FLAT CONNECTIONS AND WIGNER–YANASE–DYSON METRICS

Anna Jenčová

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

On the manifold of positive definite matrices, we investigate the existence of pairs of flat affine connections, dual with respect to a given monotone metric. The connections are defined either using the α-embeddings and finding the duals with respect to the metric, or by means of contrast functionals. We show that in both cases, the existence of such a pair of connections is possible if and only if the metric is given by the Wigner–Yanase–Dyson skew information.

Keywords: monotone metrics, flat affine connections, duality, generalized relative entropies, WYD metrics.

1. Introduction

An important feature of the classical information geometry is the uniqueness of its structures, the Fisher metric and the family of affine α-connections on a manifold $\mathcal{P}$ of probability distributions [5, 1]. In the case of finite quantum systems, this uniqueness does not take place: it was shown by Chentsov and Morozova [6] and later by Petz [22] that there are infinitely many Riemannian metrics, which are monotone with respect to stochastic maps. As for the affine connections, there were several definitions of the α-connections [16, 19, 12, 14].

In the commutative case, two equivalent definitions of the connections were used by Amari [1]. First, the connections can be defined using α-embeddings (α-representations) given by the family of functions

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

On the other hand, the connections can be defined as mixtures of the exponential and the mixture connections,

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

Such connections are torsion-free and the α and $-\alpha$ connections are dual with respect to the Fisher metric. Moreover, in the case of a finite system, that is on the manifold of all (non-normalized) multinomial distributions, the α-connections are flat for all α.

The definition involving α-representations can be easily generalized to noncommutative case to obtain a family of flat connections $\nabla^{(\alpha)}$ on the manifold of positive definite matrices. This definition was treated also by the present author in [16] and [17]. The dual of such α-connection with respect to a given monotone metric is in general different from the $-\alpha$-connection. The duals have vanishing Riemannian curvature, but are not always torsion-free and hence not flat. The condition that the dual of the α-connection with respect to a monotone metric is torsion-free restricts α to the interval $[-3, 3]$ and, for such α, singles out a monotone metric λ_α, which belongs to the family of Wigner–Yanase–Dyson (WYD) metrics. This is also equivalent to the condition that the dual of $\nabla^{(\alpha)}$ is $\nabla^{(-\alpha)}$, see also [10]. A brief description of these results is given in Sections 2 and 3.

For $\alpha = \pm 1$, we get the Kubo–Mori–Bogoljubov metric, with respect to which the mixture $\nabla^{(m)}$ and exponential $\nabla^{(e)}$ connections are dual. As in the classical case, we may use mixtures of $\nabla^{(e)}$ and $\nabla^{(m)}$ to define a family of torsion-free connections, having the required duality properties with respect to the BKM metric. In our approach, however, the value of α in (2) will be restricted to the interval $[-1, 1]$, but the proofs in Section 5 suggest that our results hold more generally. Convex mixtures were considered also by Grasselli and Streater, see the Discussion in [11]. We will answer the questions discussed there in proving that, for $\alpha \in (-1, 1)$, affine connections defined by (2) are different from the α-connections and are not flat. A simple direct proof of this fact can be found at the end of Section 5.

Another way to define an affine connection was proposed by Eguchi in [9], by means of a contrast functional on $\mathcal{P}$. A functional $\phi : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ is a contrast functional if it satisfies $\phi(p, q) \geq 0$ for all p, q and $\phi(p, q) = 0$ if and only if $p = q$. Using such a functional, a metric tensor and affine connection can be defined. Let $\theta_1, \ldots, \theta_p$ be a smooth parametrization of $\mathcal{P}$ and let ∂_i, $i = 1, \ldots, p$, be the corresponding vector fields, then the metric tensor is given by

$$g_{ij}^\phi = -\partial_i \partial_j' \phi(p(\theta), p(\theta'))|_{\theta=\theta'}.$$

An affine connection ∇^ϕ is defined by

$$\Gamma_{ijk}^\phi = g^\phi(\nabla_{\partial_i}^\phi \partial_j, \partial_k) = -\partial_i \partial_j \partial_k' \phi(p(\theta), p(\theta'))|_{\theta=\theta'}.$$

Consider a special class of contrast functionals ϕ_g, related to convex functions g satisfying $g(1) = 0$ by

$$\phi_g(p, q) = \int g\left(\frac{q}{p}\right) dp.$$

In this case, it was shown [1] that the corresponding metric is the Fisher metric (multiplied by $g''(1)$) and the affine connection is the α-connection, $\alpha = 2g'''(1) + 3$.

As the quantum counterpart of such functionals we will use the relative g-entropies H_g defined by Petz [20],

$$H_g(\rho, \sigma) = \text{Tr} \, \rho^{1/2} g(L_\sigma/R_\rho)(\rho^{1/2}),$$

where g is an operator convex function and $g(1) = 0$. It was shown that [18]:

(a) In the normalized case (or if $\text{Tr} \, \rho = \text{Tr} \, \sigma$), $H_g(\rho, \sigma) \geq 0$ and $H_g(\rho, \sigma) = 0$ if and only if $\rho = \sigma$.

(b) $H_g(\lambda\rho, \lambda\sigma) = \lambda H_g(\rho, \sigma)$ for each $\lambda > 0$.

(c) H_g is jointly convex in ρ and σ.

(d) H_g is monotone, that is, it decreases under stochastic maps.

(e) H_g is differentiable.

We see that H_g is a contrast functional on the manifold of quantum states, and we will show that we can use it to define the geometrical structures as above, even in the non-normalized case. The relative g-entropies were used by Lesniewski and Ruskai [18], who proved that the Riemannian structure given by H_g is monotone for each g and, conversely, each monotone metric is obtained in this way. A short account on some of their results is in Section 4.

In Section 5 we will use H_g to define an affine connection and show that this definition contains both the α-connections, defined from α-embeddings, and the convex mixtures of $\nabla^{(m)}$ and $\nabla^{(e)}$. We will show that for each monotone metric there is a family of such connections (the p-connections) parametrized by $p \in [0, 1]$, such that they are torsion-free and the p- and $(1 - p)$-connections are dual. We will then use the theory of statistical manifolds by Lauritzen [2] to investigate the Riemannian curvature of the connections.

Finally, in the last section we will show that a pair of dual flat connections exists if and only if the metric is one of the WYD metrics λ_α. The flat connections are then the $\pm\alpha$-connections. This result holds for the connections given by the relative g-entropies. It is known from [1] that dual flat connections give rise to divergence functionals on the manifold, it is therefore reasonable to consider connections defined from functionals having the properties (a)–(e). The class of g-entropies seems to be large enough, although it does not contain all such functionals (see [18]). The main results of the present paper can be summarized as follows: If a pair of dual flat connections is required, the structures of information geometry are unique even in the quantum case, at least if we consider only connections defined by the relative g-entropies. These structures are provided by the family of Wigner–Yanase–Dyson metrics and the α-connections.

2. The manifold and monotone metrics

Let $\mathcal{M}_n(\mathbb{C})$ be the space of $n \times n$ complex matrices, $\mathcal{M}_h$ be the real linear subspace of hermitian matrices and let $\mathcal{M} \subset \mathcal{M}_h$ denote the set of positive definite matrices. As an open subset in a finite-dimensional real vector space, $\mathcal{M}$ inherits the structure of a differentiable manifold. The tangent space T_ρ of $\mathcal{M}$ at ρ is the

linear space of directional (Fréchet) derivatives in the direction of smooth curves in $\mathcal{M}$ and it can be identified with $\mathcal{M}_h$ in an obvious way. In the present paper, the elements of the tangent space, seen as directional derivative operators, will be denoted by $\mathcal{X}, \mathcal{Y}$, etc., while the corresponding capital letters will mean their representations $X = \mathcal{X}(\rho)$ etc. in $\mathcal{M}_h$. The map $\mathcal{X} \mapsto X$ is the same as Amari's -1-representation of the tangent space in the classical case [1], see also the next section. The vector fields on $\mathcal{M}$ are represented by $\mathcal{M}_h$-valued functions on $\mathcal{M}$. If $\mathcal{X}, \mathcal{Y}$ are vector fields, then the bracket $[\mathcal{X}, \mathcal{Y}]$ is unrelated to the usual commutator of the representing matrices and these two should not be confused. In the present paper we will use $[\cdot, \cdot]$ only in the first (vector fields) meaning.

A Riemannian structure is introduced in $\mathcal{M}$ by

$$\lambda_\rho(X, Y) = \operatorname{Tr} X J_\rho(Y), \qquad X, Y \in T_\rho,$$

where J_ρ is a suitable operator on matrices. We say that the metric λ is monotone if it is monotone with respect to stochastic maps, that is, we have

$$\lambda_{T(\rho)}(T(X), T(X)) \leq \lambda_\rho(X, X), \qquad \rho \in \mathcal{M}, \ X \in T_\rho,$$

for a stochastic map T. It is an important result of Petz [22], that this is equivalent to

$$J_\rho = R_\rho^{-1/2} F(L_\rho / R_\rho)^{-1} R_\rho^{-1/2},$$

where $F : \mathbb{R}^+ \to \mathbb{R}$ is an operator monotone function, which is symmetric, $F(x) = x F(x^{-1})$, and normalized, $F(1) = 1$. The operators L_ρ and R_ρ are the left and right multiplication operators. Clearly, $J_\rho(X) = \rho^{-1} X$ if X and ρ commute, so that the restriction of λ to commutative submanifolds is the Fisher metric.

EXAMPLE 2.1. Let J_ρ be the symmetric logarithmic derivative, given by $J_\rho(X) = Y$, $Y\rho + \rho Y = 2X$, then the metric λ is monotone, with $F(x) = \frac{1+x}{2}$. This metric is sometimes called the Bures metric and it is the smallest monotone Riemannian metric.

EXAMPLE 2.2. The largest monotone metric is given by the operator monotone function $F(x) = \frac{2x}{1+x}$. In this case $J_\rho(X) = \frac{1}{2}(\rho^{-1} X + X\rho^{-1})$ is the right logarithmic derivative (RLD).

EXAMPLE 2.3. An important example of a monotone metric is the Kubo–Mori–Bogoljubov (BKM) metric, obtained from

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

In this case $F(x) = \frac{x-1}{\log(x)}$.

3.   The α-representation and α-connections

Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function and let $\rho \in \mathcal{M}$. Let us define the operator $L_f[\rho] : \mathcal{M}_h \to \mathcal{M}_h$ by

$$L_f[\rho](X) = \frac{d}{ds} f(\rho + sX)|_{s=0}.$$

This operator has the following properties [16]:

(i) The chain rule: $L_{f \circ g}[\rho] = L_f[g(\rho)]L_g[\rho]$. In particular, if f is invertible then $L_f[\rho]$ is invertible and $L_f[\rho]^{-1} = L_{f^{-1}}[f(\rho)]$.

(ii) $L_f[\rho]$ is a self-adjoint operator in $\mathcal{M}_h$, with respect to the Hilbert–Schmidt inner product $\langle X, Y \rangle = \mathrm{Tr}\, X^*Y$.

(iii) If $X\rho = \rho X$, then $L_f[\rho](X) = f'(\rho)X$, $f'(x) = \frac{d}{dx}f(x)$.

Let now f_α be given by (1). The map

$$\ell_\alpha : \mathcal{M} \ni \rho \mapsto f_\alpha(\rho) \in \mathcal{M}_h$$

will be called the α-embedding of $\mathcal{M}$. The α-embedding induces the map

$$T_\rho \ni X \mapsto \mathcal{X}(f_\alpha(\rho)) = L_\alpha[\rho](X) \in \mathcal{M}_h,$$

where $L_\alpha[\rho] := L_{f_\alpha}[\rho]$, it will be called the α-representation of the tangent vector X. We will often omit the indication of the point in the square brackets, if no confusion is possible.

Let λ be a monotone metric and let $Y_1 = L_\alpha(X_1)$ and $Y_2 = L_\alpha(X_2)$ be the α-representations of the tangent vectors X_1 and X_2, then

$$\lambda_\rho(X_1, X_2) = \mathrm{Tr}\, Y_1 K_\alpha(Y_2), \qquad (3)$$

where $K_\alpha = L_\alpha^{-1} J_\rho L_\alpha^{-1}$.

EXAMPLE 3.1. The family of Wigner–Yanase–Dyson (WYD) metrics λ_α is defined by $J_\rho = L_{-\alpha}L_\alpha$. In [15], it was shown that such metrics are monotone for $\alpha \in [-3, 3]$ and that there are no other monotone metrics, satisfying

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

for some functions f and f^*. The corresponding operator monotone function is

$$F_\alpha(x) = \frac{1 - \alpha^2}{4} \frac{(x - 1)^2}{(x^{\frac{1+\alpha}{2}} - 1)(x^{\frac{1-\alpha}{2}} - 1)}.$$

As special cases we obtain the BKM metric for $\alpha = \pm 1$ and RLD metric for $\alpha = \pm 3$. The smallest metric in this class is the Wigner-Yanase (WY) metric,

corresponding to $\alpha = 0$, here $F_0(x) = \frac{1}{4}(1 + \sqrt{(x)})^2$, the Bures metric is not included. For the metric λ_α, $\alpha \in [-3, 3]$, we have $K_\alpha = L_{-\alpha} L_\alpha^{-1}$. It can be shown that $K_\alpha^{-1} = K_{-\alpha}$ if and only if $\lambda = \lambda_\alpha$.

The connection $\nabla^{(\alpha)}$ is defined by

$$L_\alpha((\nabla_{\mathcal{X}}^{(\alpha)} \mathcal{Y})(\rho)) = \mathcal{X} \mathcal{Y} f_\alpha(\rho)$$

for smooth vector fields $\mathcal{X}, \mathcal{Y}$. Clearly, a vector field is parallel with respect to this connection if and only if its α-representation is a constant hermitian matrix-valued function on $\mathcal{M}$. For $\alpha = -1$ and $\alpha = 1$, we get the mixture and exponential connections, sometimes denoted by $\nabla^{(m)}$ and $\nabla^{(e)}$. The mixture connection coincides with the natural flat affine structure inherited from $\mathcal{M}_h$.

For each α there is a coordinate system $\xi_1, \ldots, \xi_N$, such that $f_\alpha(\rho(\xi)) = \sum_i \xi_i Z_i$, where $Z_i \in \mathcal{M}_h$, $i = 1, \ldots, N$, form a basis of $\mathcal{M}_h$. Clearly, such coordinate system is $\nabla^{(\alpha)}$-affine. The existence of an affine coordinate system is equivalent to flatness of the connection $\nabla^{(\alpha)}$, that is, the connections are torsion-free and the Riemannian curvature tensor vanishes. Thus we have a one-parameter family of flat α-connections, just as in the classical case. But, contrary to the classical case, the $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$ are not dual for a general monotone metric.

Let us define the connection $\nabla^{(\alpha)*}$ by

$$L_\alpha^{-1} J_\rho((\nabla_{\mathcal{X}}^{(\alpha)*} \mathcal{Y})(\rho)) = \mathcal{X} L_\alpha^{-1} J_\rho(Y) = \mathcal{X} K_\alpha L_\alpha(Y).$$

It can be easily seen from (3) that the connections $\nabla^{(\alpha)}$ and $\nabla^{(\alpha)*}$ are dual with respect to λ. It follows that $\nabla^{(\alpha)*}$ is also curvature free and it is torsion-free if and only if [16]

$$\mathcal{X} L_\alpha^{-1} J_\rho(Y) = \mathcal{Y} L_\alpha^{-1} J_\rho(X) \tag{4}$$

for all vector fields satisfying $[\mathcal{X}, \mathcal{Y}] = 0$.

THEOREM 3.1 ([17]). *Let $\alpha \in [-3, 3]$. The following conditions are equivalent:*

(i) $(\nabla^{(\alpha)})^*$ *is torsion-free.*

(ii) $J_\rho = L_\alpha L_{-\alpha}$.

(iii) $(\nabla^{(\alpha)})^* = \nabla^{(-\alpha)}$.

Proof: (i)$\Rightarrow$(ii): Let $\theta \mapsto \rho(\theta)$ be a smooth parametrization of $\mathcal{M}$ and let $\partial_i = \frac{\partial}{\partial \theta_i}$, $i = 1, \ldots, N$. Let us denote $X_i(\theta) = \partial_i(\rho(\theta))$. Let $\nabla^{(\alpha)*}$ be torsion-free and let $F_i(\theta) = L_\alpha^{-1} J_{\rho(\theta)}(X_i(\theta))$, $i = 1, \ldots, N$. Then we get from (4) that $\partial_j F_i = \partial_i F_j$ for all i, j.

Let $A_1, \ldots, A_N$ be a basis of $\mathcal{M}_h$ and let $F_i(\theta) = \sum_k f_{ik}(\theta) A_k$, then $\partial_i f_{jk}(\theta) = \partial_j f_{ik}(\theta)$ for all k, i and j. This implies the existence of functions $\phi_1, \ldots, \phi_N$, such that $f_{ik}(\theta) = \partial_i \phi_k(\theta)$. Let $\phi(\theta) = \sum_k \phi_k(\theta) A_k$, then $F_i = \partial_i \phi$. Moreover, if $\rho_t = \rho(\theta(t))$ is a curve in $\mathcal{M}$, then

$$\frac{d}{dt} \phi(\theta(t)) = \sum_i \frac{d}{dt} \theta_i(t) F_i(\theta(t)) = L_\alpha^{-1}[\rho_t] J_{\rho_t}\left(\frac{d}{dt} \rho_t\right).$$

Let now $\rho \in \mathcal{M}$ and let us consider the curve $\rho_t = \rho(\theta(t)) = t\rho + (1 - t)$. Using the fact that $\frac{d}{dt}\rho_t = \rho - 1$ and ρ_t commute for all t, we have

$$\phi(\theta(1)) - \phi(\theta(0)) = \int_0^1 \frac{d}{dt}\phi(\theta(t))dt = \int_0^1 L_\alpha^{-1}[\rho_t]J_{\rho_t}(\rho - 1)dt$$

$$= \int_0^1 (1 + t(\rho - 1))^{\frac{\alpha-1}{2}}(\rho - 1)dt = f_{-\alpha}(\rho) - f_{-\alpha}(I).$$

Therefore, $\phi(\theta) = f_{-\alpha}(\rho(\theta)) + c$. It follows that

$$L_\alpha^{-1} J_{\rho(\theta)}(X_i(\theta)) = F_i(\theta) = \partial_i f_{-\alpha}(\rho(\theta)) = L_{-\alpha}(X_i(\theta))$$

and $J_\rho = L_\alpha L_{-\alpha}$.

(ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (i) are quite clear.   □

The statement for $\alpha = \pm 1$ was already proved in [3]. The equivalence (ii) $\Longleftrightarrow$ (iii) was proved (by a different method) in [11] for $\alpha = \pm 1$ and in [10] for $\alpha \in (-1, 1)$.

REMARK 3.1. Let $\mathcal{D} = \{\rho \in \mathcal{M} : \text{Tr}\, \rho = 1\}$ be the submanifold of quantum states. The connections induced on $\mathcal{D}$ are orthogonal projections of the above connections. The Riemannian curvature is given by [16]

$$R_\rho^\alpha(X, Y, Z, W) = \frac{1 - \alpha^2}{4}\{\text{Tr}\, Y J_\alpha(Z)\text{Tr}\, X J_\rho(W) - \text{Tr}\, X J_\alpha(Z)\text{Tr}\, Y J_\rho(W)\},$$

where $\rho \in \mathcal{D}$, $X, Y, Z, W \in T_\rho(\mathcal{D})$, and thus $R^\alpha = 0$ if and only if $\alpha = \pm 1$. Therefore, the α-connections are not flat on $\mathcal{D}$, unless $\alpha = \pm 1$, which corresponds to the classical results.

4. Relative g-entropies and monotone metrics

Let G be the set of all operator convex functions $(0, \infty) \to \mathbb{R}$, satisfying $g(1) = 0$ and $g''(1) = 1$. For $g \in G$, we define the relative g-entropy $H_g : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ by [20]

$$H_g(\rho, \sigma) = \text{Tr}\, \rho^{1/2}g(L_\sigma/R_\rho)(\rho^{1/2}).$$

The set G is the set of functions of the form

$$g(u) = a(u - 1) + \int_{[0,\infty]} (u - 1)^2\frac{1 + s}{u + s}d\mu(s), \tag{5}$$

where μ is a positive finite measure on $[0, \infty]$ satisfying $\int_{[0,\infty]} d\mu(s) = 1/2$ and $a = g'(1)$ is a real number. We will denote by $b = \mu(\{\infty\})$ and $c = \mu(\{0\})$ the possible atoms in 0 and ∞, then

$$g(u) = a(u - 1) + b(u - 1)^2 + c\frac{(u - 1)^2}{u} + \int_0^\infty (u - 1)^2\frac{1 + s}{u + s}d\mu(s).$$

For an operator convex function g we define its transpose $\hat{g}(u) = ug(u^{-1})$. Clearly, $g \in G$ implies $\hat{g} \in G$, with the positive measure $\hat{\mu}$ satisfying $d\hat{\mu}(s) = d\mu(s^{-1})$ and $\hat{a} = -a$. We say that g is symmetric if $g = \hat{g}$. For each symmetric function $h \in G$, we denote by $G_h \subset G$ the convex subset of functions such that $g + \hat{g} = 2h$. If $g \in G_h$, then clearly $\hat{g} \in G_h$ and $H_{\hat{g}}(\rho, \sigma) = H_g(\sigma, \rho)$.

THEOREM 4.1 ([18]). *For each $\rho, \sigma \in \mathcal{M}$,*

$$H_g(\rho, \sigma) = a\mathrm{Tr}\,(\sigma - \rho)$$

$$+ \mathrm{Tr}\,(\sigma - \rho)\left\{b\rho^{-1} + c\sigma^{-1} + \int_0^\infty \frac{1+s}{L_\sigma + sR_\rho}d\mu(s)\right\}(\sigma - \rho)$$

$$= a\mathrm{Tr}\,(\sigma - \rho) + \mathrm{Tr}\,(\sigma - \rho)R_\rho^{-1}k(L_\sigma/R_\rho)(\sigma - \rho),$$

where

$$k(u) = \int_0^\infty \frac{1+s}{u+s}d\mu(s) = \frac{g(u) - a(u-1)}{(u-1)^2}. \tag{6}$$

Theorem 4.1 implies that if $a = 0$, H_g is a contrast functional on $\mathcal{M}$. The value of $g'(1) = a$ does not influence the Riemannian structure and connections defined by H_g, so that we may also use functions with $g'(1) \neq 0$, as it is sometimes more convenient, for example $g(u) = -\log u$.

Let us consider the mixture connection $\nabla^{(m)}$ on $\mathcal{M}$. A vector field on $\mathcal{M}$ is parallel with respect to $\nabla^{(m)}$ if and only if its -1-representation is a constant $\mathcal{M}_h$-valued function over $\mathcal{M}$. In the rest of the paper, we will deal only with such vector fields. The symbol $\mathcal{X}$ will denote the vector field such that the constant value of the -1-representation is X, similarly $\mathcal{Y}$, etc. Note that for such vector fields we have $[\mathcal{X}, \mathcal{Y}] = 0$.

Let us define the Riemannian metric λ^g on $\mathcal{M}$ by

$$\lambda_\rho^g(X, Y) = -\frac{\partial^2}{\partial s \partial t}H_g(\rho + sX, \rho + tY)|_{s,t=0}, \quad \forall X, Y \in T_\rho. \tag{7}$$

Then [18]

$$\lambda_\rho^g(X, Y) = \mathrm{Tr}\,X R_\rho^{-1}k_{\mathrm{sym}}(L_\rho/R_\rho)(Y),$$

with

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

Moreover, the function k_{sym} is operator monotone decreasing, hence λ^g is a monotone metric, with $F = 1/k_{\mathrm{sym}}$ the corresponding operator monotone function. Note also that if h is a fixed symmetric function in G, then λ^g defines the same monotone metric for each $g \in G_h$.

Conversely, if λ is a monotone metric with the operator monotone function F, then

$$h(u) = \frac{1}{2}\frac{(u-1)^2}{F(u)} \tag{8}$$

is a symmetric operator convex function with $h(1) = 0$, so that $\lambda = \lambda^h$. The condition $h''(1) = 1$ is equivalent to the normalization condition $F(1) = 1$. This gives a one-to-one correspondence between the monotone metrics and the convex sets G_h, with symmetric $h \in G$.

5. The p-connections

Let us fix a monotone metric λ and let h be given by (8). Let us choose some $g \in G_h$, then $\lambda = \lambda^g$. We define the affine connection $\nabla^{(g)}$ by

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

just as in the classical case. It is clear that the restriction of $\nabla^{(g)}$ to submanifolds of mutually commuting elements coincides with the classical α-connection, with $\alpha = 2g'''(1) + 3$. In contrast with the classical case, the condition $g \in G$ leads to a restriction on α. Indeed, we have

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

From this, $0 \geq g'''(1) \geq -3$ and therefore $\alpha \in [-3, 3]$ for each $g \in G$.

PROPOSITION 5.1. *The connections $\nabla^{(g)}$ and $\nabla^{(\hat{g})}$ are dual with respect to λ. Moreover, the connections are torsion-free.*

Proof: We have

$$\mathcal{X}\lambda_\rho(\mathcal{Y}, \mathcal{Z}) = -\frac{d}{du}\frac{\partial^2}{\partial t \partial s} H_g(\rho + uX + sY, \rho + uX + tZ)|_{s,t,u=0}$$

$$= -\frac{\partial^3}{\partial s \partial t \partial u} H_g(\rho + uX + sY, \rho + tZ)|_{s,t,u=0}$$

$$- \frac{\partial^3}{\partial s \partial t \partial u} H_{\hat{g}}(\rho + uX + tZ, \rho + sY)|_{s,t,u=0}$$

$$= \lambda_\rho(\nabla_{\mathcal{X}}^{(g)}\mathcal{Y}, \mathcal{Z}) + \lambda_\rho(\mathcal{Y}, \nabla_{\mathcal{X}}^{(\hat{g})}\mathcal{Z}),$$

so that duality is proved. Moreover, as $[\mathcal{X}, \mathcal{Y}] = 0$, the connection is torsion-free if $\nabla_{\mathcal{X}}^{(g)}\mathcal{Y} - \nabla_{\mathcal{Y}}^{(g)}\mathcal{X} = 0$, which is obvious. $\square$

If the function g is symmetric, then from Proposition 5.1, $\nabla^{(g)}$ is self-dual and torsion-free, hence it is the metric connection $\bar{\nabla}$. For $g \neq \hat{g}$, let us define $g_p = pg + (1 - p)\hat{g}$, then $g_p \in G_h$ for $p \in [0, 1]$ and $\hat{g}_p = g_{1-p}$. For λ and g fixed, the connection given by g_p will be called the p-connection and denoted by $\nabla^{(p)}$. Clearly, $\nabla^{(p)}$ is a convex mixture of $\nabla^{(g)}$ and $\nabla^{(\hat{g})}$,

$$\nabla^{(p)} = p\nabla^{(g)} + (1 - p)\nabla^{(\hat{g})}.$$

Thus we have a one-parameter family of torsion-free p-connections, satisfying $(\nabla^{(p)})^* = \nabla^{(1-p)}$. We have $\nabla^{(1/2)} = \tilde{\nabla}$ for all $g \in G_h$. In the rest of this section we will investigate the Riemannian curvature of the p-connections.

EXAMPLE 5.1. We see from (5) that the extreme boundary of G consists of functions

$$g_s(u) = \frac{1+s}{2} \frac{(u-1)^2}{u+s} \quad \text{for } s \geq 0,$$

$$g_\infty(u) = \frac{1}{2}(u-1)^2.$$

We have $\hat{g}_s = g_{s^{-1}}$ for $s > 0$ and $\hat{g}_0 = g_\infty$. In this case

$$G_{h_s} = \{g_p = pg_s + (1-p)\hat{g}_s, \quad p \in [0,1]\},$$

where $h_s = \frac{1}{2}(g_s + \hat{g}_s)$. For the corresponding metric we obtain a unique family of p-connections. In particular, if $s = 1$, $g_1 = h_1$ is symmetric and $G_{h_1} = \{h_1\}$. The corresponding metric is the Bures metric. Hence we see that for the Bures metric, we obtain only the metric connection, which is known to be not flat, see for example [7].

EXAMPLE 5.2. Let

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

Then $g_\alpha \in G$ for $\alpha \in [-3, 3]$ and $\hat{g}_\alpha = g_{-\alpha}$. The relative entropies H_{g_α} are (up to a linear term) the α-divergences defined by Hasegawa in [13]. It was also proved that $\lambda^{g_\alpha} = \lambda_\alpha$, the WYD metric, and $\nabla^{g_\alpha} = \nabla^{(\alpha)}$, the α-connection from Section 3, see also [14]. Hence, $\nabla^{(g_\alpha)}$ is flat. In particular, for $\alpha = \pm 1$ we get the BKM metric and the mixture and exponential connection. The family of p-connections for $g(u) = -\log(u)$ is

$$\nabla^{(p)} = p\nabla^{(m)} + (1-p)\nabla^{(e)}$$

In the classical case, this is an equivalent definition of the α-connection, $p = (1 - \alpha)/2$. In our case however, these connections are different from the α-connections which, by Theorem 3.1, have torsion-free duals with respect to the BKM metric if and only if $\alpha = \pm 1$.

To compute the Riemannian curvature tensor of $\nabla^{(p)}$, we use the theory of statistical manifolds due to Lauritzen [2]. A statistical manifold is a triple $(M, \lambda, \tilde{D})$, where M is a differentiable manifold, λ is a metric tensor and $\tilde{D}$ is a symmetric covariant 3-tensor called the skewness.

On M, a class of connections is introduced by

$$\nabla_X^{(p)} Y = \bar{\nabla}_X Y - \frac{1-2p}{2} D(X, Y), \tag{9}$$

where X, Y are smooth vector fields, $\bar{\nabla}$ is the metric connection and the tensor D is given by $\tilde{D}(X, Y, Z) = \lambda(D(X, Y), Z)$. Such connections are torsion-free, this is equivalent to symmetry of $\tilde{D}$, resp. D. Moreover, $(\nabla^{(p)})^* = \nabla^{(1-p)}$. Let R^p be the corresponding Riemannian curvature. The manifolds satisfying $R^p = R^{1-p}$ for all p are called conjugate symmetric. It was proved in [2] that the manifold is conjugate symmetric if and only if the tensor $F = \bar{\nabla}\tilde{D}$ is symmetric. From symmetry of $\tilde{D}$, it follows that F is symmetric if (and only if) it is symmetric in X and Y. We also have that if there is some $p \neq 1/2$, such that $R^p = R^{1-p}$, then the manifold is conjugate symmetric.

Let $g \in G$, then $(\mathcal{M}, \lambda^g, \tilde{D})$, where $D(\mathcal{X}, \mathcal{Y}) = \nabla_{\mathcal{X}}^{(g)}\mathcal{Y} - \nabla_{\mathcal{X}}^{(\hat{g})}\mathcal{Y}$, is a statistical manifold. The connections defined by (9) coincide with the p-connections if $p \in [0, 1]$. For simplicity, we denote this manifold by $(\mathcal{M}, g)$. If g is symmetric, then $\tilde{D} \equiv 0$ and $\nabla^{(p)} = \bar{\nabla}$ for all p; in this case, the manifold is trivially conjugate symmetric.

PROPOSITION 5.2. *Let us denote* $\bar{R} = R^{1/2}$. *Then*

$$R^p(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \bar{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) + \frac{1-2p}{2}\{F(\mathcal{Y}, \mathcal{X}, \mathcal{Z}, \mathcal{W}) - F(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W})\}$$

$$+ \frac{(1-2p)^2}{4}\{\lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})) - \lambda(D(\mathcal{X}, \mathcal{Z}), D(\mathcal{Y}, \mathcal{W}))\}.$$

Proof: We have $[\mathcal{X}, \mathcal{Y}] = 0$ and therefore

$$R^p(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \lambda(\nabla_{\mathcal{X}}^{(p)}\nabla_{\mathcal{Y}}^{(p)}\mathcal{Z} - \nabla_{\mathcal{Y}}^{(p)}\nabla_{\mathcal{X}}^{(p)}\mathcal{Z}, \mathcal{W}).$$

Let us now recall that

$$F(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \mathcal{X}\tilde{D}(\mathcal{Y}, \mathcal{Z}, \mathcal{W}) - \tilde{D}(\bar{\nabla}_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}, \mathcal{W}) - \tilde{D}(\mathcal{Y}, \bar{\nabla}_{\mathcal{X}}\mathcal{Z}, \mathcal{W}) - \tilde{D}(\mathcal{Y}, \mathcal{Z}, \bar{\nabla}_{\mathcal{X}}\mathcal{W}).$$

From (9) we get

$$\nabla_{\mathcal{X}}^{(p)}\nabla_{\mathcal{Y}}^{(p)}\mathcal{Z} = \bar{\nabla}_{\mathcal{X}}\bar{\nabla}_{\mathcal{Y}}\mathcal{Z} - \frac{1-2p}{2}\{\bar{\nabla}_{\mathcal{X}}D(\mathcal{Y}, \mathcal{Z}) + D(\mathcal{X}, \bar{\nabla}_{\mathcal{Y}}\mathcal{Z})\}$$

$$+ \frac{(1-2p)^2}{4} D(\mathcal{X}, D(\mathcal{Y}, \mathcal{Z})).$$

Moreover, from self-duality of $\bar{\nabla}$,

$$\lambda(\bar{\nabla}_{\mathcal{X}}D(\mathcal{Y}, \mathcal{Z}) + D(\mathcal{X}, \bar{\nabla}_{\mathcal{Y}}\mathcal{Z}), \mathcal{W}) = \mathcal{X}\tilde{D}(\mathcal{Y}, \mathcal{Z}, \mathcal{W}) - \tilde{D}(\mathcal{Y}, \mathcal{Z}, \bar{\nabla}_{\mathcal{X}}\mathcal{W}) + \tilde{D}(\mathcal{X}, \bar{\nabla}_{\mathcal{Y}}\mathcal{Z}, \mathcal{W})$$

and

$$\lambda(D(\mathcal{X}, D(\mathcal{Y}, \mathcal{Z})), \mathcal{W}) = \tilde{D}(\mathcal{X}, D(\mathcal{Y}, \mathcal{Z}), \mathcal{W}) = \lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})),$$

this follows from symmetry of the tensor $\tilde{D}$. Subtracting the expression with interchanged $\mathcal{X}$ and $\mathcal{Y}$ and using symmetry of $\bar{\nabla}$ completes the proof. □

COROLLARY 5.1. *Let $g \neq \hat{g}$ and let the connection $\nabla^{(g)}$ be flat. Then the manifold $(\mathcal{M}, g)$ is conjugate symmetric. Moreover, if $R^{p_0} = 0$ for some $p_0 \in (0, 1)$ then $R^p = 0$ for all $p \in [0, 1]$.*

Proof: If $\nabla^{(g)}$ is flat, then also its dual $\nabla^{(\hat{g})}$ is flat, therefore $0 = R^1 = R^0$ and the manifold is conjugate symmetric. From Proposition 5.2, we see that

$$0 = \bar{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) + \frac{1}{4}\{\lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})) - \lambda(D(\mathcal{X}, \mathcal{Z}), D(\mathcal{Y}, \mathcal{W}))\},$$

and therefore

$$R^p(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = p(p-1)\{\lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})) - \lambda(D(\mathcal{X}, \mathcal{Z}), D(\mathcal{Y}, \mathcal{W}))\}.$$

If this vanishes for some $p_0 \neq 0, 1$, then the term in brackets must be zero. □

Let λ be the BKM metric and $g(u) = -\log(u)$, then $\nabla^{(g)} = \nabla^{(m)}$ is flat. It is known [21] that in this case, the metric connection is not flat, hence $\bar{R} = R^{1/2} \neq 0$. It follows that $p\nabla^{(m)} + (1-p)\nabla^{(e)}$ is flat if and only if $p = 0$ or $p = 1$.

6. Operator calculus

In the following sections, we are going to prove that the connection $\nabla^{(g)}$ is flat if and only if $\nabla^{(g)} = \nabla^{(\alpha)}$ for some $\alpha \in [-3, 3]$. To do this, we will need to compute the derivatives of functions of the form $c(L_\rho, R_\rho)$. We use the same method as in [8].

Let c be a function, defined and complex analytic in a neighbourhood of $(\mathbb{R}^+)^2$ in $\mathbb{C}^2$. As the operators L_ρ and R_ρ commute and have the same spectrum as ρ, by the operator calculus we have

$$c(L_\rho, R_\rho) = \frac{1}{(2\pi i)^2} \int \int c(\xi, \eta) \frac{1}{\xi - L_\rho} \frac{1}{\eta - R_\rho} d\xi d\eta,$$

where we integrate twice around the spectrum of ρ. We have

$$\frac{d}{dt}c(L_{\rho+tX}, R_\rho)|_{t=0} = \frac{1}{(2\pi i)^2} \int \int c(\xi, \eta) \frac{1}{\xi - L_\rho} L_X \frac{1}{\xi - L_\rho} \frac{1}{\eta - R_\rho} d\xi d\eta,$$

$$\frac{\partial^2}{\partial s \partial t}c(L_{\rho+sX+tY}, R_\rho)|_{s,t=0} = \frac{1}{(2\pi i)^2} \int \int c(\xi, \eta) \left\{ \frac{1}{\xi - L_\rho} L_Y \frac{1}{\xi - L_\rho} L_X \frac{1}{\xi - L_\rho} \right.$$
$$\left. + \frac{1}{\xi - L_\rho} L_X \frac{1}{\xi - L_\rho} L_Y \frac{1}{\xi - L_\rho} \right\} \frac{1}{\eta - R_\rho} d\xi d\eta,$$

$$\frac{\partial^2}{\partial s \partial t} c(L_{\rho+sX}, R_{\rho+tY})|_{s,t=0} = \frac{1}{(2\pi i)^2} \int \int c(\xi, \eta) \frac{1}{\xi - L_\rho} L_X \frac{1}{\xi - L_\rho}$$

$$\times \frac{1}{\eta - R_\rho} R_Y \frac{1}{\eta - R_\rho} d\xi d\eta.$$

We express the derivatives in the form of divided differences [4]. Let us denote

$$T(x, y|z) = \frac{c(x, z) - c(y, z)}{x - y}, \tag{10}$$

$$T(z|x, y) = \frac{c(z, x) - c(z, y)}{x - y}, \tag{11}$$

$$T(x, y, z|w) = \frac{T(x, y|w) - T(y, z|w)}{x - z}, \tag{12}$$

$$T(x, y|z, w) = \frac{T(x, y|z) - T(x, y|w)}{z - w} = \frac{T(x|z, w) - T(y|z, w)}{x - y}. \tag{13}$$

Then we have:

(i) $T(x, y|z)$, $T(z|x, y)$, $T(x, y|z, w)$ are symmetric in x, y and z, w. $T(x, y, z|w)$ is symmetric in x, y, z.

(ii) $T(x, x|z) = \frac{\partial}{\partial x} c(x, z)$ and $T(z|x, x) = \frac{\partial}{\partial x} c(z, x)$,

(iii) $T(x, x, z|w) = \frac{\partial}{\partial x} T(x, z|w)$,

(iv) $T(x, x, x|w) = \frac{1}{2} \frac{\partial^2}{\partial x^2} c(x, w)$.

Let $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ be the spectral decomposition of ρ. Let $e_{ij} = |\psi_i\rangle\langle\psi_j|$, then $\{e_{ij} \mid i, j = 1, \ldots, n\}$ is a basis of $\mathcal{M}_n(\mathbb{C})$. Let $u_{ij} = L_{e_{ij}}$, $v_{ij} = R_{e_{ji}}$. Then $u_{ij}e_{kl} = \delta_{jk}e_{il}$ and $v_{ij}e_{kl} = \delta_{jl}e_{ki}$. We also have

$$L_\rho = \sum_i \lambda_i u_{ii}, \qquad R_\rho = \sum_i \lambda_i v_{ii}, \qquad c(L_\rho, R_\rho) = \sum_{i,j} c(\lambda_i, \lambda_j) u_{ii} v_{jj}.$$

Let $X = \sum_{i,j} x_{ij} e_{ij}$. Inserting this into the expressions for derivatives, we get

$$\frac{d}{dt} c(L_{\rho+tX}, R_\rho)|_{t=0} = \sum_{i,j,k} T(\lambda_i, \lambda_j|\lambda_k) x_{ij} u_{ij} v_{kk}. \tag{14}$$

Similarly,

$$\frac{\partial^2}{\partial s \partial t} c(L_{\rho+sX+tY}, R_\rho)|_{s,t=0} = \sum_{i,j,k,l} T(\lambda_i, \lambda_j, \lambda_k|\lambda_l)(x_{ij}y_{jk} + y_{ij}x_{jk}) u_{ik} v_{ll}, \tag{15}$$

$$\frac{\partial^2}{\partial s \partial t} c(L_{\rho+sX}, R_{\rho+tY})|_{s,t=0} = \sum_{i,j,k,l} T(\lambda_i, \lambda_j|\lambda_k, \lambda_l) x_{ij} y_{lk} u_{ij} v_{kl}. \tag{16}$$

7. Conjugate symmetry

Let $g \in G$ and let k be given by (6). Let us define the function $c : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ by $c(x, y) = (1/y)k(x/y)$. As we see from the integral representation, the function k is operator monotone decreasing, therefore it has an analytic extension to the right halfplane in $\mathbb{C}$. It follows that c is complex analytic in a neighbourhood of $(\mathbb{R}^+)^2$ and we may use the results of the previous section. Note also that for $\hat{g}$, $\hat{k}(u) = u^{-1}k(u^{-1})$ and $\hat{c}(x, y) = c(y, x)$. Moreover, $\bar{c}(x, y) = c(x, y) + \hat{c}(x, y) = (1/y)k_{\mathrm{sym}}(x/y)$ is the Chentsov–Morozova function. As follows from (6), $k(1) = \int_{[0,\infty]} d\mu = 1/2$, therefore $c(x, x) = \frac{1}{2x}$ for all g.

LEMMA 7.1. *Let c and $\hat{c}$ be as above. Then*

$$\lambda_\rho(\nabla_{\mathcal{X}}^{(g)}\mathcal{Y}, \mathcal{Z}) = 2\mathrm{Re}\frac{d}{ds}\mathrm{Tr}\left\{X\hat{c}(L_{\rho+sY}, R_\rho)(Z) + Y\hat{c}(L_{\rho+sX}, R_\rho)(Z) - \right.$$
$$\left. - Xc(L_{\rho+sZ}, R_\rho)(Y)\right\}\Big|_{s=0}.$$

Proof: From Theorem 4.1 we compute

$$\lambda_\rho(\nabla_{\mathcal{X}}^{(g)}\mathcal{Y}, \mathcal{Z}) = -\frac{\partial^3}{\partial s \partial t \partial u}\mathrm{Tr}\,(uZ - sX - tY)c(L_{\rho+uZ}, R_{\rho+sX+tY})(uZ - sX - tY)\big|_{s,t,u=0}$$

$$= -\frac{d}{ds}\mathrm{Tr}\left\{Xc(L_{\rho+sZ}, R_\rho)(Y) + Yc(L_{\rho+sZ}, R_\rho)(X)\right.$$
$$\left. - Xc(L_\rho, R_{\rho+sY})(Z) - Zc(L_\rho, R_{\rho+sY})(X)\right.$$
$$\left. - Yc(L_\rho, R_{\rho+sX})(Z) - Zc(L_\rho, R_{\rho+sX})(Y)\right\}\Big|_{s=0}.$$

For $\sigma, \rho \in \mathcal{M}$, $c(L_\sigma, R_\rho)$ is a positive operator on $\mathcal{M}_n(\mathbb{C})$ endowed with the inner product $\langle A, B\rangle = \mathrm{Tr}\,A^*B$. For hermitian X and Y we have

$$\mathrm{Tr}\,Xc(L_\sigma, R_\rho)(Y) + \mathrm{Tr}\,Yc(L_\sigma, R_\rho)(X) = 2\mathrm{Re}\mathrm{Tr}\,Xc(L_\sigma, R_\rho)(Y).$$

Clearly, for all $X \in \mathcal{M}_h$ and sufficiently small s, $\rho + sX \in \mathcal{M}$. Moreover,

$$\mathrm{Re}\,\mathrm{Tr}\,Xc(L_\rho, R_{\rho+sY})(Z) = \mathrm{Re}\,\mathrm{Tr}\,(Xc(L_\rho, R_{\rho+sY})(Z))^*$$
$$= \mathrm{Re}\,\mathrm{Tr}\,X\hat{c}(L_{\rho+sY}, R_\rho)(Z). \qquad \square$$

LEMMA 7.2. *Let $D(\mathcal{X}, \mathcal{Y}) = \nabla_{\mathcal{X}}^{(g)}\mathcal{Y} - \nabla_{\mathcal{X}}^{(\hat{g})}\mathcal{Y}$ and let $\tilde{D}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) = \lambda(D(\mathcal{X}, \mathcal{Y}), \mathcal{Z})$. Let us denote $c_r(x, y) = \hat{c}(x, y) - c(x, y) = c(y, x) - c(x, y)$ and let*

$$Q(X, Y, Z) = \frac{d}{ds}\mathrm{Tr}\,Xc_r(L_{\rho+sY}, R_\rho)(Z).$$

Then

$$\tilde{D}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) = 2\mathrm{Re}\{Q(X, Y, Z) + Q(Y, X, Z) + Q(X, Z, Y)\} = 6Q_{\mathrm{sym}}(X, Y, Z),$$

where Q_{sym} is the symmetrization of Q over X, Y, Z.

Proof: Straightforward from Lemma 7.1.   □

Let us now denote by $\bar{T}(x, y|z)$ resp. $R(x, y|z)$, etc. the expressions (10)–(13) for $c = \bar{c}$, resp. $c = c_r$. From Section 6 we find

$$Q(X, Y, Z) = \sum_{i,j,k} R(\lambda_i, \lambda_j|\lambda_k) x_{ki} y_{ij} z_{jk}. \tag{17}$$

Further,

$$\mathcal{X}Q(Y, Z, W) = \sum_{i,j,k,l} R(\lambda_i, \lambda_j, \lambda_k|\lambda_l)(x_{ij} z_{jk} + z_{ij} x_{jk}) w_{kl} y_{li} \tag{18}$$

$$+ \sum_{i,j,k,l} R(\lambda_i, \lambda_j|\lambda_k, \lambda_l) z_{ij} w_{jl} x_{lk} y_{ki}.$$

Clearly, $\mathcal{X}\tilde{D}(\mathcal{Y}, \mathcal{Z}, \mathcal{W})$ is the symmetrization of (18) over $\mathcal{Y}, \mathcal{Z}, \mathcal{W}$.

PROPOSITION 7.1. *Let*

$$S(x, y|z) = \frac{1}{2\bar{c}(x, y)}\{\bar{T}(x, z|y) + \bar{T}(y, z|x) - \bar{T}(x, y|z)\}.$$

Then the -1-representation $\bar{\nabla}_{\mathcal{X}}\mathcal{Y}(\rho) = \sum_{\alpha,\beta} d_{\alpha\beta} e_{\alpha\beta}$, *where*

$$d_{\alpha\beta} = \sum_i S(\lambda_\alpha, \lambda_\beta|\lambda_i)(x_{\alpha i} y_{i\beta} + y_{\alpha i} x_{i\beta}).$$

Proof: Let $h = \frac{1}{2}(g + \hat{g})$, then $\bar{\nabla} = \nabla^{(h)}$. In this case $c = \frac{1}{2}\bar{c} = \hat{c}$. From Lemma 7.1 and Eq. (14) we see that

$$\lambda_\rho(\bar{\nabla}_{\mathcal{X}}\mathcal{Y}(\rho), Z) = \mathrm{Re} \sum_{i,j,k} \bar{T}(\lambda_i, \lambda_j|\lambda_k)\{x_{ki} y_{ij} z_{jk} + y_{ki} x_{ij} z_{jk} - x_{ki} z_{ij} y_{jk}\}. \tag{19}$$

Let us denote $f_{\alpha\alpha}^1 = e_{\alpha\alpha}$, for $\alpha = 1, \ldots n$, $f_{\alpha\beta}^2 = e_{\alpha\beta} + e_{\beta\alpha}$, $\alpha \neq \beta$, and $f_{\alpha\beta}^3 = i(e_{\alpha\beta} - e_{\beta\alpha})$, $\alpha \neq \beta$. Then $\{f_{\alpha\alpha}^1, \alpha = 1, \ldots, n, f_{\alpha\beta}^k, k = 2, 3, \alpha < \beta = 2, \ldots, n\}$ forms a basis of T_ρ with elements mutually orthogonal with respect to each monotone metric λ. Moreover,

$$\lambda(f_{\alpha\beta}^k, f_{\alpha\beta}^k) = \begin{cases} \bar{c}(\lambda_\alpha, \lambda_\alpha), & k = 1, \\ 2\bar{c}(\lambda_\alpha, \lambda_\beta), & k \neq 1. \end{cases}$$

Suppose that

$$\bar{\nabla}_{\mathcal{X}}\mathcal{Y}(\rho) = \sum_{k;\alpha \leq \beta} a_{\alpha\beta}^k f_{\alpha\beta}^k,$$

then $\bar{\nabla}_{\mathcal{X}}\mathcal{Y}(\rho) = \sum_{\alpha,\beta} d_{\alpha\beta} e_{\alpha\beta}$, where $d_{\alpha\alpha} = a^1_{\alpha\alpha}$, $d_{\alpha\beta} = a^2_{\alpha\beta} + i a^3_{\alpha\beta}$, if $\alpha < \beta$ and $d_{\beta\alpha} = \bar{d}_{\alpha\beta}$. From (19) we compute

$$a^1_{\alpha\alpha} = 2\mathrm{Re}\sum_j S(\lambda_\alpha, \lambda_\alpha | \lambda_j) x_{\alpha j} y_{j\alpha},$$

$$a^2_{\alpha\beta} = \mathrm{Re}\sum_j S(\lambda_\alpha, \lambda_\beta | \lambda_j)\{x_{\alpha j} y_{j\beta} + y_{\alpha j} x_{j\beta}\},$$

$$a^3_{\alpha\beta} = \mathrm{Im}\sum_j S(\lambda_\alpha, \lambda_\beta | \lambda_j)\{x_{\alpha j} y_{j\beta} + y_{\alpha j} x_{j\beta}\}. \qquad \square$$

As we know from Section 5, $(\mathcal{M}, g)$ is conjugate symmetric if and only if

$$\mathcal{X}\tilde{D}(\mathcal{Y}, \mathcal{Z}, \mathcal{W}) - \mathcal{Y}\tilde{D}(\mathcal{X}, \mathcal{Z}, \mathcal{W}) + \tilde{D}(\mathcal{X}, \bar{\nabla}_{\mathcal{Y}}\mathcal{Z}, \mathcal{W}) \qquad (20)$$
$$+ \tilde{D}(\mathcal{X}, \mathcal{Z}, \bar{\nabla}_{\mathcal{Y}}\mathcal{W}) - \tilde{D}(\mathcal{Y}, \bar{\nabla}_{\mathcal{X}}\mathcal{Z}, \mathcal{W}) - \tilde{D}(\mathcal{Y}, \mathcal{Z}, \bar{\nabla}_{\mathcal{X}}\mathcal{W}) = 0.$$

Using Lemma 7.2, (17), (18) and Proposition 7.1, we express the above equality in terms of the divided differences and then insert the basis elements $f^k_{\alpha\beta}$. This and other further lengthy computations are best performed using some software suitable for symbolic calculations, like Maple or Mathematica.

The equalities $\bar{c}(x, y) = \bar{c}(y, x)$, $c_r(x, y) = -c_r(y, x)$ and the definition and properties of divided differences imply that

$$R(x, y|x) = \frac{1}{x - y} c_r(x, y) = R(x, y|y), \qquad (21)$$

$$R(x, x|x) = -\frac{\alpha}{6x^2}, \quad \text{where} \quad \alpha = 2g'''(1) + 3, \qquad (22)$$

$$R(x, y|z, w) = -R(z, w|x, y), \qquad (23)$$

$$\bar{T}(x, y|z, w) = \bar{T}(z, w|x, y), \qquad (24)$$

$$S(x, y|x) = \frac{1}{2}\frac{\partial}{\partial x}\log\bar{c}(x, y), \qquad (25)$$

$$S(x, x|y) = \frac{1}{2}\left\{2\frac{1 - x\bar{c}(x, y)}{x - y} - x\frac{\partial}{\partial x}\bar{c}(x, y)\right\}, \qquad (26)$$

$$S(x, x|x) = -\frac{1}{4x} \qquad (27)$$

for all $x, y, z, w > 0$.

THEOREM 7.1. *Let* $g \neq \hat{g}$ *and let* $\bar{g} = g + \hat{g}$, $g_r = \hat{g} - g$. *If* $(\mathcal{M}, g)$ *is conjugate symmetric, then*

$$-\alpha\bar{g}(u) = 2u g'_r(u) - g_r(u) + 2au + 2a \qquad (28)$$

for all $u > 0$, *where* $a = g'(1)$ *and* $\alpha = 2g'''(1) + 3$.

Proof: Let us write the equality (20) for the basis elements $f^k_{\alpha\beta}$ with $\alpha, \beta \in \{1, 2\}$, in this case, the resulting expression depends only on eigenvalues λ_1 and λ_2 of ρ. Let us put $X = Z = e_{11}$ and $Y = W = e_{12} + e_{21}$, and let $\lambda_1 = x$, $\lambda_2 = y$. We get

$$R(x, x, x|y) - R(x, x, y|x) + R(x, y|x, x) - R(x, x|x, y)$$
$$+ 3R(x, x|x)S(x, x|y) - S(x, x|x)(2R(x, y|x) + R(x, x|y)) = 0.$$

We have

$$c_r(x, y) = \frac{y g_r(x/y)}{(x - y)^2} + \frac{2a}{x - y},$$

$$\bar{c}(x, y) = \frac{y \bar{g}(x/y)}{(x - y)^2}.$$

From this and from (i)–(iv) and (21)–(27), we get the equation

$$2g_r''\left(\frac{x}{y}\right)\frac{x}{y} + 2a + \alpha \bar{g}'\left(\frac{x}{y}\right) + g_r'\left(\frac{x}{y}\right) = 0.$$

Putting $u = x/y$ and integrating this, taking into account that $\bar{g}(1) = 0$, $g_r(1) = 0$ and $g_r'(1) = -2a$, we get (28). $\square$

REMARK 7.1. Let $g \neq \hat{g}$, α and a be as above. According to Theorem 7.1, if $(\mathcal{M}, g)$ is conjugate symmetric, then

$$\frac{1 + \alpha}{2} \bar{g}(u) = g'(u^{-1}) + u g'(u) - au - a. \tag{29}$$

If h is symmetric, then $(\mathcal{M}, h)$ is, of course, conjugate symmetric. In such a case, $\alpha = a = 0$ and Eq. (29) reads

$$h(u) = h'(u^{-1}) + u h'(u),$$

which is fulfilled for all symmetric $h \in G$.

EXAMPLE 7.1. It is easily checked that (29) is satisfied for all $pg_\alpha + (1 - p)g_{-\alpha}$, $p \in [0, 1]$, $\alpha \in [-3, 3]$ (as it should be). On the other hand, it is not true for g_s from the extreme boundary of G, unless $s = 1$, which is symmetric (the Bures case), or $s = 0$, which corresponds to g_α, $\alpha = 3$.

8. Flat connections

As we know from Corollary 5.1 and Proposition 5.2, the connection $\nabla^{(g)}$ is flat if and only if
   (a) $(\mathcal{M}, g)$ is conjugate symmetric,
   (b) $\bar{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) + \frac{1}{4}\{\lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})) - \lambda(D(\mathcal{X}, \mathcal{Z}), D(\mathcal{Y}, \mathcal{W}))\} = 0$.
This holds also for symmetric g, in that case (a) is satisfied and $D = 0$.

LEMMA 8.1.

$$\bar{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \mathcal{X}\lambda(\bar{\nabla}_{\mathcal{Y}}\mathcal{Z}, \mathcal{W}) - \mathcal{Y}\lambda(\bar{\nabla}_{\mathcal{X}}\mathcal{Z}, \mathcal{W}) + \lambda(\bar{\nabla}_{\mathcal{X}}\mathcal{Z}, \bar{\nabla}_{\mathcal{Y}}\mathcal{W}) - \lambda(\bar{\nabla}_{\mathcal{Y}}\mathcal{Z}, \bar{\nabla}_{\mathcal{X}}\mathcal{W}).$$

Proof: The statement is proved similarly as Proposition 5.2, using self-duality and symmetry of $\bar{\nabla}$. □

As before, we compute

$$\mathcal{X}\lambda(\bar{\nabla}_{\mathcal{Y}}\mathcal{Z}, \mathcal{W}) = \text{Re}\{\mathcal{X}\bar{Q}(Y, Z, W) + \mathcal{X}\bar{Q}(Z, Y, W) - \mathcal{X}\bar{Q}(Y, W, Z)\},$$

where

$$\mathcal{X}\bar{Q}(Y, Z, W) = \sum_{i,j,k,l} \bar{T}(\lambda_i, \lambda_j, \lambda_k|\lambda_l)(x_{ij}z_{jk} + z_{ij}x_{jk})w_{kl}y_{li} \tag{30}$$

$$+ \sum_{i,j,k,l} \bar{T}(\lambda_i, \lambda_j|\lambda_k, \lambda_l)z_{ij}w_{jl}x_{lk}y_{ki}.$$

Moreover,

$$\lambda(X, Y) = \text{Tr}\, X\bar{c}(L_\rho, R_\rho)(Y) = \sum_{i,j} \bar{c}(\lambda_i, \lambda_j)x_{ji}y_{ij}. \tag{31}$$

The second term in (b) can be written in a form using $\tilde{D}$: let $\{b_j \mid j = 1, \ldots N\}$ be the orthonormal basis obtained by normalization of $\{f^k_{\alpha\beta} \mid k = 1, 2, 3, \alpha \leq \beta = 1, \ldots, n\}$, then

$$\lambda(D(\mathcal{X}, \mathcal{W}), D(\mathcal{Y}, \mathcal{Z})) - \lambda(D(\mathcal{X}, \mathcal{Z}), D(\mathcal{Y}, \mathcal{W})) \tag{32}$$

$$= \sum_j \{\tilde{D}(X, W, b_j)\tilde{D}(Y, Z, b_j) - \tilde{D}(X, Z, b_j)\tilde{D}(Y, W, b_j)\}.$$

Using Lemma 8.1, (30), (31), (32) and Proposition 7.1, we get from (b) an equation involving divided differences, and we may proceed in the same way as in Section 6.

PROPOSITION 8.1. *Let $g \in G$. If the connection $\nabla^{(g)}$ is flat, then*

$$(\alpha^2 - 1)\bar{g}(u) + \bar{g}'(u)(u - 1) - 2\bar{g}''(u)u(1 + u) + \alpha(g'_r(u) + 2a)(u - 1) + 8 = 0 \tag{33}$$

for all $u > 0$.

Proof: Let $X = Z = e_{11}$ and $Y = W = e_{12} + e_{21}$. From (b) we get the equation

$$2\bar{T}(x, x, x|y) - 2\bar{T}(x, x|x, y) - 2\bar{c}(x, y)S(x, y|x)^2 + 4\bar{c}(x, x)S(x, x|x)\bar{S}(x|y)$$

$$- 3\frac{R(x, x|x)}{\bar{c}(x, x)}(2R(x, y|x) + R(x, x|y)) + \frac{1}{\bar{c}(x, y)}(2R(x, y|x) + R(x, x|y))^2 = 0.$$

For $X = Z = e_{12} + e_{21}$, $Y = W = i(e_{12} - e_{21})$, Eq. (b) reads

$$4\bar{T}(y, y, x|x) + 4\bar{T}(x, x, y|y) - 8\bar{T}(x, y|x, y) + 4\bar{c}(x, x)S(x, x|y)^2$$
$$+ 4\bar{c}(y, y)S(y, y|x)^2 - (2R(x, y|x) + R(x, x|y))^2 - (2R(x, y|y) + R(y, y|x))^2 = 0.$$

As in the proof of Theorem 7.1, we get after some rearrangements

$$u\left[(g'_r(u) + 2a)^2 - (\bar{g}'(u))^2\right] + \bar{g}(u)\{2u\bar{g}''(u) + \bar{g}'(u) + \alpha(g'_r(u) + 2a)\} = 0$$

from the first equation, and

$$u\left[(g'_r(u) + 2a)^2 - (\bar{g}'(u))^2\right] + \{g'_r(u)u - g_r(u) + 2a\}^2 - \{\bar{g}'(u)u - \bar{g}(u)\}^2 + 8\bar{g}(u) = 0,$$

from the second equation.

If g is symmetric, then in the above two equations $\alpha = a = 0$ and $g_r = 0$. From this we get

$$\bar{g}(u)\{-\bar{g}(u) + \bar{g}'(u)(u - 1) - 2\bar{g}''(u)u(1 + u) + 8\} = 0,$$

which is (33).

Let now $g \neq \hat{g}$. From (a), $(\mathcal{M}, g)$ is conjugate symmetric, and therefore (28) holds. From this

$$g'_r(u)u - g_r(u) + 2a = -\alpha\bar{g}(u) - u\{g'_r(u) + 2a\}.$$

Inserting this into the second equation and after some further computation we get (33). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We are now in a position to prove our main theorem.

THEOREM 8.1. *Let* $g \in G$ *and* $\alpha = 2g'''(1) + 3$. *Then* $\alpha \in [-3, 3]$ *and the connection* $\nabla^{(g)}$ *is flat if and only if* $\nabla^{(g)} = \nabla^{(\alpha)}$.

Proof: Let g be symmetric and suppose that $\nabla^{(g)}$ is flat. Then $\bar{g} = 2g$ and we get from (33) that g is a solution of

$$-g(u) + g'(u)(u - 1) - 2g''(u)u(1 + u) + 4 = 0$$

with the initial conditions $g(1) = 0$, $g'(1) = 0$. The unique solution of this equation is

$$g(u) = 2(1 - \sqrt{u})^2 = g_0.$$

If $g \neq \hat{g}$, then from (28) and (33) we get that g_r is the solution of

$$(\alpha^2 - 1)g_r(u) - (\alpha^2 - 1)(1 + u)g'_r(u) + 4u(u + 2)g''_r(u)$$
$$+ 4u^2(u + 1)g'''_r(u) - 4a(\alpha^2 - 1) + 8\alpha = 0$$

with $g_r(1) = 0$, $g'_r(1) = -2a$ and $g''_r(1) = 0$. If $\alpha \neq \pm 1$, the unique solution is

$$g_r(u) = \frac{4}{1-\alpha^2}(u^{\frac{1+\alpha}{2}} - u^{\frac{1-\alpha}{2}}) - \left(\frac{4\alpha}{1-\alpha^2} + 2a\right)(u-1)$$

$$= g_{-\alpha}(u) - g_\alpha(u) - 2\big(a - g'_\alpha(1)\big)(u-1),$$

and from (28) we get $\bar{g} = g_\alpha + g_{-\alpha}$.

If $\alpha = -1$, then the solution of the above equation is

$$g_r(u) = \log(u)(u+1) - 2(a - g'_{-1}(1))(u-1),$$

and from (28) we get

$$\bar{g}(u) = \log(u)(u-1).$$

It follows that $g = g_\alpha$, up to an additional linear term $(g'(1) - g'_\alpha(1))(u-1)$. □

COROLLARY 8.1. *Let* λ *be a monotone Riemannian metric and let* $\bar{\nabla}$ *be the metric connection. Then* $\bar{\nabla}$ *is flat if and only if* λ *is the WY metric* ($\alpha = 0$).

Proof: Let G_h be the convex subset of G, corresponding to λ. Then $\bar{\nabla} = \nabla^{(h)}$ and $h = \hat{h}$ implies that $h'''(1) = -\frac{3}{2}$. The proof now follows from Theorem 8.1. □

Acknowledgement

REFERENCES

[1] S. Amari: *Differential-geometrical Methods in Statistics*, Lecture Notes in Statistics **28**, Springer, Berlin 1985.

[2] S. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen and C. R. Rao: *Differential Geometry in Statistical Inference*, IMS Lecture Notes-Monograph Series 10, Mayward 1987.

[3] S. Amari and H. Nagaoka: *Method of Information Geometry*, AMS monograph, Oxford University Press, 2000.

[4] R. Bhatia: *Matrix Analysis*, Springer, New York 1997.

[5] N. N. Chentsov: *Statistical Decision Rules and Optimal Inferences*, Translation of Math. Monog. **53**, AMS, Providence, 1982.

[6] N. N. Chentsov and Morozova: Markov invariant geometry on state manifolds, *Itogi Nauki i Tekhniki* **36** (1990), 69.

[7] J. Dittmann: *J. Geom. Phys.* **31** (1999), 16.

[8] J. Dittmann: *Lin. Alg. Appl.* **315** (2000), 83.

[9] S. Eguchi: *Ann. Statist.* **11** (1983), 793.

[10] M. R. Grasselli: Monotonicity, Duality and Uniqueness of the WYD Metrics, submitted to IDAQP, math-ph/0212022.

[11] M. R. Grasselli and R. F. Streater: IDAQP, **4** (2001), 173, math-ph/0006030

[12] H. Hasegawa: *Rep. Math. Phys.* **39** (1997), 49.

[13] H. Hasegawa: *Rep. Math. Phys.* **33** (1993), 87.

[14] H. Hasegawa: On the dual geometry of Wigner–Yanase–Dyson information contents, 2002, to appear in IDAQP.

[15] H. Hasegawa and D. Petz: Non-commutative extension of information geometry II, in: *Quantum Communication, Computing and Measurement*, Hirota et al. (eds.), Plenum Press, New York 1997.

[16] A. Jenčová: *Rep. Math. Phys.* **47** (2001), 121.

[17] A. Jenčová: Dualistic properties of the manifold of quantum states, in: *Disordered and Complex Systems*, AIP Conf. Proc. **553**, AIP, Melville, NY, 2001.

[18] A. Lesniewski and M. B. Ruskai: *J. Math. Phys.* **40** (1999), 5702, math-ph/9808016.

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

[20] D. Petz: *Rep. Math. Phys.* **23** (1986), 57.

[21] D. Petz: *J. Math. Phys.* **244** (1994), 780.

[22] D. Petz: *Linear Algebra Appl.* **244** (1996), 81.