Quantum Communication, Computing, and Measurement
Edited by Hirota *et al.*, Plenum Press, New York, 1997

NON-COMMUTATIVE EXTENSION OF INFORMATION GEOMETRY II

Hiroshi Hasegawa[1] and Dénes Petz[2]

[1] Research Center for Quantum Communications
Tamagawa University
Tamagawa-gakuen, Machida, Tokyo 194, Japan

[2] Department for Mathematical Analysis
Budapest University of Technology and Economics
Műegyetem rakpart 3–9
H–1111 Budapest, Hungary
e-mail: petz@math.bme.hu

The Fisher information provides a canonical Riemannian metric in the geometric approach to classical statistics. It seems that the quantum analogue of the Fisher information is not uniquely defined, and it is necessary to study the possible candidates and to compare them on physical grounds. Description of monotone metrics under coarse graining has been given by Petz and this class of metrics fixes many candidates. Here we show that the skew information $I_p(\rho, K) \equiv -\frac{1}{2}\mathrm{Tr}\,[\rho^p, K][\rho^{1-p}, K]$ first introduced by Wigner, Yanase and Dyson (WYD) many years ago yields a monotone metric for all values of p; $-1 \le p \le 2$ (for $p = 0, 1$ under a proper limiting procedure and beyond the limits with a change of the sign of I_p). Furthermore, we argue that the symmetry between I_p and I_{1-p} is identical to the quantum version of Amari's duality concept for smooth statistical manifolds.

1 Introduction

This presentation is a summary of our collaboration on the extension of *information geometry* in classical statistics to the quantum mechanical setting since the last QCM workshop held in 1994. We deal with the construction of a Riemannian geometry on a manifold of finite dimensional density matrices parametrized by $\theta = (\theta^1, \theta^2, \ldots, \theta^r) \in \Theta$ in a matrix algebra $\mathcal{M}_n$. References on this topic from a general point of view are [1] and [2]. In the present workshop, we revisit the old notion of *skew information* [3] from the newly developed quantum geometric standpoint.

In 1963, Wigner and Yanase [4] investigated a trace quantity of matrices in $\mathcal{M}_n$,

$$-\mathrm{Tr}\,[\rho^{1/2}, k]^2 \qquad ([\,.\,,\,.\,] \text{ represents a commutator}), \tag{1.1}$$

where ρ and k are self-adjoint elements in $\mathcal{M}_n$, in particular, ρ is positive definite and $\mathrm{Tr}\,\rho = 1$, i.e. a density matrix, while k is supposed to represent a conserved quantity of the quantum system under consideration: they proved that the above quantity is convex with respect to ρ like $\mathrm{Tr}\,\rho \log \rho$ and represents a kind of information content. They further supplemented to this work in [5], saying that the amount of this information depends on the non-commutativity of ρ and k that is needed for a measurement of non-commuting observables with k by means of external apparatus; hence the origin of the naming "information content relative to k". It is important to note that Dyson [4] was said to propose a generalizing idea to expression (1.1) of Wigner and Yanase in regard to retaining the nature of information i.e.

$$I_p(\rho, k) = -\mathrm{Tr}\,[\rho^p, k][\rho^{1-p}, k], \qquad 0 < p < 1, \tag{1.2}$$

guessing that it would be possible to prove the convexity of this expression with respect to ρ: this is the so-called Wigner–Yanase–Dyson conjecture whose correctness was verified later fully by Lieb [6] in a more general context. (For various modifications of (1.2), see [3]).

Being motivated by the classical information geometry discussed by Amari in his Lecture Notes [7], Hasegawa [8] proposed to define the quantum mechanical α-divergence $D_\alpha(\rho, \sigma)$ for two density matrices ρ and σ:

$$D_\alpha(\rho, \sigma) = \frac{4}{1-\alpha^2}\mathrm{Tr}\,(\rho - \sigma^{\frac{1+\alpha}{2}}\rho^{\frac{1-\alpha}{2}}), \qquad \alpha \in \mathbb{R} = (-\infty, \infty), \tag{1.3}$$

$$\mathrm{Tr}\,\rho = \mathrm{Tr}\,\sigma = 1, \tag{1.4}$$

which is a natural generalization of the α-divergence $D_\alpha(p, q)$ for classical probability density functions $p(x)$ and $q(x)$ in Amari's formulation. It is well-known that $D_\alpha(p, q)$ is a globally defined distance function between the two probabilities p and q but lacking the symmetry between them, i.e. $D_\alpha(p, q) \neq D_\alpha(q, p)$. What is significant about this function for the geometry of the statistical manifold is, when the two density functions are infinitesimally distant from each other, the Riemannian metric structure arises in such a way that

$$\begin{aligned} D_\alpha(p + dp, p) = D(p, p + dp) &= \frac{1}{2}E_p(d\log p, d\log p) \\ &= \frac{1}{2}E_p\left(\frac{\partial \log p}{\partial \theta^i}\frac{\partial \log p}{\partial \theta^j}\right) d\theta^i d\theta^j \text{ up to } O(d\theta)^2, \end{aligned} \tag{1.5}$$

for the expansion

$$dp = \frac{\partial p}{\partial \theta^i} d\theta^i, \tag{1.6}$$

where

$$E_p\left(\frac{\partial \log p}{\partial \theta^i}\frac{\partial \log p}{\partial \theta^j}\right) = \int p(x)\frac{\partial \log p}{\partial \theta^i}\frac{\partial \log p}{\partial \theta^j}\, d\mu(x) \equiv g^F_{ij}(\theta) \tag{1.7}$$

represents the Fisher information metric tensor expressed as a scalar product of the tangent vector

$$\partial_i :\ \partial_i \log p = \frac{\partial}{\partial \theta_i}\log p. \tag{1.8}$$

The success of the definition of the quantum α-divergence (1.3) and (1.4) can be seen in the full parallelism of its Riemannian metric structure compared to the classical formulation above for the two infinitesimally distant density matrices ρ and $\rho + d\rho$. Here, $d\rho$ consists of a commutative part $d^c\rho$ with ρ and a non-commutative part such that

$$\begin{aligned} d\rho &= d^c\rho + [\rho, \Delta] \qquad (\Delta \text{ is an anti-selfadjoint}) \qquad (1.9)\\ &= \left(\frac{\partial \rho}{\partial \theta_i} + [\rho, \Delta_i]\right) d\theta^i, \end{aligned}$$

$$d^c\rho = \frac{\partial \rho}{\partial \theta^i} d\theta_i \quad \text{and} \quad \Delta = \Delta_i d\theta^i \tag{1.10}$$

in terms of which

$$\begin{aligned} D_\alpha(\rho + d\rho, \rho) &= D_\alpha(\rho, \rho + d\rho)\\ &= \frac{1}{2}\mathrm{Tr}\, \rho(d^c \log \rho d^c \log \rho) + \frac{2}{1-\alpha^2}\mathrm{Tr}\, [\rho^{\frac{1-\alpha}{2}}, \Delta][\rho^{\frac{1+\alpha}{2}}, \Delta]\\ &= \frac{1}{2}\left(\mathrm{Tr}\, \rho\frac{\partial \log \rho}{\partial \theta^i}\frac{\partial \log \rho}{\partial \theta^j} + \frac{4}{1-\alpha^2}\mathrm{Tr}\, [\rho^{\frac{1-\alpha}{2}}, \Delta_i][\rho^{\frac{1+\alpha}{2}}, \Delta_j]\right) d\theta^i d\theta^j\\ &\equiv \frac{1}{2} g^{(\alpha)}_{ij} d\theta^i d\theta^j \qquad (1.11) \end{aligned}$$

up to $O(d\theta)^2$ and a $\alpha \neq \pm 1$. The limit $\alpha \to \pm 1$ yields the familiar Kubo–Mori metric, which should be rendered to a full treatment separately (see [9] and [10]).

We can observe that the non-commutative part of $d\rho$ with ρ gives rise to an additional Riemannian metric in the form

$$\frac{1}{2p(1-p)}\mathrm{Tr}\, [\rho^p, \Delta][\rho^{1-p}, \Delta], \qquad p = \frac{1-\alpha}{2}, \tag{1.12}$$

which, apart from the factor $1/p(1-p)$, is just equal to the WYD skew information $I_p(\rho, i\Delta)$.

From this observation, we see that the Riemannian metric tensor for quantum information geometry is of such a structure that it consists of two parts; the first part arising from the commutative $d\rho$ and the second from the non-commutative $d\rho$ with ρ, the first part being identified precisely with the classical Fisher form but the second part depending on the starting definition of the inner product on matrix spaces that induces the metric. We need, therefore, some axiomatic specification of inner products on $\mathcal{M}_n$ for a Riemannian metric to be well-defined as a quantum information metric. The *theory of monotone metrics on matrix spaces* by Petz [11] meets this need, which we will discuss in the next two sections. In Section 4 we will discuss a problem of characterizing the WYD metric in terms of Amari's notion of *duality* in the information geometry.

2 Monotone Riemannian Metrics for Quantum Statistics [11]

We consider $\mathcal{M}_n(\mathbb{C})$ which can be regarded as a vector space whose elements are denoted by $A, B, \ldots$ and their adjoint by $A^*, B^*, \ldots$. An inner product of two elements A, B can be defined by a sesquilinear form $K(B, A)$ (linear in A and anti-linear in B), which may be written in terms of a linear superoperator K on $\mathcal{M}_n(\mathbb{C})$, $A \mapsto K(A) \in \mathcal{M}_n(\mathbb{C})$, and the standard Hilbert-Schmidt inner product as

$$K(B, A) = \langle B, K(A) \rangle = \operatorname{Tr} B^* K(A). \tag{2.1}$$

In case that the linear superoperator K is positive definite and self-adjoint on $\mathcal{M}_n(\mathbb{C})$ (in the Hilbert-Schmidt sense), a positive quantity $K(A, A)$ induces a length of A which is called here a metric of A. We suppose further that the superoperator K depends on a density matrix $\rho \in \mathcal{M}_n^{++}$ (all positive definite self-adjoint elements in $\mathcal{M}_n$) and $\operatorname{Tr} \rho = 1$, which is denoted by K_ρ; also the corresponding positive form K by K_ρ. We then impose a physical condition on the metric defined by K_ρ and A that by any stochastic mapping T of elements ρ and A (i.e. a completely positive map of ρ and A) the resulting metric does not exceed the original one: it means that a coarse-graining of a (quantum) state may reduce but may not increase the metric. All these requirements are now listed as follows.

(a) $(A, B) \mapsto K_\rho(A, B)$ is sesquilinear,

(b) $K_\rho(A, A) \geq 0$ and the equality holds if and only if $A = 0$,

(c) $\rho \mapsto K_\rho(A, A)$ is continuous on $\mathcal{M}_n^{++}$ for every fixed A,

(d) monotonicity condition: $K_{T(\rho)}(T(A), T(A)) \leq K_\rho(A, A)$ for every stochastic mapping $T : \mathcal{M}_n(\mathbb{C}) \mapsto \mathcal{M}_m(\mathbb{C})$.

The metric $K_\rho(A, A)$ which satisfies (a)–(d) will be called a *monotone metric.* In condition (d) the stochastic mapping T may be taken to be any unitary conjugation $U^* \cdot U$ for which the inequality must be replaced by the equality, implying the unitary covariance

(d′) $K_{U^*\rho U}(U^*AU, U^*AU) = K_\rho(A, A)$.

It implies the fact that the density matrix ρ of K_ρ may be understood to be represented always as diagonal.

For the purpose of application to geometry, it is sufficient to restrict the metric tensor to be real i.e. $g_{ji} = g_{ij} =$ real. For this to be the case, it is more convenient to restrict the vectors A, B, $\ldots$ always to be self-adjoint: then expression (2.1) may be replaced by

$$\begin{aligned} K'(B, A) &= \frac{1}{2}(K(B, A) + K(A, B)) \\ &= K'(A, B), \qquad A^* = A, \quad B^* = B, \end{aligned} \tag{2.2}$$

which is a symmetric, bilinear form of any pair of self-adjoint elements in $\mathcal{M}_n(\mathbb{C})$, and

$$K'_\rho(A,A) = \sum_{i=1}^{n} c(\lambda_i)A_{ii}^2 + 2\sum_{i<j} c(\lambda_i,\lambda_j)|A_{ij}|^2, \tag{2.3}$$

$$\rho = \sum_{i=1}^{n} \lambda_i e_{ii} \qquad (0<\lambda_i<1).$$

The two-variable function $c(\lambda,\mu) = c(\mu,\lambda)$ is called "Morozova–Chentsov function" after the names of Russian mathematicians who obtained the expression for the first time [12]. Conditions (b), (c) and (d) are shown to restrict the possible form of the two functions $c(\lambda\ )$ and $c(\lambda,\mu)$ in (2.3), reducing their degree of freedom to one. Namely,

THEOREM 1 *Suppose that* (b), (c) *and* (d) *hold for a real, bilinear form* $K'_\rho(A,B)$ *on self-adjoint element in* $\mathcal{M}_n$. *Then, the two functions* $c(\lambda)$ *and* $c(\lambda,\mu)$ *on* $\mathbb{R}^+\times\mathbb{R}^+$ *in the* ρ*-diagonalized representation of the metric* $K'_\rho(A,A)$ *satisfy the following:*

(i) $c(\lambda)$, $c(\lambda,\mu) = c(\mu,\lambda)$ *are continuous, positive functions,*

(ii) $\lim_{\mu\to\lambda} c(\lambda,\mu) = c(\lambda) = c\lambda^{-1}$, *with a positive constant* $c=1$ *hereafter.*

(iii) $c(\lambda,\mu)$ *is homogeneous of order* -1 *in* λ *and* μ, *implying* $c(x\lambda,x\mu) = \dfrac{1}{x}c(\lambda,\mu)$ *for any* $x>0$.

Different from the approach in [12], Petz has given a proof of Theorem 1 by means of *operator monotone function* initiated by Löwner in 1932. We do not go into detail, but in order to show the power of this notion for the present problem we summarize a result which clarifies the Morozova–Chentsov (M–C) function $c(\lambda,\mu)$ by means of the operator monotonicity in Theorem 2 after giving its definition:

DEFINITION 1 A real non-negative function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is operator monotone, if, for any x and $y \in \mathcal{M}_n^{++}$ and for every $n \in N$, $x \le y$ implies $f(x) \le f(y)$. (See for example [13], [15]).

THEOREM 2 *There exists a one-to-one correspondence (apart from normalization) between the M–C function* $c(\lambda,\mu)$ *for a symmetric monotone metric* $K'_\rho(A,A)$ *and an operator monotone function* f *as follows*

$$\text{i)}\quad f(x) = \frac{1}{c(x,1)}, \qquad \text{ii)}\quad c(\lambda,\mu) = \frac{1}{\mu f(\lambda/\mu)}, \tag{2.4}$$

$$\textit{where by virtue of symmetry } c(\lambda,\mu) = c(\mu,\lambda),\ xf(1/x) = f(x). \tag{2.5}$$

That is, for a given M–C function $c(\lambda,\mu)$, (x) defined by i) is operator monotone and, conversely, for a given operator monotone function f the function $c(\lambda,\mu)$ defined by ii) yields an M–C function.

Previously, three important monotone metrics have been given together with many other examples [11]:

Symmetric logarithmic-derivative metric K^{Sl}

$$c(\lambda,\mu)=\frac{2}{\lambda+\mu}\qquad\text{for which}\quad f(x)=\frac{1}{2}(x+1); \tag{2.6}$$

Kubo–Mori metric K^{KM}

$$c(\lambda,\mu)=\frac{\log\lambda-\log\mu}{\lambda-\mu}\quad\text{for which}\quad f(x)=\frac{x-1}{\log x}; \tag{2.7}$$

Right logarithmic-derivative metric K^{Rl}

$$c(\lambda,\mu)=\frac{1}{2}\left(\frac{1}{\lambda}+\frac{1}{\mu}\right)\quad\text{for which}\quad f(x)=\frac{2x}{(x+1)}. \tag{2.8}$$

From a viewpoint of *operator mean*, Kubo and Ando [13] obtained a remarkable result that any normalized operator monotone function f lies between the maximum $f_{\max}(x)=\frac{1}{2}(x+1)$ (arithmetic mean) and the minimum $f_{\min}(x)=\frac{2x}{x+1}$ (harmonic mean). Correspondingly, for every monotone metric $K_\rho(A,A)$ with a fixed A the following inequalities hold

$$K_\rho^{Sl}\le K_\rho\le K_\rho^{Rl}. \tag{2.9}$$

Therefore, our present interest is whether the WYD metrics are monotone metrics or not and, if they are, in what ranges they lie in the inequalities (2.9). The full answer will be seen in the next section.

3 Wigner–Yanase–Dyson Metrics with Monotonicity

We begin by determining the M–C function $c(\lambda,\mu)$ for the WYD metric (1.12) and the corresponding f-function. If this function can be proved to be operator monotone, then we can say that the WYD metric is surely a monotone metric by virtue of Theorem 2. The determination can be easily done by taking the ρ-diagonal representation of the matrix under the trace in (1.12), showing that

$$c(\lambda_j,\lambda_k)|A_{jk}|^2=\frac{1}{p(1-p)}(\lambda_j^p-\lambda_k^p)(\lambda_j^{1-p}-\lambda_k^{1-p})|\Delta_{jk}|^2,\qquad j\neq k.$$

But since $A_{jk}=(\lambda_j-\lambda_k)\Delta_{jk}$ and $\Delta_{jk}^*=-\Delta_{kj}$, we get

$$c(\lambda,\mu)=\frac{1}{p(1-p)}\frac{(\lambda^p-\mu^p)(\lambda^{1-p}-\mu^{1-p})}{(\lambda-\mu)^2}, \tag{3.1}$$

for which properties i), ii) and iii) in Theorem 1 can be checked. Hence

$$f(x)=p(1-p)\frac{(x-1)^2}{(x^p-1)(x^{1-p}-1)}\equiv f_p(x), \tag{3.2}$$

and changing the parameter p to α, $p=\frac{1-\alpha}{2}$

$$f_\alpha(x) = \frac{1-\alpha^2}{4}\frac{(x-1)^2}{(x^{\frac{1-\alpha}{2}}-1)(x^{\frac{1+\alpha}{2}}-1)}. \tag{3.2$'$}$$

Let us ask now whether this function is operator monotone. An affirmative answer has been obtained for the case $0 < p < 1$ [14]. Here we outline the full answer both for $0 < p < 1$ and for $-1 < p < 0$ on an equal footing. (Case for $1 < p < 2$ reduces to the latter case.) The keyword of the proof is an integral representation of the inverse of $f_p(x)$ in (3.2), which can be given by

$$\frac{1}{f_p(x)} = \frac{\sin p\pi}{\pi}\int_0^\infty d\lambda\lambda^{p-1}\int_0^1 ds\int_0^1 dt\frac{1}{x((1-t)\lambda+(1-s))+(t\lambda+s)}, \quad 0 < p < 1, \tag{3.3}$$

$$= \frac{\sin p\pi}{p\pi}\int_0^\infty d\lambda\lambda^{-|p|}\int_0^1 ds\int_0^1 dt\frac{x(1-t)+t}{[x((1-t)\lambda+(1-s))+(t\lambda+s)]^2}, \quad -1 < p < 0. \tag{3.3$'$}$$

This is based on the integral representation of a fractional power of a positive quantity, namely

$$z^{p-1} = \frac{\sin p\pi}{\pi}\int_0^\infty \frac{\lambda^{p-1}}{\lambda+z}d\lambda, \qquad 0 < p < 1, \tag{3.4}$$

$$= \frac{\sin p\pi}{p\pi}\int_0^\infty \frac{\lambda^{-|p|}}{(\lambda+z)^2}d\lambda, \qquad -1 < p < 0. \tag{3.4$'$}$$

(The latter is obtainable from the former by replacing $p \to 1-p$ and a differentiation). Accordingly, we perform the λ-integration first in (3.3) and (3.3$'$) and then the rest double integrations which can be made elementarily to obtain the desired expressions.

In order to assure the operator monotonicity of $f_p(x)$ in both cases, let us set up another definition in comparison with Definition 1.

DEFINITION 2 We say, a real non-negative function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ is operator monotone-decreasing, if, for any x and $y \in \mathcal{M}_n^{++}$ and far every $n \in N$, $x \le y$ implies $f(x) \ge f(y)$.

Evidently, by Definitions 1 and 2, the inverse of an operator monotone function is operator monotone-decreasing and *vice versa*. Then, it is easy to see that $f_p(x)$ given in (3.3) is operator monotone, because the integrand is operator monotone-decreasing (a linear function $ax+b$, $a > 0$, $b \ge 0$, is operator monotone and so $(ax+b)^{-1}$ is operator monotone-decreasing). Since the triple integration there is made always with positive coefficients, the result $\frac{1}{f_p(x)}$ is operator monotone-decreasing, and hence $f_p(x)$ is operator monotone. Thus, the same reasoning also applies to the expression (3.3$'$), if its integrand is shown to be operator monotone-decreasing, and this is indeed the case: a function of the form $(cx+d)/(ax+b)^2$, $a, c > 0$ and $b, d \ge 0$, is operator monotone-decreasing [15].

We thus establish the operator monotonicity of the f-functions associated with the WYD metrics given by (3.2) or (3.2$'$) for all cases. Strictly speaking, the above proof may lose the validity for some special parameter values because of a divergence of the integration (3.3) or (3.3′); the case $p = \pm 1$, and 0 listed as

$p =$	-1	0	$+1$
$\alpha =$	3	1	-1
	(Right log. derivative	Kubo–Mori	Kubo–Mori)

The operator monotonicity for these situations has been verified by other means [11]. It is remarkable to see the absence of $f_{\max}$ in the above list (symmetric logarithmic-derivative metric): there exists maximum $f_{\max}$ of all the WYD f-functions which satisfies

$$f_{\max}^{\rm WYD}(x) = \frac{1}{4}(x^{1/2}+1)^2 < f_{\max}(x) = \frac{x+1}{2}.$$

This corresponds to the original Wigner–Yanase metric (1.1). There exists a gap between these two maxima which can not be filled by any WYD f-function. The feature can be seen in *Fig. 1*: This figure illustrates how the functions $f_\alpha(x)$ behave on the real line $[0,1]$ with α from 0 to 3, showing that $f_\alpha;\ \alpha > 3$ is forbidden.

4 Characterization of the WYD Metrics by Duality

Let ρ be a density matrix and $\mathcal{T}_\rho$ be the tangent space at ρ in the quantum statistical manifold. It was seen above that the monotone metric is uniquely determined in the subspace

$$\mathcal{T}_\rho^c = \{x \in \mathcal{T}_\rho : \rho X - X\rho = 0\}, \tag{4.1}$$

and therefore, it is worthwhile to decompose the tangent space as $\mathcal{T}_\rho = \mathcal{T}_\rho^c \oplus \mathcal{T}_\rho^X$, where $\mathcal{T}_\rho^X = \{{\rm i}[\rho, X] \in \mathcal{T}_\rho : X = X^*\}$. Using the superoperator $\mathcal{L}_\rho : X \mapsto {\rm i}[\rho, X]$, we may put this decomposition in the form ${\rm Ker}\mathcal{L}_\rho \oplus {\rm Rng}\,\mathcal{L}_\rho$.

It is a useful observation that the kernel and the range of $\mathcal{L}_\rho$ remain unchanged when ρ is replaced by $g(\rho)$ and g is a monotone function, $\mathbb{R}^+ \to \mathbb{R}$. We say that a symmetric metric K_ρ admits dual affine connections if, for a pair of such monotone functions g and g^*,

$$K_\rho(\mathcal{L}_\rho(X), \mathcal{L}_\rho(Y)) = \langle \mathcal{L}_{g(\rho)}(X), \mathcal{L}_{g^*(\rho)}(Y)\rangle. \tag{4.2}$$

Note that (4.2) is equivalently written as

$$K_\rho(A,B) = \frac{\partial^2}{\partial t \partial s} {\rm Tr}\, g(\rho + tA) g^*(\rho + sB)|_{t=s=0}, \qquad A = {\rm i}[\rho, X],\ B = {\rm i}[\rho, Y].$$

It follows that

$$\delta_C K_\rho(A,B) = \langle \delta_C \delta_A g(\rho), \delta_B g^*(\rho)\rangle + \langle \delta_A g(\rho), \delta_C \delta_B g^*(\rho)\rangle, \tag{4.3}$$

which stands for the quantum version of Amari's duality concept in affine connections. Namely, the first term of the right-hand side of (4.3) represents an affine connection as regards three directions A, B and C and then the second term its dual connection as regards B, A and C, both adding up to yield a derivative of the metric in the direction C [7]. The WYD skew information is seen to have this form, where both g and g^* are the power function, and we want to show that this is the unique possibility for the metrics with this dual structure.

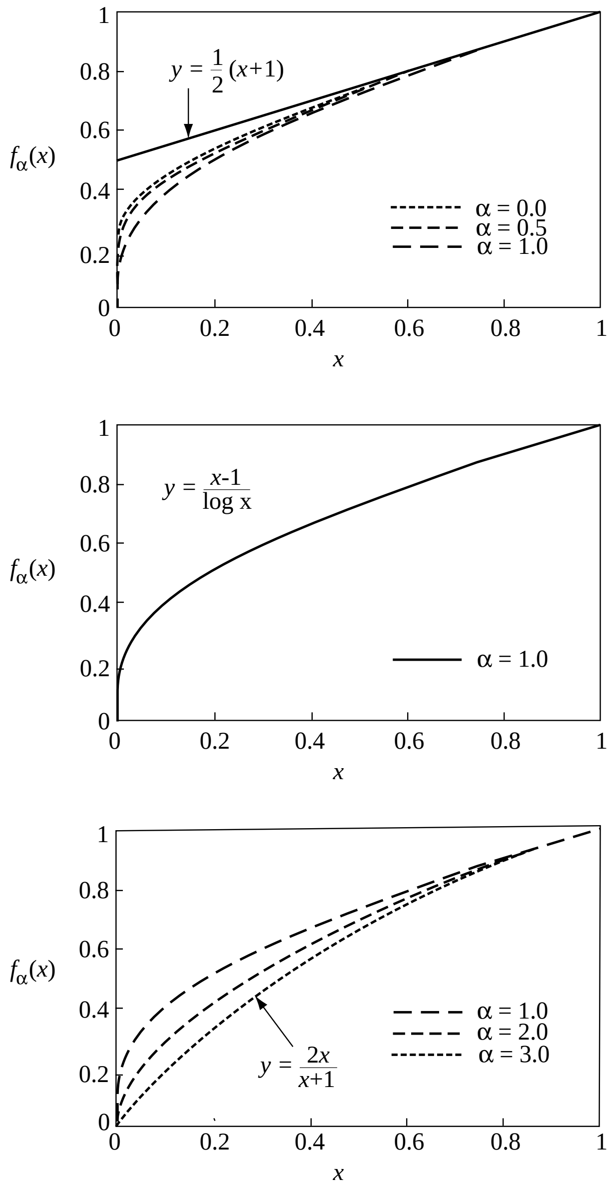

Figure 1:

THEOREM 3 *In the class of symmetric monotone metrics, the Wigner–Yanase–Dyson skew information is characterized by the property that it admits dual affine connections.*

In the sketch of the proof of this theorem we assume that the functions g and g^* appearing in definition (4.2) are continuous and smooth on $\mathbb{R}^+$. It is an analogous task to the derivation of expression (3.1) to compute the Morozova–Chentsov function for the metric (4.2), and we get

$$c(\lambda, \mu) = \frac{(g(\lambda) - g(\mu))(g^*(\lambda) - g^*(\mu))}{(\lambda - \mu)^2}. \tag{4.4}$$

From the property $c(t\lambda, t\mu) = t^{-1}c(\lambda, \mu)$ we deduce that, under the condition $g(0)g^*(0) = 0$, $g(t\lambda)g^*(t\lambda) = tg(\lambda)g^*(\lambda)$ must hold. This implies that

$$g(x)g^*(x) = cx \qquad (x \in \mathbb{R}^+). \tag{4.5}$$

Another necessary condition comes from the property that $\lim_{\lambda \to \mu} c(\lambda, \mu) = \mu^{-1}$. In this way, we arrive at the condition

$$g'(x)g^{*\prime}(x) = x^{-1} \qquad (x > 0). \tag{4.6}$$

(4.5) and (4.6) together have the solution $g(x) = ax^p$ and $g^*(x) = bx^{1-p}$, $ab = c = \frac{1}{p(1-p)}$, and the possible limit $\lim_{p \to 0, \text{ or } 1}$ allowing x and $\log x$.

Note that the operator-monotonicity condition fixes the set of the possible values of p as $-1 \le p \le 2$ (or, $|\alpha| \le 3$). From the present standpoint, we can say that the reason of the absence of the metric for the *symmetric logarithmic-derivative* and other those lying in the gap in *Fig. 1* as a WYD type is because the above duality is lacking in these metrics. In a later publication we shall give a more detailed account of this dual structure.

References

[1] H. HASEGAWA, Non-commutative extension of the information geometry, in *Proc. Intern. Workshop on Quantum Communications and Measurement*, Nottingham 1994, eds. V.P. Belavkin, O. Hirota and R.L. Hudson, Plenum, 1995.

[2] D. PETZ and CS. SUDÁR, Geometry of quantum states, *J. Math. Phys.*, 37 (1996), 2662–2673.

[3] M. OHYA and D. PETZ, *Quantum Entroppy and Its Use*, Springer-Verlag, Heidelberg, 1993.

[4] E.P. WIGNER and M.M. YANASE, Information content of distributions, *Proc. Nat. Acad. Sci.*, USA 49 (1963), 910–918.

[5] E.P. WIGNER and M.M. YANASE, On the positive semidefinite nature of a certain matrix expression, *Canad. J. Math.*, 16 (1964), 397–406.

[6] E.H. LIEB, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Math.* **11** (1973), 267–288.

[7] S. AMARI, Differential-Geometrical Methods in Statistics, *Lecture Notes in Statistics*, 28 (1985).

[8] H. HASEGAWA, α-divergence of the non-commutative information geometry, *Rep. Math. Phys.*, 33 (1933), 87–93.

[9] D. PETZ, Geometry of canonical corretation on the state space of a quantum system, *J. Math. Phys.*, 35 (1994), 780–795.

[10] H. Hasegawa, Exponential and mixture families in quantum statistics, *Rep. Math. Phys.*, 39 (1997), 49–68.

[11] D. Petz, Monotone metrices on matrix spaces, *Linear Algebra Appl.*, 244 (1996), 81–96.

[12] E.A. Morozova and N.N. Chentsov, Markov invariant geometry on state manifold (in Russia), *Itogi Naukini Tekhniki*, 36 (1990), 69–102.

[13] F. Kubo and T. Ando, Means of positive linear operators, *Math Ann.*, 246 (1980) 205–224.

[14] D. Petz and H. Hasegawa, On the Riemannian metric of α-entropies of density matrices, *Lett. Math. Phys.*, 38 (1996), 221–255.

[15] F. Hansen and G.K. Pedersen, Jensen's inequality for operators and Löwner's theorem, *Math. Ann.*, 258 (1982), 229–241.