Proceedings of the 9th WSEAS International Conference on Applied Mathematics, Istanbul, Turkey, May 27-29, 2006 (pp227-235)

Noncommutative Analysis: Application to Quantum Information Geometry II. Uniqueness Theorem

HIROSHI HASEGAWA
Institute of Quantum Science, College of Science and Technology
Nihon University,Chiyoda-ku,Tokyo 101-8308,
JAPAN
h-hase@mxj.mesh.ne.jp

Abstract: The quasi-entropy $S_g(\rho,\sigma) = \mathrm{Tr}\rho^{1/2}g(L_\sigma R_\rho^{-1})\rho^{1/2}$ introduced by Petz is used as a class of contrast functionals of two invertible density matrices ρ and σ in finite state spaces to show the uniqueness of the Wigner-Yanase-Dyson metric and the dual α-connections with respect to it: namely, by expanding $S_g(\rho,\sigma)$ in the difference between ρ and σ to show that, with Fréchet derivatives D_ρ and D_σ up to third order on an asymmetric convex operator function $g(x) \in \mathcal{G}_{asym}$ for which $S_g(\rho,\sigma) \neq S_g(\sigma,\rho)$,

$$D_\rho^2 S_g(\rho,\sigma)(A,B)\Big|_{\sigma=\rho} = -D_\sigma D_\rho S_g(\rho,\sigma)(A,B)\Big|_{\sigma=\rho} = g''(1)K^{(g)}(A,B),$$

$$-D_\rho^2 D_\sigma S_g(\rho,\sigma)(A,C,B)\Big|_{\sigma=\rho} = g''(1)\Gamma^{(g)}(A,C,B) \text{ with its dual form } \rho \leftrightarrow \sigma; \Gamma^{(g)} \leftrightarrow \Gamma^{(g)*}$$

hold, if and only if the resulting metric $K^{(g)}$ is identical to one of the WYD metrics $K^{WYD(\alpha)}$; and the dual connection $\Gamma^{(g)} = \Gamma^{(\pm\alpha)}$ with α being specified by $|\alpha| = 3 + 2\frac{g'''(1)}{g''(1)}$.

Key–Words: g-divergence, convex set, extremal point, Wigner-Yanase-Dyson metrics, dual connection, selfdual vs non-selfdual quasi-entropy.

1 Introduction

One of the major results in classical information geometry initiated by Chentsov[3] and reformulated by Amari[4] is the uniqueness of the Fisher information metric and a pair of dual connections with respect to this metric: it selects these two objects among all those derivable by differentiations up to third order of invariant divergence functionals of two probabilities $D(p,q)$ (a functional $D(p,q)$ is invariant by a transformation of the random variable x in p,q, if $D(p_1,q_1) = D(p,q)$ for $p \mapsto p_1, q \mapsto q_1$[4]). Namely,

"the Fisher information K^F is the only metric introduced by any invariant g-divergence $D_g(p,q)$ with unspecified convex function g, and the $\pm\alpha$-connections $\Gamma^{(\pm\alpha)}$ are the only connections introduced by the same divergence, where α is given by the following identity"

$$|\alpha| = 3 + 2\frac{g'''(1)}{g''(1)} : \tag{1.1}$$

a concise statement of a theorem by Amari, who represented it in two identities, as

$$\textbf{(1)} \quad \partial_i\partial_j D_g(p(\theta),p(\theta'))\Big|_{\theta'=\theta} = -\partial_i\partial_j' D_g(p(\theta),p(\theta'))\Big|_{\theta'=\theta} = g''(1)K_{ij}^F(\theta)$$

$$\textbf{(2)} \quad -\partial_i\partial_j\partial_k' D_g(p(\theta),p(\theta'))\Big|_{\theta'=\theta} = g''(1)\Gamma_{ij,k}^{(\alpha)};$$

$$\Gamma_{ij,k}^{(-\alpha)} \quad \text{by interchanging} \partial_i \leftrightarrow \partial_i' \text{ etc..}$$

In the above identities, $\partial_i \equiv \frac{\partial}{\partial\theta^i}, \partial_i' \equiv \frac{\partial}{\partial\theta'^i}$ etc. for the parameter sets $\{\theta^i\}$ and $\{\theta'^i\}$, defining a given statistical model; we abbreviate $p(x,\{\theta^i\})$ by $p(\theta)$. The implication of **(1)** is that either 1st or 2nd term of the left side in **(1)** is equal to $g''(1)K_{ij}^F(\theta)$, and that of **(2)** to represent the coefficient of connection $\Gamma_{ij,k}^{(\alpha)}(\theta)$ which indicates the duality $\Gamma_{ij,k}^{(-\alpha)}(\theta) = \Gamma_{ij,k}^{(\alpha)*}(\theta)$, when each pair of primed and unprimed derivatives are interchanged before setting $\theta' = \theta$.

The present paper is a continuation of the preceding article,**Noncommutative Analysis: Application to Quantum Information Geometry**, referred as **I** hereafter. Our purpose here is to extend the above uniqueness theorem to a non-commutative(quantum) framework of information geometry. We show that, if the underlying statistical manifolds are generalized to finite quantum states (i.e. finite dimensional matrix manifolds), the uniqueness theorem above holds with three modifications in equalities **(1)** and **(2)**:

(a) the g-divergence $D_g(p,q)$ is replaced by the quasi-entropy $S_g(\rho,\sigma)$ of Petz(1986)[12] for finite quantum states, with p, q being replaced by two invertible density matrices ρ, σ, respectively, and the g-function in $S_g(\rho,\sigma)$ is operator convex, which is assumed to belong to *asymmetric class*; $g^{dual}(x)(= xg(x^{-1})) \neq g(x), x \in \mathcal{M}_h^+$[cf. **I**]

(b) each ∂_i is interpreted properly to be the respective Fréchet partial derivative [cf. **I**]

(c) The Fisher metric $K_{ij}^F(\theta)$ expressed in **(1)** is replaced by the class of Wigner-Yanase-Dyson(WYD) metrics[7][8] $K^{WYD(\alpha)}(\theta)$ with α being specified in the same way as eq.(1.1) i.e. $|\alpha| = 3 + 2\frac{g'''(1)}{g''(1)}$ which is shown to ensure $|\alpha| \leq 3$ automatically.

With these modifications, we verify that **(1)** and **(2)** are true if and only if the metric is identical to one of the WYD metrics $K^{WYD(\alpha)}(\theta)$ apart from a normalization constant $g''(1)$, and $\Gamma^{(\pm\alpha)}$ are a pair of torsionless dual connections with respect to $K^{WYD(\alpha)}$.

In **I**, we presented a feature of quantum information geometry on the basis of the Fréchet differential formula (D) for matrix-valued function $\varphi(\rho)$ of a hermitian variable ρ in terms of the sum of its commutative part and the commutator part, and discussed information metrics and relative entropies both having monotone decrease by stochastic maps. Here we specialize to the question of uniqueness of the *Wigner-Yanase-Dyson metric* and the α *quasi-entropy* which enables us to achieve the above-stated result. From **I**, we can foresee two obstacles to be overcome which do not exist in the classical theory[4]: the limited allowed range of the WYD metrics, and the distinction of asymmetric vs symmetric convex functions.

Before going, it is necessary to add one more comment on the prerequisite for the convex g-functions, either classical or quantum, to ensure the specification of the α-value in eq.(1), implied by the stated uniqueness theorem. That is, the function g in a pertinent convex set must be an *extremal point* of this set, because otherwise it cannot exclude the possibility of convex combinations of g's with different α. So, we first discuss several preliminaries in **Sec.2** and **3**. In **Sec.4** we present two characterization theorems in detail. Then, we proceed to the statement of the result shown in *abstract* and its parametrized scheme.

2 Preliminary 1: Convex functions and their convex set

2.1 Basic matters (some elementary facts in functional analysis)

Definition 2.1 A real, continuous function $f(X)$;

$$X \in \mathcal{M}_n^h \mapsto \mathcal{M}_n^h \text{ is operator monotone of order } n,$$
$$\text{if } X \leq Y \ (Y - X \geq 0) \text{ implies } f(X) \leq f(Y).$$

If f is operator monotone of order n for all $n \in I$, f is said just operator monotone [1].

Definition 2.2 A real, continuous function $g(X); X \in \mathcal{M}_n^h \mapsto \mathcal{M}_n^h$ is operator convex of order n, if for all $n \times n$ hermitians X and Y and for all real numbers $0 \leq \lambda \leq 1$, the following inequality holds:

$$g(\lambda X + (1-\lambda)Y) \leq \lambda g(X) + (1-\lambda)g(Y). \quad (2.1)$$

If g is operator convex of all orders, g is said just oprator convex[1]. We consider the set of operator monotone functions and operator convex functions which are denoted by $\mathcal{F}$ and $\mathcal{G}$, respectively. The concept of operator monotonicity and operator convexity on matrix functions are closed by taking linear combinations with positive coefficients, and pointwise limits in sequences: namely, consider a subset $\{f_n\} \in \mathcal{F}\, or \{g_n\} \in \mathcal{G}; n = 1, 2, ..$; and a set of positive numbers $\{a_n \geq 0\}$. Then,

$$f_1,\ f_2 \quad or\ g_1,\ g_2 \quad \text{and}\, a_i \geq 0\, (i = 1, 2) \Rightarrow$$
$$a_1 f_1 + a_2 f_1 \in \mathcal{F};\ or\ a_1 g_1 + a_2 g_2 \in \mathcal{G}. \quad (2.2)$$

In particlar, the case $\sum_i^n a_i = 1$ is called a convex combination. Also, consider a sequence $\{f_n; n = 1, 2, ..; f_n \in \mathcal{F}\}$ and $\{g_n; g_n \in \mathcal{G}\}$, which are assumed to converge pointwise to f and g, respectively. Then, $\lim_{n\to\infty} f_n \in \mathcal{F}$, and $\lim_{n\to\infty} g_n \in \mathcal{G}$ [1].

2.2 power functions

It is known that power functions of the form x^p; $(x \in \mathbf{R}^+)$ $p \in \mathbf{R}$ yield primitive examples of operator monotone/convex function for which the range of index p is limited: this is specified in a list of the respective set of linearly independent power functions as follows.

list of operator power functions

$$(m)\ -x^p \quad (-1 \le p \le 0); \quad x^p \quad (0 \le p \le 1);$$

$$-x^p, \ (1 \le p \le 2) \ \text{(operator monotone)} \qquad (2.3)$$

$$(c)\ x^p \quad (-1 \le p \le 0); \quad -x^p \quad (0 \le p \le 1);$$

$$x^p \quad (1 \le p \le 2) \quad \text{(operator convex)}. \qquad (2.4)$$

In **I**, we have introduced a specific operator convex function $g(x)$ for the purpose to define a generalized relative entropy[14](we call it "quasi-entropy" following Petz[12]) denoted by $S_g(\rho, \sigma)$: it satisfies $g(1) = 0$. Let us denote such a convex function in the class (2.4) by $g_p(x)$: it may be of the form $g_p(x) = c_p(1 - x^p)$, where c_p is defind properly as

$$\mathcal{G}_{power} = \{g_p(x)\}; \quad g_p(x) = \frac{1 - x^p}{p(1-p)}$$

$$g(1) = 0; \ \text{and}\, g''(1) = 1 \text{ for normalization.} \qquad (2.5)$$

This is a subset of all continuous functions $\mathbf{R}^+ \mapsto \mathbf{R}$; $\mathcal{C}(\mathbf{R}^+)$, and we consider its convex subsets $\mathcal{G}$ such that $\mathcal{G}_{power} \subset \mathcal{G} \subset \mathcal{C}(\mathbf{R}^+)$. We follow two definitions in functional analysis[16].

Definition 2.3 extremal element of a convex set $\mathcal{G}$: it is an element $\in \mathcal{G}$ which cannot be expressed as $\lambda g_1 + (1 - \lambda) g_2$; $g_1\, g_2 \in \mathcal{G}$ for any $\lambda; 0 < \lambda < 1$.

Definition 2.4 convex hull of a subset V: $\mathcal{G}_V = \{\sum_{i=1}^m a_i v_i; a_i \ge 0; \sum_i a_i = 1\}$.
A convex hull of a subspace V of a vecor space is identified to be the minimum convex subset that contains V, and we consider the convex hull of $\mathcal{G}_{power}$, eq.(2.5), which is dennoted by $\mathcal{C}(\mathcal{G}_{power})$ so that $\mathcal{G}_{power} \subset \mathcal{C}(\mathcal{G}_{power}) \subset \mathcal{C}\mathbf{R}^+$ holds. Then, we have

Lemma 2.1
Each function $g_p(x) \in \mathcal{G}_{power}$ is an extremal element of $\mathcal{C}(\mathcal{G}_{power})$ i.e. for any real number λ; $0 < \lambda < 1$, $g_p \ne \lambda g^{(1)} + (1 - \lambda) g^{(2)}$; for any $g^{(1)}, g^{(2)} \in \mathcal{G}_{power}$ and $g^{(1)} \ne g^{(2)}$.

proof. Let $g^{(1)} = \sum_i a_i^{(1)} g_{p_i}$ and $g^{(2)} = \sum_i a_i^{(2)} g_{p_i}$; $0 \le a_i^{(1or2)} \le 1$, and consider the equality $g_p = \lambda g^{(1)} + (1 - \lambda) g^{(2)}$. The right-hand side of this equality is the sum of terms $(\lambda a_i^{(1)} + (1 - \lambda) a_i^{(2)}) g_{p_i}$ so that each coefficient is of the form $\lambda a^{(1)} + (1-\lambda) a^{(2)} < a^{(1)} or a^{(2)}$, depending on $a^{(1)} > a^{(2)}$ *or vice versa*, and < 1. Therefore, the coefficient of the term g_p on the right-hand side is nonnegative and smaller than 1. Thus, this term, if nonzero, can be moved to the left-hand side to rearrange the equality as $g_p = \sum_i b_i g_{p_i}$. However, this equality implies that g_p is linearly dependent on all other g_p' s, which contradicts the linear independence of the set (2.4). This means that $g_p \ne \lambda g^{(1)} + (1 - \lambda) g^{(2)}$.
end of proof.

2.3 Symmetric and asymmetric class of convex functions

In **I**, we have introduced two classes of operator convex functions $\mathcal{G}_{sym} = \{g; g(x) = g^{dual}(x)\}$ and

$\mathcal{G}_{asym} = \{g; g(x) \ne g^{dual}(x)\}$, where besides continuity and $g(1) = 0$ for $g(x)$,

$$g^{dual}(x) \equiv x g(x^{-1}) \quad (g(x) = (g^{dual})^{dual}(x) \qquad (2.6)$$

exemplified by $g_p^{dual}(x) = g_{1-p}(x))$.

It should be remarked that the distinction of the two classes $\mathcal{G}_{sym}, \mathcal{G}_{asym}$ is indispenssable to get the proper solution of the uniqueness problem which is absent in the classical theory[4].

Definition 2.3: equivalent class of dual pairs (Gibilisco and Isola[16]) $(\varphi(x), \chi(x))$
is called a dual pair, if $\chi(x) = \varphi^{dual}(x)$. In general, for a given dual pair $(\varphi_0(x), \chi_0(x))$, any pair $(\varphi = A\varphi_0 + B, \chi = C\chi_0 + D)(x)$ with $AC = 1$ can be a dual pair to yield the same metric form by each Fréchet differentiation.

3 Preliminary 2: Second and third Fréchet derivatives

$$D_X^2 \varphi(X)(A, A) =$$

$$[[\varphi(X), \Delta(A)], \Delta(A)] - D_X \varphi(X) \left([[X, \Delta_A], \Delta_A]\right) \qquad (D_2)$$

$$\begin{aligned} D_X^3 \varphi(X)(A, A, A) = {} & [[[\varphi(X), \Delta_A], \Delta_A], \Delta_A] \\ & -3 D_X^2 \varphi(X)([X, [X, \Delta_A]], [X, \Delta_A]) \\ & -D_X([[[X, \Delta_A], \Delta_A], \Delta_A]). \end{aligned} \qquad (D_3)$$

These formulas are given by Bhatia and Sinha, who discussed a general construction of the relations between derivations and Fréchet derivatives[17].

3.1 Another derivation of Lesniewski-Ruskai formula (Theorem 5.3 in I)

By taking an inner product of the left and right sides of (D_2) with an arbitrary function $h(X)$, we have

$$\mathrm{Tr}D_X^2\varphi(X)\chi(X) = \mathrm{Tr}[\varphi(X),\Delta_A][\chi(X),\Delta_A] + \mathrm{Tr}[X,\Delta_A][\varphi'(X)\chi(X),\Delta_A]. \quad (3.1)$$

Lemma 3.1 *In the tangent space $\mathcal{T}_x\mathcal{M}$ at a fixed X, the equality between the inner product of $D_X^2\varphi$ and χ and the Riemannian metric form of Morozova-Chentzov and Petz type i.e.*

$$\mathrm{Tr}D_X^2(A,B)\chi(X)(= \langle D_X^2\varphi(X)(A,B),\chi(X)\rangle) = \mathrm{Tr}(\varphi'(X)\chi'(X) + [\varphi(X),\Delta_A][\chi(X),\Delta_B]), \quad (3.2)$$

holds, if and only if the pair of the operator functions $\varphi(X)$ and $\chi(X)$ satisfies $\varphi'(X)\chi(X) = \mathrm{constant}\times 1$.

proof. From eq.(3.1) it can be observed that the desired equality holds, if and only if the second term on the right-hand side vanishes i.e. $[\varphi'(X)\chi(X),\Delta_A]=0$ for all tangent vectors. This is possible if and only if $\varphi'(X)\chi(X) = \mathrm{constant}\times 1$.
end of proof.

Remark 3.1 An operator dual pair $(\varphi(X) = X^p, \chi(X) = X^{1-p})$meets the statement of this lemma. However, this does not imply to ensure that the resulting metric satisfies the monotonicity condition, as it does not specify the range of the p-indices. It only indicates the first equality in the extended Lesniewski-Ruskai formula (5.13) of Theorem 5.3 in **I**. Therefore, not only the WYD metrics but also the power-mean metrics in **I** are shown to satisfy the identity (3.2), where it has been seen that the latter metrics can be represented as a power seris of oprator dual pairs with indices unrestricted to the list (2.3).

3.2 Representation by means of divided differences[1][18]

$$D_X\varphi(X)(A) = \sum_{i,j}\varphi^{[1]}(\lambda_i,\lambda_j)A_{ij}e_{ij}, \quad \text{where}$$

$$\varphi^{[1]}(\lambda,\mu) = \begin{cases} \frac{\varphi(\lambda)-\varphi(\mu)}{\lambda-\mu} & \lambda\neq\mu \\ \varphi'(\lambda) & \lambda=\mu \end{cases} \quad (3.3)$$

$$D_X^2\varphi(X)(A,B) = \sum_{i,j,k}\varphi^{[2]}(\lambda_i,\lambda_j,\lambda_k)A_{ij}A_{jk}e_{ik},$$

$$\text{where } \varphi^{[2]}(\lambda_i,\lambda_j,\lambda_k) = \frac{\varphi^{[1]}(\lambda_i,\lambda_j)-\varphi^{[1]}(\lambda_j,\lambda_k)}{\lambda_i-\lambda_k}$$

$$= \frac{\varphi(\lambda_i)-\varphi(\lambda_j)}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_k)} - \frac{\varphi(\lambda_j)-\varphi(\lambda_k)}{(\lambda_j-\lambda_k)(\lambda_i-\lambda_k)}, \quad \text{and}$$

$$\varphi^{[2]}(\lambda_i,\lambda_i) = \frac{1}{2}\varphi''(\lambda_i),\ \varphi^{[3]}(\lambda_{i,..i,..i}) = \frac{1}{6}\varphi'''(\lambda_i),.. \quad (3.4)$$

3.3 Dual affine connections[4]

Hereafter, we recover density matrices as the non-commutative object variable X to discuss quantum info geometry in third order of derivatives in the tangent space i.e. i.e. *affine connections*: by relaxing the normalization $\mathrm{Tr}\rho = 1$ instead $0<\rho$ and $\mathrm{Tr}\rho<\infty$.

Given a Riemannian metric of Morozova-Chentzov and Petz type[5][6] in terms of a pair of functions $(\varphi(\rho),\chi(\rho))$ be given:

$$K_\rho(A,B) = \mathrm{Tr}D_\rho\varphi(\rho)(A)D_\rho\chi(\rho)(B).$$

Another pair of quantities associated with the above metric form: $(\Gamma_\rho^K(C,A,B),\ \Gamma_\rho^{K*}(C,B,A))$ is called *affine connections* which are *dual* with respect to the metric K_ρ i.e.

$$\Gamma_\rho^K(C,A,B) = \mathrm{Tr}D_\rho^2\varphi(\rho)(C,A)D_\rho\chi(\rho)(B);$$

$$\Gamma_\rho^{K*}(C,B,A) = \mathrm{Tr}D_\rho\varphi(\rho)(A)D_\rho^2\chi(\rho)(C,B). \quad (3.5)$$

Note that the second derivative $D^2\cdot(C,A)$ in ordinary situation is symmetric with respect to the two tangent vectors A and B: this case is called "torsionless". Our task in **Sec.5** is to verify the unique identification of torsionless connections to the $\pm\alpha$ connections with duality.

4 Characterization theorems in dual geometry

4.1 Characterization theorem for the WYD metrics

Theorem 4.1(Hasegawa[9] suplemented by Gibilisco-Isola[15]; an improved version of Theorem 3.3 in **I**)
The Wigner-Yanase-Dyson information is a symmetric monotone metric defined on matrix spaces whose MC function is given in terms of a dual pair $(\varphi(x),\chi(x))$ such that

$$c(\lambda,\mu) = \frac{(\varphi(\lambda)-\varphi(\mu))(\chi(\lambda)-\chi(\mu))}{(\lambda-\mu)^2} \quad (4.1)$$

where the product $\varphi(x)\chi(x)$ satisfies $\varphi(0+)\chi(0+)=0$. $(4.1a)$

Conversely, if a metric is defined on matrix spaces in terms of a pair (φ, χ) of the form $\mathrm{Tr} D_\rho \varphi(\rho)(A) D_\rho \chi(\rho)(B)$ and is monotone in the sense of Moroziva-Chentsov and Petz (Theorem 3.1 in **I***) with property (4.1a). Then, it is identical to one of the WYD metrics.*

proof.

the former part We recall the first-order divided difference representation of the Fréchet derivative (eq.(3.3))

$$D_\rho(\varphi(\rho))(A) = \sum_{i,j} \Big(\frac{\varphi(\lambda_i) - \varphi(\lambda_j)}{\lambda_i - \lambda_j} A_{ij} \Big) e_{ij}, \quad (4.2)$$

and set $\varphi(\rho) = \rho^p$ $(\rho = \sum_i \lambda_i e_{ii})$. Similarly, $\chi(\rho) = \rho^{1-p}$. Then,

$$\mathrm{Tr} D_\rho \rho^p(A) D_\rho \rho^{(1-p)}(B) = \sum_{i,j} c(\lambda_i, \lambda_j) A^*_{ij} B_{ij},$$

with MC function, and Petz function being

$$\text{given by} \quad c(\lambda, \mu) = \frac{(\lambda^p - \mu^p)(\lambda^{1-p} - \mu^{1-p})}{(\lambda - \mu)^2}$$

$$\text{and } f(x) = \frac{p(1-p)(x-1)^2}{(x^p - 1)(x^{1-p} - 1)}, \quad (4.3)$$

respectively. We show an explicit proof of the monotonicity of this metric by using the integral representation of the latter function. Namely,(cf.[9])

$$\frac{1}{f_{WYD(p)}(x)} = \frac{\sin p\pi}{\pi} \int_0^\infty d\lambda \lambda^{p-1} \times$$

$$\int_0^1 ds \int_0^1 dt \frac{1}{x((1-t)t\lambda + (1-s) + (t\lambda + s)}, \quad (4.4)$$

$(0 < p < 1)$

$$= \frac{\sin p\pi}{p\pi} \int_0^\infty d\lambda \lambda^{-|p|} \int_0^1 \int_0^1 \times$$

$$dt \frac{x(1-t) + t}{(x((1-t)t\lambda + (1-s) + (t\lambda + s))^2} \quad (-1 < p < 0). \quad (4.5)$$

which are obtainable from the integral representation of fractional-power functions: namely,

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

$$= \frac{\sin p\pi}{p\pi} \int_0^\infty \frac{\lambda^{-|p|}}{(\lambda + z)^2} d\lambda \ (-1 < p < 0), \quad (4.6)$$

together with additional twice elementary integrations(See [9] for details). The fact that the integrand of eqs.(4.5) and (4.6) is an operator monotone-decreasing function of x enables us to conclude that in the left hand side $f_{WYD(p)}(x)$ is opearator monotone. We note that the special cases of p values at $0, 1$ can be fixed by taking the limits in eq.(4.3): Namely,

$$f_{WYD;(p=0)}(x) = \frac{x-1}{\log x} \quad \text{(BKM metric[23])}. \quad (4.7)$$

The monotonicity of this metric is ensured by the closedness property presented in **Sec.2**.

$$\text{For } p \neq 0, 1, \quad \varphi(x)\chi(x) = (p(1-p))^{-1} x,$$

$$\text{and for } p = 0, or 1, \varphi(x)\chi(x) = x \log x \overset{x \to 0+}{\to} 0.$$

the latter(converse) part If a Riemannian metric on matrix spaces is given of the form $\mathrm{Tr} D_\rho \varphi(\rho)(A) D_\rho \chi(\rho)(B)$ which is monotone in the sense of Morozova-Chentsov and Petz, the MC function (4.1)must satisfy condition

(i) continuity

(ii) Fisher form for $\lambda \to \mu; c(\lambda, \mu) \to c(\lambda) = 1/\lambda$

(iii) (-1) th uniformity $c(t\lambda, t\mu) = t^{-1} c(\lambda, \mu)$. In terms of φ, χ functions $+1$-th uniformity $\varphi(t\lambda)\chi(t\lambda) = t\varphi(\lambda)\chi(\lambda)$.

(i) We assume futher the smoothness(differentiability) of the function $\varphi(x)$ and $\chi(x)$.
(ii) This condition leads to $\varphi'(x)\chi'(x) = \frac{1}{x}$, as $(\varphi(y) - \varphi(x))_{y \to x} = \varphi'(x)$ and same for χ.
(iii) By taking the uniformity parameter t large, it can be seen that a function $h(x) \equiv \varphi(x)\chi(x)$ determines its behavior near $x \simeq 0+$, and we use the condition (4.1a) which states that $\varphi(0+)\chi(0+) = h(0+) = 0$. This implies that $\frac{1}{t} h(tx) = h(x)$ on one hand, and $\frac{1}{tx} h(tx)|_{x \to 0+} = h'(x)$ on the other, so that we obtain $\frac{h'(x)}{h(x)} = \frac{1}{x}$ leading to $h(x) = cx$; with $c (> 0)$, a positive constant.

Accordingly, we have a simultaneous differential equation for $\varphi(x)$ and $\chi(x)$ as

$$\varphi'(x)\chi'(x) = \frac{1}{x}, \quad \varphi(x)\chi(x) = cx$$

$$\Longrightarrow y^2 - \frac{y}{x} + \frac{1}{cx^2} = 0, \quad \text{if we set} \quad y = \frac{\varphi'(x)}{\varphi(x)}. \quad (4.8)$$

Possible solutions for eq.(4.8) are given, via

$y = \frac{1 \pm \sqrt{1-4/c}}{2x}$, by

$$\varphi(x) = x^p \quad \chi(x) = x^{1-p} \quad \text{with}$$

$$p = \frac{1+\sqrt{1-4/c}}{2}; \quad 1-p = \frac{1-\sqrt{1-4/c}}{2}$$

for $c < \infty$, obtaining

$$c = \frac{1}{p(1-p)}; \quad 0 < p, 1-p < 1. \tag{4.9}$$

For $c = \infty$ $(p = 0\, or 1)$,

$$\varphi(x) = x \quad \chi(x) = \log x, \tag{4.10}$$

$$\text{because} \quad \varphi'(x)\chi'(x) = \frac{1}{x}$$

$$\text{and} \quad \varphi'(x) = 1 \Rightarrow \chi'(x) = \frac{1}{x}.$$

However, eq.(4.8) is not the all possibilities to characterize the indices of the set up dual pair $(\varphi(x), \chi(x))$ for the MC function, as the above classification excludes the case $c = p(1-p) < 0$.

We have another possibility of differential equation for $\varphi(x)$ and $\chi(x)$:

$$\varphi'(x)\chi'(x) = \frac{1}{x}, \ \varphi(x)\chi(x) = cx$$

$$\Longrightarrow y^2 + \frac{y}{x} + \frac{1}{cx^2} = 0, \quad \text{if we set} \quad y = \frac{\chi'(x)}{\chi(x)}.$$

This is because $\chi(x) = c(\varphi(x))^{-1}$ yields $\frac{\chi'}{\chi}(x) = -\frac{\varphi'}{\varphi}(x)$. On the other hand, since the first two equalities in eq.(4.8) are totally symmetric with respect to the constituent functions of the pair, φ and χ, the quantity $\frac{\chi'}{\chi}(x) = y$ must also satisfy the same quadratic equation as (4.8) so that $y(= \chi'/\chi$, and $= \varphi'/\varphi)$ should obey both equations simultaneously:

$$(a) y^2 + \frac{y}{x} + \frac{1}{cx^2} = 0; \quad (b) y^2 - \frac{y}{x} + \frac{1}{cx^2} = 0. \tag{4.11}$$

Hence,

$$p = \frac{1+\sqrt{1-4/c}}{2}, \quad \bar{p} = \frac{-1-\sqrt{1-4/c}}{2} = -p, \text{ and}$$
$$1-p = \frac{1-\sqrt{1-4/c}}{2} \quad 1-\bar{p} = \frac{-1+\sqrt{1-4/c}}{2}.$$

The possible range of the indices fixed is given by

$$(a)(c > 0): \ 0 < p < 1; \quad 0 < 1-p < 0;$$

$$(b)(c \leq 0): \ -1 \leq \bar{p} \leq 0; \quad 1 \leq 1-\bar{p} \leq 2. \tag{4.12}$$

end of proof.

Corollary 4.1
In the paired monotone metrics $(\varphi(x), \chi(x))$, the Morozova-Chentsov condition of (-1)th uniformity ((iii) of theorem 3.1 in **I**) *is fullfilled by the p or, $(1-p)$-th uniformity of each function of the pair: namely*

$$\varphi(tx) = t^p \varphi(x), \quad \chi(tx) = t^{1-p}\chi(x);$$

in particular, for the BKM metric

$$\varphi(tx) = t\varphi(x); \chi(tx) = \chi(x) \quad (or,\ vice\ versa). \tag{4.13}$$

Remark 4.1 In [15], Gibilisco and Isola obtained the result of this corollary for case (b) in eq.(4.12) (case of negative c in the above proof) independently of the second differential equation (b), but in a context of *regularly varying functions*.

Corollary 4.2(Gibilisco and Isola[15])
Difference between two cases (a) and (b) can be characterized by the behavior of the dual pair $\varphi(x), \chi(x)$

$$(a)\varphi(0+) = 0 \ \chi(0+) = 0;$$
$$(b)\varphi(0+) = -\infty, \chi(0+) = 0\, or, vice\, versa. \tag{4.14}$$

4.2 Characterization theorem for the quasi-entropy with duality

We begin by recalling the correspondence theorem of Lesniewski-Ruskai[14](Theorem 5.2 in **I**); namely, *There exists one-to-one correspondence between (an operator) monotone decreasing function $k(x)$ and a symmetrized quasi-entropy $S_g(\rho, \sigma) + S_{g^{dual}}(\rho, \sigma)$:*

$$k(x)\Big(= \frac{1}{f(x)}\Big) = \frac{g(x) + g^{dual}(x)}{(x-1)^2}; \quad g(x) \in \mathcal{G}_{asym}. \tag{4.15}$$

However, it does not state the one-to-one correspondence between the monotone-metric function and the asymmetric convex function $g(\neq g^{dual})$ associated with S_g. This question is of a special interest from the viewpoint whether the power index p that specifies the WYD metric also specifies uniquely the asymmetric convex function $g_\alpha(x)$, eq.(5.9) in **I**. The affirmative answer to this question is now stated, which is revised from the context in [9].

Theorem 4.2(revised version of Hasegawa[9] Theorem 4.1)
Let a real, continuous convex function $g(x)$ which belongs to $\mathcal{G}_{asym}$ be given. Then, its Lesniewski-Ruskai correspondece equation (4.15) provides the WYD metric k-function, if and only if the pair $(g(x), g^{dual}(x))$ is identical to an equivalent class which satisfies $g(1) = 0$ of one of the dual pairs (x^p, x^{1-p}); or $(-\log x, x\log x)$ (this itself satisfies

$g(1) = 0$).

proof.
the "if part" For $p \neq 0, 1$ the dual pair $(c_p x^p, x^{1-p})$ has an equivalent class $(c_p(1-x^p), (1-x^{1-p}))$ with $c_p = \frac{1}{p(1-p)}$, and their product yields

$$k_{WYD(p\neq 0,1)}(x) = \frac{c_p(1+x-x^p-x^{1-p})}{(x-1)^2} = \frac{c_p(x^p-1)(x^{1-p}-1)}{(x-1)^2},$$

$$k_{WYD(p=0,1)}(x) = \frac{-\log x + x\log x}{(x-1)^2} = \frac{\log x}{x-1} \quad (\text{ BKM metric}). \tag{4.16}$$

the "only if part" We may assume a pair $(\varphi(x), \chi(x))$ in the representation of the MC function $c(\lambda, \mu)$ in eq.(4.1), and ask under what circumstance it agrees with the equivalent class of the power dual pair (x^p, x^{1-p}), if the convex g-function is written in the form

$$g(x) = \frac{c(1-\varphi(x))(1-\chi(x))}{(1-x)^2} \in \mathcal{G}_{asym}, \tag{4.17}$$

where the dual pair $(\varphi(x), \chi(x))$ is conditioned, besides $\varphi(1) = \chi(1) = 1$, by

$$i)\ +1\text{th uniformity } \varphi(tx)\chi(tx) = t\varphi(x)\chi(x);$$
$$ii)\ \text{duality} \quad \chi(x) = \varphi^{dual}(x) = x\varphi(x^{-1}). \tag{4.18}$$

By Corollary 4.1, we may also require that the uniformity for $\varphi(x)$ i.e. with some parameter p $i')$ $\varphi(tx) = t^p\varphi(x)$. Accordingly, we now need to show that, under requirements in eq.(4.18) with i) strengthened by $i')$, the dual pair (φ, χ) is identified to be (x^p, x^{1-p}). The numerator of expression (4.17) for $g(x)$ can be rewritten as $1+\varphi(x)\chi(x)-\varphi(x)-\chi(x) = 1 + x\varphi(x)\varphi(x^{-1}) - \varphi(x) - x\varphi(x^{-1})$, and noting that $1^{dual} = x$; $x^{dual} = 1$, we see that, in order for eq.(4.17) to agree with the Lesniewski-Ruskai formula (4.15), it is necessary and sufficient that the functional equation $\varphi(x)\varphi(x^{-1}) = 1$ holds. We verify the following statement. Namely, consider a real, continuous function $\varphi(x); x \in \mathbf{R}^+$ with property

$$i')\ \ p\text{-th uniformity} \quad \varphi(tx) = t^p\varphi(x);$$
$$ii)\ \ \varphi(x)\varphi(x^{-1}) = 1. \tag{4.19}$$

The only function equipped with this property is the power function $\varphi_p(x) = x^p$, because from ii) $\varphi(1)^2 = 1 \Rightarrow \varphi(1) = \pm 1$ which is inserted into $i')$ to yield $\varphi(t) = \pm t^p$: we take the $+$ convention and replace t by x. The operator convexity requires $p \in [-1, 2]$.
end of proof.

Corollary 4.3
In the gap region between the Bures metric and the minimum of the WYD metric ($p = 1/2$) which is filled by the power-mean metrics, no member of $\mathcal{G}_{asym}$ with duality exists.

5 Uniqueness Theorem

5.1 On the uniqueness of the Lesniewski-Ruskai asymmetric g-functions

Theorem 5.1
The only convex operator functions $g(x) \in \mathcal{G}_{asym}$ are those to define quantum α-divergence[13] (identical to those defined in the classical version [4] except the forbiddenness $|\alpha| > 3$) *i.e.*

$$g_\alpha(x) = \frac{4}{1-\alpha^2}(1 - x^{\frac{1+\alpha}{2}});\ \alpha \neq \pm 1,\ |\alpha| \leq 3,\ \textit{and}$$
$$\begin{cases} x\log x & \alpha = 1 \\ -\log x & \alpha = -1 \end{cases} \quad \textit{all} \in \mathcal{G}_{asym}. \tag{5.1}$$

proof. This is obvious, since by Corollary 4.3 no WYD metric f-functions are located in the gap region, as can be visualized in Fig.1 in **I**.
end of proof.

5.2 Uniqueness of the dual $\pm\alpha$ connections

In [9], we made an argument on the uniqueness of the dual affine connections associated with the WYD metrics under the assumption that the extent of monotone metrics on matrix spaces are paired metrics: this was not sufficient. The result is expected to hold, when the underlying metrics are extended to general monotone metrics. (See related papers on the uniqueness of dual connections[22][23][24].) Here, we show that this is the case. To avoide an unnecessary complication, we assume that second Fréchet derivatives are torsionless.

Theorem 5.2
The only monotone metric on matrix spaces(in the sense of Morozova-Chentsov and Petz) *with respect to which each of the two connections derivedby a first, and a second Fréchet differential, are dual to each other is the WYD-metric parametrized by* $\pm\alpha$; $0 < |\alpha| \leq 3$. ($\alpha = 0$ is excluded from the dual-connction viewpoint.)

First, we show a proposition from Theorem 5.3 in **I**, also Lemma 2.1 and 2.2 as follows.

Proposition 5.3
Every monotone metric defined on (finite dimensional) matrix spaces can be expressed as a sum, either finite or infinite, of trace funcions of the form

$$\sum_{n\geq 0} C_n \mathrm{Tr}[\rho^{pn}, \Delta_A][\rho^{1-pn}, \Delta_B \rho^{1-pn}] \qquad (5.2)$$

which can be classified into two cases as follows.

Case A. finite sum the sum belongs to the convex hull of $\mathcal{G}_p$, which is identified to be a convex combination of the WYD($\alpha \neq 0$) metrics, and is classified into *non-selfdual class*; $S(\sigma, \rho) \neq S(\rho, \sigma)$, where the dual connection is possible.

Case B. infinite sum the infinite part belongs to *self-dual class* $S(\rho, \sigma) = S(\sigma, \rho)$: the only connection possible for this is the metric connection [4].

The BKM metric(WYD metric with $\alpha = \pm 1$) for which the dual connection is possible[23] may be classified in **A**, as a point-wise limit of the WYDα.

proof of Theorem 5.2 By definition of 5.1 in **I**, a non-selfdual quasi-entropy $S(\rho, \sigma)$ is associated with an asymmetric convex function $g \neq g^{dual}$, and Theorem 5.1 tells us that it is identified to be one of α-divergence with $0 < |\alpha| \leq 3$. **end of proof.**

(Details of Proposition 5.3 will be given in a paper of a WSEAS journal.)

6 Concluding Remarks: Amari form of uniqueness theorem

nonparametric form

(1) $D_\rho^2 S_g(\rho, \sigma)(A, B)\Big|_{\sigma=\rho}$

$= -D_\sigma D_\rho S_g(\rho, \sigma)(A, B)\Big|_{\sigma=\rho} = g''(1) K^{(g)}(A, B),$

(2) $- D_\rho^2 D_\sigma S_g(\rho, \sigma)(A, C, B)\Big|_{\sigma=\rho} = g''(1) \times$
$\Gamma^{(g)}(A, C, B)$ with its dual form
$\rho \leftrightarrow \sigma; \Gamma^{(g)}(A, C, B) \leftrightarrow \Gamma^{(g)*}(A, C, B).$

A satisfaction of equalities (1) and (2) by an unspecified operator convex function g is possible, if and only if it is selected from the convex hull in the sense of **Sec.2.1** *i.e.* $\mathcal{C}(\mathcal{G}_{asym})$, *restricted to its extremal element, which identifies itself with one of quantum α-divergence and the resulting metric with the corresponding Wigner-Yanase-Dyson metric.*

Remark 6.1 on eq.(1.1)

$$|\alpha| = 3 + 2\frac{g'''(1)}{g''(1)}.$$

This equality is valid also in the present non-commutative version of information geometry: it stems from the fact that every Fréchet derivative $D, D^2, D^3, ..$ contains the first commutative part denoted by $D^c, (D^c)^2, (D^c)^3, ..,$ each reduces to the ordinary derivative by the projection onto the commutative part of the tangent space. We can write the expansion $S_g(\rho, \rho + d\rho)$ up to $O(d\rho)^3$ as[4]

$$S_g(\rho, \rho{+}d\rho) = \frac{1}{2} K^g(\rho, \rho)(d^c\rho)^2 {+} \frac{1}{6} H^g(\rho)(d^c\rho)^3 {+}..,$$

where for a dual pair φ, χ, the above expansion can be written explicitly (cf. eq.(3.4)), as

$$g'''(1)\rho^{-2} = -g''(1)\left(1 + \frac{1 - \pm\alpha}{2}\right)\rho^{-2}, \qquad (6.1)$$

from which $\pm\alpha = 3 + 2\frac{g'''(1)}{g''(1)}$ can be deduced. Then, the only noncommutativ/commutative difference in g is the operator convexity of g, or equivalently the operator monotonicity of $-g$, which is known to be *completely monotone* [1] with ocsillating successive derivatives i.e. $g'(1) \leq 0, g''(1) \geq 0, g'''(1) \leq 0, ...$ Thus, it ensures the inequality $|\alpha| \leq 3$.

parametrized form based on Sec.4.2 in I

(1) $\partial_i \partial_j S_g(\rho(\theta), \rho(\theta'))\Big|_{\theta'=\theta}$
$= -\partial_i \partial'_j S_g(\rho(\theta), \rho(\theta'))\Big|_{\theta'=\theta} = g''(1) K_{ij}^F(\theta)$

(2) $- \partial_i \partial_j \partial'_k S_g(\rho(\theta), \rho(\theta'))\Big|_{\theta'=\theta}$
$= g''(1)\Gamma_{ij,k}^{(\alpha)};\ \Gamma_{ij,k}^{(-\alpha)}$(by interchange $\partial_i \leftrightarrow \partial'_i$etc.),

where Fréchet partial derivative is defined by

$$\partial_i \varphi(\rho(\theta))(A) = \frac{\partial^c \varphi}{\partial \theta^i} A_i^c + [\varphi(\rho), \Delta_{A_i}].$$

$K(A, B)_{ij} =$
$\langle A_i^c \rho^{-1} \partial_i \rho \partial_j \rho B_j^c \rangle + \mathrm{Tr}([\varphi(\rho), \Delta_{A_i}][\chi(\rho), \Delta_{B_j}]).$

In conclusion, we wish to point out that the uniqueness of the α-divergence contributes to the same question for the relative entropy in Tsallis statistics[19]-[21].

Acknowledgements The author thanks F.Hansen for useful comments on the author's publication[9].

References (through **I** and **II**)

1. R.Bhatia, *Matrix Annalysis*(Springer Verlag, Berlin,1997).
2. T.M.Flett, *Differential Analysis*(Cambridge University Press, 1980).
3. N.N.Chentsov,*Statistical Decision Rules and Optional Inference* Translations of Mathematical Monograph**53**, AMS, Providence, R.I. (1982); the original paper in Russia (1972).
4. S.Amari, *Differential-geometrical methods in statistics*, Lecture Notes in Statistics, Vol**28** Springer, Berlin (1985); S.Amari and H. Nagaoka,*Methods of Information Geometry*,AMS monograph**191**(2000).
5. E.A.Morozova and N.N.Chentsov, *Markov invariant geometry on manifolds of states*,Itogi NaukiniTeckhniki,**36**(1990)69-102;English translate,J.Soviet Math.**56**(1991)2648-2669.
6. D.Petz,*Monotone metrics on matrix spaces*, Linear Alg.Appl.**244**(1996)81-96.
7. E.P.Wigner M.M.Yanase,*Information content of distributions*,Proc.Nat.Acad.Sci.USA**49**(1963) 910-918.
8. E.H.Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Advances in Math.**11**(1973)267-288.
9. H.Hasegawa, *Dual geometry of the Wigner-Yanase-Dyson information content*, Infinite Dimensional Analysis, Quantum Probability and Related Topics(IDAQP) **6**(2003)413-430.
10. F.Kubo and T.Ando,*means of positive linear operators*,Math. Ann.**246**(1980)205-224.
11. M.K.Murray and J.W.Rice, *Differential Geometry and Statistics*, Monograph on Statistics and Applied Probability**48**, Chapman and Hall(1993).
12. D. Petz,*Quasi-entropies for finite quantum systems*, Rep.Math.Phys.**23**(1986) 57-65.
13. H.Hasegawa,*α-divergence of the noncommutative information geometry*, Rep.Math.Phys. **33**(1993)87-93.
14. A.Lesniewski and M.B.Ruskai,*Monotone Riemannian metrics and relative entropy on noncommutative probability spaces*,J.Math. Phys.**40**(1999),5702-5723.
15. P.Gibilisco and T.Isola, *On the characterization of paired monotone metrics*, Annals of the Institute of Statistical Mathematics**56**(2004) 369-381.
16. N.Dunford and J.T.Schwartz,*Linear Operators*, PART I,Interscience Publisher, New York, London: Wiley(1958).
17. R.Bhatia and K.B.Sinha,*Derivations,derivatives and chain rules*,Linear Alg.Appl.**302-303** (1999)231-244.
18. F.Hansen,*Operator convex functions of several variables*, Pub. RIMS, Kyoto Univ.**33**(1997)443-463.
19. C.Tsallis,*Possible generalization of Boltzmann statistics*,J.Stat.Phys.**52**(1988)479-487.
20. S.Abe,*q-deformed entropies and Fisher metrics* Proc.5th Wigner Symposium, Vienna(1997)(World Scientific,1998)
21. S.Abe *Nonadditive generalization of Kullback Leibller divergence for measuring the degree of purification*.Phys.Rev.**68**(2003)032302.
22. P.Gibilisco and T.Isola,*Connections on statistical manifolds of density operators by geometry of noncommutative L^p-spaces*, Infinite Dimensional Analysis, Quantum Probability and Related Topics(IDAQP)**2**(1999) 169-178.
23. M.Grasselli and R.F.Streater, *On the uniqueness of the Chentsov metric in quantum information geometry*, Infinite Dim. Anal.Prob.Related Topics**4**(2001) 172-182.
24. A.Jenčová, *Geometry of quantum states: dual connections and divergence functions*, Rep. Math. Phys.**47**(2001)121-138.
25. D.Petz and Jenčová, *On quantum Fisher information*,J.Elec.Eng.**50**(1999). (This paper contains a brief proof of additivity for quantum Fisher metrics.)
26. H.Hasegawa,*Quantum Fisher Information and q-deformed relative entropies-additivity vs nonadditivity*, Suppl.Prog.Theor.Phys.(2006), to be published.
27. A.Jenčová, unpublished note communicated to H.Hasegawa(2004).