0\})$, $b = \mu(\{\infty\})$.

If g is an operator convex function, we define its transpose $\hat{g}$ by $\hat{g}(w) = w g(w^{-1})$. It is clear that $\hat{g} \in \mathcal{G}$ if $g \in \mathcal{G}$ and that $g \mapsto \hat{g}$ induces the map $\mathcal{P} \to \mathcal{P}$, given by $\mu \mapsto \hat{\mu}$, where $d\hat{\mu}(s) = d\mu(s^{-1})$.

If $g = \hat{g}$, we say that g is symmetric. The subset of symmetric functions in $\mathcal{G}$ will be denoted by $\mathcal{G}_{\text{sym}}$. Let $\sim$ be the equivalence relation on $\mathcal{G}$

$$g_1 \sim g_2 \iff g_1 + \hat{g}_1 = g_2 + \hat{g}_2.$$

The quotient space $\mathcal{G}|_\sim$ is isomorphic to $\mathcal{G}_{\text{sym}}$. Similarly, $\mathcal{P}_{\text{sym}}$ denotes the subset of measures symmetric with respect to the transform $s \mapsto s^{-1}$ and we have an equivalence relation $\sim$ on $\mathcal{P}$. Let us denote by $\mathcal{G}_h$ the equivalence class containing $\frac{1}{2}h$, where $\frac{1}{2}h \in \mathcal{G}_{\text{sym}}$, and similarly $\mathcal{P}_m$.

In Petz (1986), see also Lesniewski and Ruskai (1999), the following class of generalized relative entropies on $\mathcal{D}^+$ was introduced.

Definition 2.1. Let $g \in \mathcal{G}$. The relative g-entropy $H_g : \mathcal{D}^+ \times \mathcal{D}^+ \to \mathbb{R}$ is defined by

$$H_g(\rho, \sigma) = \operatorname{Tr} \rho^{\frac{1}{2}} g\left(\frac{L_\sigma}{R_\rho}\right)(\rho^{\frac{1}{2}})$$

Proposition 2.1. (Lesniewski and Ruskai, 1999).
Let $g \in \mathcal{G}$ and let a, b, c and μ be as above. Then

$$\begin{aligned} H_g(\rho, \sigma) &= a\operatorname{Tr}(\sigma - \rho) \\ &\qquad + \operatorname{Tr}(\sigma - \rho)\left\{b\rho^{-1} + c\sigma^{-1} + \int_0^\infty \frac{1+s}{L_\sigma + sR_\rho}\, d\mu(s)\right\}(\sigma - \rho) \\ &= a\operatorname{Tr}(\sigma - \rho) + \operatorname{Tr}(\sigma - \rho)R_\rho^{-1} k\left(\frac{L_\sigma}{R_\rho}\right)(\sigma - \rho) \end{aligned}$$

where

$$k(w) = \int_{[0,\infty]} \frac{1+s}{w+s}\, d\mu(s) = \frac{g(w) - a(w-1)}{(w-1)^2}.$$

The relative g-entropy can be used to define a Riemannian structure on $\mathcal{D}^+$ as follows. Let $X, Y \in T_\rho$, then

$$\lambda_\rho(X, Y) = -\frac{\partial^2}{\partial s \partial t} H_g(\rho + sX, \rho + tY)|_{s=t=0} = \operatorname{Tr} X R_\rho^{-1} k_{\text{sym}}\left(\frac{L_\rho}{R_\rho}\right)(Y)$$

where

$$k_{\text{sym}}(w) = k(w) + w^{-1}k(w^{-1}) = \frac{g(w) + \hat{g}(w)}{(w-1)^2}.$$

It was proved that this defines a monotone metric, where the corresponding operator monotone function is $f = 1/k$. Conversely, for a given monotone metric, we may put $g(w) = \frac{(w-1)^2}{f(w)}$. The condition $g''(1) = 1$ is equivalent to the normalization condition $f(1) = 1$. Thus we have

Proposition 2.2. *There is a one-to-one correspondence between monotone Riemannian metrics and equivalence classes $\mathcal{G}_h$.*

3. AFFINE CONNECTIONS

Let $\theta \in \Theta \subseteq \mathbb{R}^N$ be a smooth parameter in $\mathcal{D}^+$ and let $\partial_i = \frac{\partial}{\partial \theta_i}$. Let us fix a monotone Riemannian metric λ on $\mathcal{D}^+$ and let $\mathcal{G}_h$ be the corresponding equivalence class. Let us choose a function $g \in \mathcal{G}_h$. In correspondence with the classical theory, we define the affine connections ∇^g by

$$\Gamma^g_{ijk}(\theta) = \lambda_\theta(\nabla_{\partial_i}\partial_j, \partial_k) = -\partial_i\partial_j \frac{\partial}{\partial \theta'_k} H_g(D(\theta), D(\theta'))|_{\theta=\theta'}$$

It is easy to show that this satisfies the transformation rules of an affine connection.

Proposition 3.1. *Let $g \in \mathcal{G}_h$. Then the connections ∇^g and $\nabla^{\hat{g}}$ are dual with respect to λ. Moreover, the connections are torsion-free.*

Proof: Consider the natural flat affine structure in $\mathcal{D}^+$ and let X be a vector field, parallel with respect to this affine structure, then X is constant over $\mathcal{D}^+$. As there is no danger of confusion, we will denote its value $X_\rho \in \mathcal{D}$ at ρ by the same letter. Let X, Y, Z be such vector fields. If $g \in \mathcal{G}_h$, then clearly $\hat{g} \in \mathcal{G}_h$ and $H_{\hat{g}}(\rho, \sigma) = H_g(\sigma, \rho)$, so that

$$\lambda_\rho\big(\nabla^{\hat{g}}_X Y, Z\big) = -\frac{\partial^3}{\partial t \partial s \partial u} H_g(\rho + uZ, \rho + sX + tY)|_{s=t=u=0}$$

Using the previous section, we get

$$\begin{aligned} X\lambda_\rho(Y, Z) &= \frac{d}{dt}\lambda_{\rho+tX}(Y, Z)|_{t=0} \\ &= \frac{d}{dt}\left(-\frac{\partial^2}{\partial s \partial u} H_g(\rho + tX + sY, \rho + tX + uZ)|_{s=u=0}\right)_{t=0} \\ &= \lambda_\rho\big(\nabla^g_X Y, Z\big) + \lambda_\rho\big(Y, \nabla^{\hat{g}}_X Z\big) \end{aligned}$$

so that the connections are dual. Torsion-freeness is obvious. □

Let $c_g : (0, \infty) \times (0, \infty) \to \mathbb{R}$ be given by $c_g(x, y) = \frac{1}{y}k(\frac{x}{y})$, where k is as in Proposition 2.1. Note that $c_g(y, x) = c_{\hat{g}}(x, y)$ is obtained from $w^{-1}k(w^{-1})$ and that $c_g^{\mathrm{sym}}(x, y) = c_g(x, y) + c_g(y, x) = \frac{1}{y}k_{\mathrm{sym}}(\frac{x}{y})$ is the Morozova-Chentsov function. From Proposition 2.1, we get

$$H_g(\rho, \sigma) = a\mathrm{Tr}\,(\sigma - \rho) + \mathrm{Tr}\,(\sigma - \rho)c_g(L_\sigma, R_\rho)(\sigma - \rho) \tag{5}$$

For $\sigma, \rho \in \mathcal{D}^+$, the operator $c_g(L_\sigma, R_\rho)$ is positive on the space of $n \times n$ complex matrices with Hilbert-Schmidt inner product $\langle X, Y\rangle = \mathrm{Tr}\, X^*Y$.

Proposition 3.2. *Let*

$$T_s(X,Y,Z) = (1+s)\,\mathrm{Tr}\,X\frac{1}{sR_\rho + L_\rho}(Y)\frac{1}{R_\rho + sL_\rho}(Z)$$

and let $T_\infty = \lim_{s\to\infty} T_s$. *Then*

$$\lambda_\rho\big(\nabla^g_X Y, Z\big) = 2\int_{[0,\infty]} \Re T_s(Z,X,Y)\,d\mu(s) -$$

$$-2\int_{[0,\infty]} (\Re T_s(Y,X,Z) + \Re T_s(X,Y,Z))\,d\hat{\mu}(s)$$

Proof: From (5), we get

$$\lambda_\rho\left(\nabla^g_X Y, Z\right) = -\frac{d}{ds}\mathrm{Tr}\,\{Xc_g(L_{\rho+sZ}, R_\rho)(Y) + Yc_g(L_{\rho+sZ}, R_\rho)(X)$$
$$- Xc_g(L_\rho, R_{\rho+sY})(Z) - Zc_g(L_\rho, R_{\rho+sY})(X)$$
$$- Yc_g(L_\rho, R_{\rho+sX})(Z) - Zc_g(L_\rho, R_{\rho+sX})(Y)\}|_{s=0}$$

Further, for $\rho,\sigma \in \mathcal{D}^+$ and $X, Y \in \mathcal{D}$,

$$\mathrm{Tr}\,Yc_g(L_\sigma, R_\rho)(X) = \langle Y, c_g(L_\sigma, R_\rho)(X)\rangle = \langle c_g(L_\sigma, R_\rho)(Y), X\rangle$$
$$= \langle X, c_g(L_\sigma, R_\rho)(Y)\rangle^- = \mathrm{Tr}\,Xc_g(L_\sigma, R_\rho)(Y)^-$$

It follows that

$$\lambda_\rho\big(\nabla^g_X Y, Z\big) = -2\Re\mathrm{Tr}\left\{X\frac{d}{ds}c_g(L_{\rho+sZ}, R_\rho)(Y)\right.$$
$$\left. - \left[X\frac{d}{ds}c_g(L_\rho, R_{\rho+sY})(Z) + Y\frac{d}{ds}c_g(L_\rho, R_{\rho+sX})(Z)\right]\right\}|_{s=0}$$

We have

$$c_g(x,y) = \mu(\{0\})x^{-1} + \mu(\{\infty\})y^{-1} + \int_0^\infty \frac{1+s}{x+sy}\,d\mu(s) \qquad (6)$$

Let us first suppose that $\mu(\{0\}) = \mu(\{\infty\}) = 0$. Then we compute

$$\frac{d}{dt}c_g(L_{\rho+tZ}, R_\rho)|_0 = -\int_0^\infty (1+s)\frac{1}{L_\rho + sR_\rho}L_Z\frac{1}{L_\rho + sR_\rho}\,d\mu(s)$$
$$\frac{d}{dt}c_g(R_\rho, L_{\rho+tY})|_0 = -\int_0^\infty s(1+s)\frac{1}{R_\rho + sL_\rho}L_Y\frac{1}{R_\rho + sL_\rho}\,d\mu(s)$$

so that

$$-\frac{d}{dt}\mathrm{Tr}\,Xc_g(L_{\rho+tZ}, R_\rho)(Y)|_0$$

$$= \int_0^\infty (1+s)\mathrm{Tr}\,\frac{1}{R_\rho + sL_\rho}(X)Z\frac{1}{sR_\rho + L_\rho}(Y)\,d\mu(s)$$

$$= \int_0^\infty T_s(Z, Y, X)\,d\mu(s)$$

and

$$-\frac{d}{dt}\mathrm{Tr}\,Xc_g(R_\rho, L_{\rho+tY})(Z)|_0$$

$$= \int_0^\infty s(1+s)\mathrm{Tr}\,\frac{1}{sR_\rho + L_\rho}(X)Y\frac{1}{R_\rho + sL_\rho}(Z)\,d\mu(s)$$

$$= \int_0^\infty (1+s)\mathrm{Tr}\,\frac{1}{R_\rho + sL_\rho}(X)Y\frac{1}{sR_\rho + L_\rho}(Z)\,d\hat{\mu}(s)$$

$$= \int_0^\infty T_s(Y, Z, X)\,d\hat{\mu}(s)$$

It follows that for each $s \in [0, \infty)$,

$$2\Re T_s(X, Y, Z) = T_s(X, Y, Z) + T_s(X, Z, Y)$$

so that $\Re T_s$ is a covariant 3-tensor, symmetric in last two variables. The statement now follows easily.

Let μ be concentrated in 0 and ∞. It is clear that $T_\infty = 0$ and we obtain by a direct computation from (6) that

$$\lambda_\rho\big(\nabla^g_X Y, Z\big) = \mu(\{0\})(T_0(Z, Y, X) + T_0(Z, X, Y)) - \mu(\{\infty\})(T_0(Y, X, Z) + T_0(Y, Z, X) + T_0(X, Y, Z) + T_0(X, Z, Y)) \qquad \square$$

3.1. Families of Connections

Let $\mathcal{G}_h$ be the equivalence class corresponding to the monotone metric λ. Let $g \in \mathcal{G}_h$. If g is symmetric, then the connection ∇^g is self dual and torsion free, hence it is the metric connection. If λ is fixed, we denote the metric connection by $\bar{\nabla}$.

Let $g \neq \hat{g}$. As $\mathcal{G}_h$ is a convex set, it contains all the functions

$$g_\alpha = \frac{1-\alpha}{2}g + \frac{1+\alpha}{2}\hat{g}$$

for $\alpha \in [-1, 1]$. If λ and g are fixed, we denote the corresponding connection by ∇^α. Then

$$\nabla^\alpha = \frac{1-\alpha}{2}\nabla + \frac{1+\alpha}{2}\nabla^*$$

where ∇ and ∇^* are the covariant derivatives corresponding to g and $\hat{g}$, respectively. The connections ∇^α and $\nabla^{-\alpha}$ are dual with respect to λ, $\nabla^{-1} = \nabla$, $\nabla^1 = \nabla^*$ and $\nabla^0 = \bar{\nabla}$ for all g. Clearly, such family of α-connections depends on the choice of $g \in \mathcal{G}_h$ and is therefore not unique.

3.2. Commutative Submanifolds

Let ρ, X, Y, and Z be all mutually commuting. Then it is easy to see that $\lambda_\rho(X, Y) = \operatorname{Tr} \rho^{-1} XY$ and

$$\lambda_\rho\big(\nabla_X^g Y, Z\big) = -\frac{1+\alpha^*}{2} \operatorname{Tr} \rho^{-2} XYZ$$

where $\alpha^* = 2g'''(1) + 3$. This corresponds to the Fisher metric and the α^*-connection in the commutative case. It seems to be a natural question to ask if, for each λ, it is possible to obtain the α^*-connections at least for $\alpha^* \in [-1, 1]$, if restricted to commutative submanifolds. From the next proposition (and examples below) it follows that this is not true.

Let the Riemannian metric λ correspond to the equivalence class $\mathcal{G}_h$, resp. $\mathcal{P}_m$. Let $\mu_{\max}$ be a measure with supp $\mu \subseteq [1, \infty]$, such that $\mu_{\max}$ coincides with m on $(1, \infty]$ and $\mu_{\max}(\{1\}) = \frac{1}{2} m(\{1\})$. Then we have

Proposition 3.3. *Let $\mu_{\max}$ be as above and let $g_{\max}$ be the corresponding operator convex function. Then $g_{\max} \in \mathcal{G}_h$ and for each $g \in \mathcal{G}_h$, we have*

$$-3 \leq \hat{g}_{\max}'''(1) \leq g'''(1) \leq g_{\max}'''(1) \leq 0$$

Proof: First, it is easy to see that $\mu_{\max}$ is a positive finite measure and $\int_{[0,\infty]} d\mu_{\max} = \frac{1}{2} \int_{[0,\infty]} dm = \frac{1}{2}$. Moreover, $\hat{\mu}_{\max}$ is concentrated in $[0, 1]$, $\hat{\mu}_{\max}$ coincides with m on $[0, 1)$ and $\hat{\mu}_{\max}(\{1\}) = \frac{1}{2} m(\{1\})$, so that $\mu_{\max} + \hat{\mu}_{\max} = m$. It follows that $g_{\max} \in \mathcal{G}_h$. Let now $g \in \mathcal{G}_h$ and let $\mu \in \mathcal{P}_m$ be the corresponding measure. Then

$$g'''(1) = -6 \int_{[0,\infty]} \frac{1}{1+s}\, d\mu(s)$$

and

$$\int_{[0,\infty]} \frac{1}{1+s}\, d\mu(s) = \int_{(1,\infty]} \frac{s}{1+s}\, d\mu(s^{-1}) + \frac{1}{2}\mu(\{1\}) + \int_{(1,\infty]} \frac{1}{1+s}\, d\mu(s)$$

$$\geq \int_{(1,\infty]} \frac{1}{1+s}\, dm(s) + \frac{1}{4} m(\{1\}) = \int_{[0,\infty]} \frac{1}{1+s}\, d\mu_{\max} \geq 0$$

and similarly,

$$\int_{[0,\infty]} \frac{1}{1+s} d\mu(s) \le \int_{[0,\infty]} \frac{1}{1+s} d\hat{\mu}_{\max}(s) \le \frac{1}{2} \qquad \square$$

4. EXAMPLE 1: THE EXTREME BOUNDARY OF $\mathcal{G}$

The extreme boundary of $\mathcal{G}$ consists of the functions

$$g_s(w) = \frac{1+s}{2}\frac{(w-1)^2}{w+s} \quad \text{for } s \ge 0$$
$$g_\infty(w) = \frac{1}{2}(w-1)^2$$

We have $\hat{g}_s = g_{s^{-1}}$ for $s > 0$ and $\hat{g}_0 = g_\infty$, g_1 being the only symmetric one of these functions. The corresponding measures are $\mu_s(t) = \frac{1}{2}\delta(s-t)$.

Let $s \in [0, 1]$. Denote $h_s = g_s + \hat{g}_s$, then

$$h_s(w) = \frac{(1+s)^2}{2}(w-1)^2 \frac{w+1}{(w+s)(sw+1)}$$

Let λ_s be the corresponding monotone metric. It is easy to see that $g_{s,max} = \hat{g}_s$ and that

$$\mathcal{G}_s := \mathcal{G}_{h_s} = \left\{ g_\alpha = \frac{1-\alpha}{2} g_s + \frac{1+\alpha}{2} g_{s^{-1}} : \alpha \in [-1, 1] \right\}$$

In particular, $\mathcal{G}_1 = \{g_1\}$. It follows that for each λ_s, we have a unique family of α-connections. If we consider commutative submanifolds, we obtain classical α^*-connections with $\alpha^* \in [-3\frac{1-s}{1+s}, 3\frac{1-s}{1+s}]$. Two important special cases, $s = 1$ and $s = 0$ will be treated below.

4.1. The Metric of Bures

Let us consider the previous example with $s = 1$. Then

$$h_1(w) = 2\frac{(w-1)^2}{w+1}$$

and the corresponding monotone metric is given by

$$\lambda_{1\rho}(X, Y) = \operatorname{Tr} X \frac{2}{L_\rho + R_\rho}(Y)$$

It is the smallest metric in the class of monotone metrics. We have already seen that the corresponding equivalence class consists of only one function g_1. It means that the only connection that we can obtain is the metric connection $\bar{\nabla}$.

4.2. The Largest Monotone Metric

Let $s = 0$. Then

$$h_0(w) = \frac{1}{2}(w-1)^2\frac{w+1}{w}$$

and λ_0 is given by

$$\lambda_{0\rho}(X, Y) = \operatorname{Tr} X\frac{1}{2}\left(R_\rho^{-1} + L_\rho^{-1}\right)(Y)$$

It is the largest monotone metric. On commutative submanifolds, we obtain α^* -connections for α^* in the largest possible interval $\alpha^* \in [-3, 3]$. It is easy to see from Proposition 3.3 that this is the only monotone metric with this property.

5. STATISTICAL MANIFOLDS

The manifold $\mathcal{D}^+$ with a monotone metric and a class of α-connections can be regarded as a statistical manifold in the sense of Lauritzen (Amari *et al.*, 1987). A statistical manifold is a triple $(M, g, \tilde{D})$, where M is a differentiable manifold, g a metric tensor and $\tilde{D}$ a symmetric covariant 3-tensor, called the skewness of the manifold. On M, a class of α-connections is introduced by

$$\nabla_X^\alpha Y = \bar{\nabla}_X Y - \frac{\alpha}{2}D(X, Y), \tag{7}$$

where $\bar{\nabla}$ is the metric connection and the tensor D is defined by $\tilde{D}(X, Y, Z) = g(D(X, Y), Z)$. These connections are torsion free, this is equivalent to symmetry of $\tilde{D}$, resp. D. The Riemannian curvature tensor is defined as

$$R^\alpha(X, Y, Z, W) = g\left(\nabla_X^\alpha\nabla_Y^\alpha Z - \nabla_Y^\alpha\nabla_X^\alpha Z - \nabla_{[X,Y]}^\alpha Z, W\right)$$

Statistical manifolds satisfying $R^\alpha = R^{-\alpha}$ for all α are called conjugate symmetric. It is proved that $R^{-\alpha} - R^\alpha = \alpha\{F(X, Y, Z, W) - F(Y, X, Z, W)\}$, where $F(X, Y, Z, W) = (\bar{\nabla}_X\tilde{D})(Y, Z, W)$, so that a statistical manifold is conjugate symmetric if and only if the tensor F is symmetric. It also follows that the condition

$$\exists\alpha_0 \neq 0, \quad R^{\alpha_0} = R^{-\alpha_0}$$

is sufficient for conjugate symmetry.

Let now λ be a monotone metric on $\mathcal{D}^+$ and let $\mathcal{G}_h$ be the corresponding equivalence class. Let $g \in \mathcal{G}_h$ such that g is not symmetric and let us consider the corresponding family of connections. Let

$$D(X, Y) = \nabla_X Y - \nabla_X^* Y$$

Then the triple $(\mathcal{D}^+, \lambda, \tilde{D})$ is a statistical manifold, with $\tilde{D}(X, Y, Z) = \lambda(D(X, Y), Z)$, and the family of connections has the form (7).

Let K be a covariant k-tensor field, then its symmetrization K^{sym} is defined as

$$K^{\text{sym}}(X_1, \ldots, X_k) = \frac{1}{k!} \sum_{\pi} K\left(X_{\pi(1)}, \ldots, X_{\pi(k)}\right)$$

where π runs over all permutations of the set $\{1, \ldots, k\}$.

Proposition 5.1. *Let T_s^{sym} be the symmetrization of $\Re T_s$. Then $\tilde{D}$ has the form*

$$\tilde{D}(X, Y, Z) = 6 \int_{[0,\infty]} T_s^{\text{sym}}(X, Y, Z)\, d(\mu - \hat{\mu})(s)$$

Proof: Straightforward from Proposition 3.2. □

Let us now compute the Riemannian curvature tensor R^α of the α-connection.

Proposition 5.2. *Let X, Y, Z, W be vector fields on $\mathcal{M}^+$ and let $\bar{R} = R^0$. Then*

$$R^\alpha(X, Y, Z, W) = \bar{R}(X, Y, Z, W) + \frac{\alpha}{2}\{F(Y, X, Z, W) - F(X, Y, Z, W)\}$$
$$+ \frac{\alpha^2}{4}\{\lambda(D(X, W), D(Y, Z)) - \lambda(D(X, Z), D(Y, W))\}$$

Proof: As we are going to establish a tensorial equality, we may suppose that $[X, Y] = 0$. As the metric connection is symmetric, we have $\bar{\nabla}_X Y - \bar{\nabla}_Y X = 0$. From (7) we obtain, using symmetry of $\tilde{D}$

$$\lambda\left(\nabla_X^\alpha \nabla_Y^\alpha Z, W\right) = \lambda(\bar{\nabla}_X \bar{\nabla}_Y Z, W) - \frac{\alpha}{2}\{\lambda(\bar{\nabla}_X D(Y, Z), W) + \lambda(D(X, \bar{\nabla}_Y Z), W)\}$$
$$+ \frac{\alpha^2}{4}\lambda(D(X, W), D(Y, Z))$$

Subtracting the expression with interchanged X and Y and using self-duality and symmetry of $\bar{\nabla}$ completes the proof. □

Corollary 5.1. *Let the manifold be conjugate symmetric. Then we have*

$$R^\alpha(X, Y, Z, W) = \bar{R}(X, Y, Z, W)$$
$$+ \frac{\alpha^2}{4}\{\lambda(D(X, W), D(Y, Z)) - \lambda(D(X, Z), D(Y, W))\}$$

If $\theta \mapsto \rho(\theta)$ is a smooth parametrization of $\mathcal{D}^+$, then

$$R_{ijkl}^\alpha(\theta) = \bar{R}_{ijkl}(\theta) + \frac{\alpha^2}{4} \sum_{\beta,\gamma} (\tilde{D}_{il\beta}\tilde{D}_{jk\gamma} - \tilde{D}_{ik\beta}\tilde{D}_{jl\gamma})\lambda^{\beta\gamma}$$

where $\lambda^{ij} = (\lambda^{-1})_{ij}$.

Corollary 5.2. *If* $\exists\alpha_0 \neq 0$*, such that* $R^{\alpha_0} = 0$*, then*

$$R^{\alpha}(X, Y, Z, W) = \frac{\alpha^2 - \alpha_0^2}{4}\{\lambda(D(X, W), D(Y, Z)) - \lambda(D(X, Z), D(Y, W))\}$$

for $\forall\alpha$*. Moreover, there exists a parametrization,* $\theta \mapsto \rho(\theta)$*, such that*

$$R^{\alpha}_{ijkl} = \frac{\alpha^2 - \alpha_0^2}{4\alpha_0^2} \sum_{\beta,\gamma} (\Gamma_{il\beta}\Gamma_{jk\gamma} - \Gamma_{ik\beta}\Gamma_{jl\gamma})\lambda^{\beta\gamma}$$

where $\Gamma_{ijk} = \lambda(\nabla^{-\alpha_0}_{\partial_i}\partial_j, \partial_k)$ *are the Christoffel symbols of* $\nabla^{-\alpha_0}$.

Proof: The connections ∇^{α} and $\nabla^{-\alpha}$ are mutually dual, therefore $0 = R^{\alpha_0} = R^{-\alpha_0}$. It follows that the manifold is conjugate symmetric and we may use Corollary 5.1.

Further, let us define

$$D_{\alpha_0}(X, Y) = \nabla^{-\alpha_0} - \nabla^{\alpha_0},$$

then $D_{\alpha_0} = \alpha_0 D$ and

$$\nabla^{\alpha} = \bar{\nabla} - \frac{\alpha}{2\alpha_0} D_{\alpha_0}$$

It follows that

$$R^{\alpha}(X, Y, Z, W) = \frac{\alpha^2 - \alpha_0^2}{4\alpha_0^2}\{\lambda(D_{\alpha_0}(X, W), D_{\alpha_0}(Y, Z)) - \lambda(D_{\alpha_0}(X, Z), D_{\alpha_0}(Y, W))\}$$

As the manifold is $\pm\alpha_0$-flat, there exists an α_0-affine parametrization $\theta \mapsto \rho(\theta)$, i.e. such that $\nabla^{\alpha_0}_{\partial_i}\partial_j = 0$ for all i, j. It follows that

$$\tilde{D}^{\alpha_0}_{ijk} = \lambda(D_{\alpha_0}(\partial_i, \partial_j), \partial_k) = \Gamma_{ijk}, \quad \forall i, j, k$$ □

Corollary 5.3. *If* $\exists\alpha_1 \neq \pm\alpha_2$ *such that* $R^{\alpha_1} = R^{\alpha_2} = 0$*, then* $R^{\alpha} = 0$ *for all* α.

Proof: We may suppose that $\alpha_1 \neq 0$ and use Corollary 5.2. □

6. EXAMPLE 2: α-DIVERGENCES

Let

$$g_{\alpha} = \begin{cases} \dfrac{4}{1 - \alpha^2}\left(\dfrac{1 + w}{2} - w^{\frac{1+\alpha}{2}}\right) & \alpha \neq \pm 1 \\ -\log w & \alpha = -1 \\ w \log w & \alpha = 1 \end{cases}$$

Then $g_\alpha \in \mathcal{G}$ for $\alpha \in [-3, 3]$. Moreover, $\hat{g}_\alpha = g_{-\alpha}$. The corresponding family of relative entropies and monotone metrics was defined by Hasegawa (1993). We have

$$\lambda_\alpha(X, Y) = \frac{\partial^2}{\partial s \partial t} \mathrm{Tr}\, f_\alpha(\rho + sX) f_{-\alpha}(\rho + tY)|_{t=s=0}$$

where f_α is the family of functions defined in Section 1. It is easy to show that the corresponding affine connections ∇^{g_α} coincide with the α-connections for λ_α defined in Jenčová (2001a). As the connections are torsion-free, this is the only case when this may happen, see also Jenčová (2001b).

There are some important special cases. For $\alpha = \pm 1$ we obtain the well known Bogoljubov-Kubo-Mori (BKM) metric. Another important example is $\alpha = \pm 3$, corresponding to the largest monotone metric, see Example 4.2. This is the unique monotone metric that is contained in both classes λ_α and λ_s from Section 4.

Let us fix $\alpha_0 \in (0, 3]$. Then

$$h_{\alpha_0}(w) = g_{\alpha_0}(w) + g_{-\alpha_0}(w) = \frac{4}{1 - \alpha_0^2} \left(1 - w^{\frac{1-\alpha_0}{2}}\right)\left(1 - w^{\frac{1+\alpha_0}{2}}\right)$$

If we proceed as in the proof of Corrolary 5.2, we see that the family of connections

$$\nabla^\alpha = \bar{\nabla} - \frac{\alpha}{2\alpha_0} D_{\alpha_0},$$

can be obtained from $\mathcal{G}_{\alpha_0} = \mathcal{G}_{h_{\alpha_0}}$ for $\alpha \in [-\alpha_0, \alpha_0]$. In particular, $\nabla^{\alpha_0} = \nabla^{g_{\alpha_0}}$. As it was shown in Jenčová (2001a), the connection $\nabla^{\pm\alpha_0}$ is flat, i.e. the Riemannian curvature tensor $R^{\pm\alpha_0}$ vanishes. Hence, for the -α_0-affine parametrization θ,

$$\begin{aligned} R^\alpha_{ijkl} &= \frac{\alpha^2 - \alpha_0^2}{4\alpha_0^2} \left\{ \lambda_{\alpha_0}\left(\nabla^{\alpha_0}_{\partial_i} \partial_l, \nabla^{\alpha_0}_{\partial_j} \partial_k\right) - \lambda_{\alpha_0}\left(\nabla^{\alpha_0}_{\partial_i} \partial_k, \nabla^{\alpha_0}_{\partial_j} \partial_l\right)\right\} \\ &= \frac{\alpha^2 - \alpha_0^2}{4\alpha_0^2} \sum_{\beta,\gamma} \left(\Gamma^{\alpha_0}_{il\beta} \Gamma^{\alpha_0}_{jk\gamma} - \Gamma^{\alpha_0}_{ik\beta} \Gamma^{\alpha_0}_{jl\gamma}\right) \lambda^{\beta\gamma} \end{aligned}$$

where

$$\Gamma^{\alpha_0}_{ijk} = \mathrm{Tr}\, \partial_i \partial_j f_{\alpha_0}(\rho) \partial_k f_{-\alpha_0}(\rho)$$

In particular, for $\alpha_0 = 1$ (the BKM metric), $\nabla^{g_{-1}}$ and ∇^{g_1} correspond to the mixture and exponential connections $\nabla^{(m)}$ and $\nabla^{(e)}$, respectively. The α-connection is then a convex mixture of the (m) and (e)-connections,

$$\nabla^\alpha = \frac{1 - \alpha}{2} \nabla^{(m)} + \frac{1 + \alpha}{2} \nabla^{(e)}$$

In the commutative case, this is an equivalent definition of the α-connections. If we consider the natural affine parametrization $\rho(\theta) = \rho_0 + \sum_i \theta_i X_i$, the coefficients of Riemannian curvature tensor can be written in the form

$$R^{\alpha}_{ijkl} = \frac{\alpha^2 - 1}{4} \mathrm{Tr} \int_0^1 \{\partial_i \partial_j \log(\rho)\rho^t \partial_j \partial_k \log(\rho)\rho^{1-t} - \partial_i \partial_k \log(\rho)\rho^t \; \partial_j \partial_l \log(\rho)\rho^{1-t}\} \, dt$$

If $\{X_i\}$ is an orthonormal basis of $\mathcal{D}$ with respect to the metric $\lambda^{BKM}_{\rho_0}$, we may compute the coefficients at $\theta = 0$ as

$$R^{\alpha}_{ijkl} = \frac{\alpha^2 - 1}{4} \sum_{\beta} (\Gamma_{il\beta}\Gamma_{jk\beta} - \Gamma_{ik\beta}\Gamma_{jl\beta})$$

where

$$\Gamma_{ijk} = \mathrm{Tr}\, \partial_i \partial_j \log(\rho) X_k = -\mathrm{Tr}\, X_k \int_0^{\infty} [(\rho_0 + s)^{-1} X_i (\rho_0 + s)^{-1} X_j (\rho_0 + s)^{-1} + (\rho_0 + s)^{-1} X_j (\rho_0 + s)^{-1} X_i (\rho_0 + s)^{-1}] \, ds$$

As it was already proved e.g. in Petz (1994), the Riemannian curvature $\bar{R}$ of the metric connection given by λ^{BKM} is not equal to 0. Using Corrolary 5.3, it follows that $R^{\alpha} = 0$ if and only if $\alpha = \pm 1$.

ACKNOWLEDGMENT

Research supported by the grant VEGA 1/0264/03.

REFERENCES

Amari, S. (1985). Differential-geometrical methods in statistic, *Lecture Notes in Statistics*, Vol. 28.

Amari, S., Barndorff-Nielsen, O. E., Kass, R. E., Lauritzen, S. L., and Rao, C. R. (1987). Differential geometry in statistical inference, *IMS Lecture notes-Monograph series*, Vol. 10.

Chentsov, N. N. (1982). Statistical decision rules and optimal inferences, *Translation of Mathematical Monograph 53*, American Mathematical Society, Providence.

Chentsov, N. N. and Morozova, E. A. (1990). Markov invariant geometry on state manifolds, *Itogi Nauki i Tekhniki* **36**, 69–102, (in Russian).

Eguchi, S. (1983). Second order efficiency of minimum contrast estimation in a curved exponential family, *Annals of Statistics* **11**, 793–803.

Hasegawa, H. (1993). α-divergence of the non-commutative information geometry, *Reports on Mathematical Physics* **33**, 87–93.

Jenčová, A. (2001). Geometry of quantum states: dual Connections and divergence functions. *Reports on Mathematical Physics* **47**, 121–138.

Jenčová, A. (2001). Dualistic properties of the manifold of quantum states. In *Disordered and Complex Systems*, P. Sollich *et al.*, eds., AIP, Melville, New York.

Lesniewski, L. and Ruskai, M. B. (1999). Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, *Journal of Mathematical Physics* **40**, 5702–5724.

Nagaoka, H. (1994). Differential geometrical aspects of quantum state estimation and relative entropy. In *Quantum Communication, Computing and Measurement*, Hirota *et al.*, eds., Plenum Press, New York.

Petz, D. (1986). Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics* **23**, 57–65.

Petz, D. (1994). Geometry of canonical correlation on the state space of a quantum system. *Journal of Mathematical Physics* **244**, 780–795.

Petz, D. (1996). Monotone metrics on matrix spaces, *Linear Algebra and its Applications* **244**, 81–96.