

Quantum Information Metrics and Relative Entropies

– Classification Problem –

Hiroshi Hasegawa¹

Institute of Quantum Science, College of Science and Technology,

Nihon University, Chiyoda-ku, Tokyo 101-8308, Japan

対称性と正值写像に対する単調減少性を備えたリーマン距離としての量子情報距離と量子相対エントロピーとの関係を述べ、それに基き Wigner-Yanase-Dyson 情報量の普遍性について論ずる。

Quantum information metric is a Riemannian metric defined on matrix spaces which measures an infinitesimal distance between two nearby density matrices ρ and $\rho + d\rho$ whose classical correspondent is the so-called Fisher information metric. In the classical version, this concept is universal among all those having monotonicity against coarse-graining (Chentsov 1982), and is derived by twice differentiations of an entropy functional, called α -divergence (Amari 1985), such that

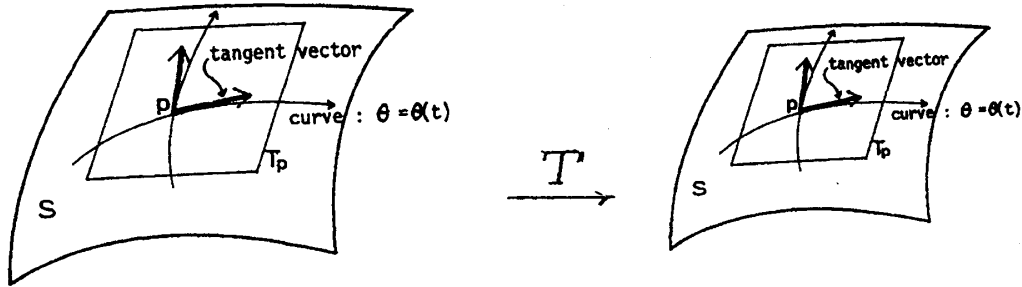
$$K_{ij}^F(\theta) = \int \partial_i \log p(x, \theta) \partial_j \log p(x, \theta) p(x, \theta) d\mu(x) = \langle \partial_i \log p(\theta) \partial_j \log p(\theta) \rangle = -\partial_i \partial_j S_{g_\alpha}(p(\theta), p(\theta'))|_{\theta'=\theta}, \quad \text{where} \quad (1)$$

$$S_{g_\alpha}(p, q) \equiv \frac{4}{1-\alpha^2} \int \left(1 - q(x, \theta)^{\frac{1+\alpha}{2}} p(x, \theta)^{\frac{-(1+\alpha)}{2}} \right) p(x, \theta) d\mu(x) \quad (2)$$

(α -divergence for two probabilities $p(x, \{\theta_i\}), q(x, \{\theta_i\})$).

Thus, the said universality is evidenced by the fact that K^F is α -independent. Unlike the classical context, quantum information metrics with monotonicity are not universal in the above sense, and we aim to get a relevant universality with an aid of relative or quasi-entropy (Petz, 1985).

Fig.1 Image of a monotone Riemannian metric.



contraction of the metric form by any coarse-graining(positive map T)

¹E-mail: h-hase@mxj.mesh.ne.jp

1 Outline of Results– with set of metrics and quasi-entropies

set of all Petz’s symmetric monotone functions (1996) $f \in \mathcal{F} \supset \mathcal{F}_{power}$ and $\mathcal{F}_{WYD}$

(“power”denotes power-mean; and “WYD” Wigner-Yanase-Dyson, for monotone metrics)

Result(1): $\mathcal{F}_{power} \cup \mathcal{F}_{WYD}$ is dense in $\mathcal{F}$ by a properly defined norm topology.

set of all convex g -functions Petz(1985) $g \in \mathcal{G} \supset \mathcal{G}_{sym}, \mathcal{G}_{asym}$ ($\mathcal{G}_{sym}, \mathcal{G}_{asym}$ with and without satisfying condition of symmetry $g(x) = g^{dual}(x) \equiv xg(x^{-1})$ and $g(x) \neq g^{dual}(x)$, respectively). Correspondence between $\mathcal{G}$ and $\mathcal{F}$ in **Sec.3**(Lesniewski-Ruskai(1999)) yields

Result(2): there exists no g -function in $\mathcal{G}_{asym}$ other than quantum α -divergence: it is the only source of WYD metrics that possess the potential duality of geometry.

Historical remark: the Wigner-Yanase-Dyson metrics

The original definition of skew information was $-\frac{1}{2}\text{Tr}[\rho^{1/2}, k]^2$ (Wigner-Yanase in Proc. Nat. Acad. Sc., 1963), to which Dyson suggested that it could be generalized to the one exponent ϑ and the other $1 - \vartheta$ for the power of ρ . Lieb(1973), substantiating their convexity for the first time, expressed them as

$$I_p(\rho, \Delta) \equiv \frac{1}{p(1-p)} \text{Tr}[\rho^p, \Delta][\rho^{1-p}, \Delta] \quad (\Delta \in \{\text{skew hermitians}\}, i\Delta \in \mathcal{M}_h)$$

$$= \frac{4}{1-\alpha^2} D_\sigma^\perp D_\rho^\perp \sigma^{\frac{1+\alpha}{2}}(A) \rho^{\frac{1-\alpha}{2}}(A) |_{\sigma=\rho} \text{ with } D_\rho^\perp(\varphi(\rho))(A) \equiv [\varphi(\rho), \Delta_A] \quad p \equiv \frac{1+\alpha}{2}.$$

This adds to the classical Fisher term to yield a full metric form in terms of a Fréchet differential denoted by D_ρ : a Fréchet derivation on a matrix function $\varphi(\rho)$ is written as $D_\rho \varphi(\rho)(A)$ with A being an arbitrary hermitian except $\text{Tr} A = 0$ a matrix tangent vector, illustration in Fig.1 with an important relation $A = [\varphi(\rho), \Delta_A]$ (Δ_A only depends on A), leading us to

$$\text{Tr} A^{\dagger} g^F A^c + I_p(\rho, \Delta_A) (g^F = \rho^{-1}) = \text{Tr} D_\rho(p^{-1} \rho^p)(A) D_\rho((1-p)^{-1} \rho^{1-p})(A). \quad (3)$$

The quantum correspondent to the Fisher metric has an additional term, as indicated above in eq.(3), which could be derived from the corresponding quantum α -divergence. Namely,

$$S_\alpha(\rho, \rho + \rho) \text{ (to 2nd order in } d\rho \text{)} = \frac{1}{2} \text{Tr} \rho d^c \log \rho d^c \log \rho + \frac{4}{1-\alpha^2} \text{Tr} [\rho^{\frac{1+\alpha}{2}}, \Delta] [\rho^{\frac{1-\alpha}{2}}, \Delta], \quad (4)$$

where symbol d^c represents the derivative of ρ which commutes with ρ : it means that an infinitesimal increment of a noncommutative variable should consist in a commutative and a noncommutative parts with ρ (former yielding the uniform Fisher term, whereas the latter nonuniform)

$$d\rho = d^c \rho + [\rho, \Delta] \quad [d^c \rho, \rho] = 0 \quad \text{and } \Delta \in \{\text{skew hermitians}\} \text{ or } i\Delta \in \mathcal{M}_h.$$

This may be regarded as a Hilbert-Schmidt orthogonal decomposition, since $\text{Tr} d^c \rho [\rho, \Delta] = 0$ holds. It is then easy to see the following expression $d\varphi(\rho)$ for any analytic function of ρ $\varphi(\rho) = \frac{1}{2\pi i} \oint_C \frac{dz}{z-\rho}$; quantum analog of tangent vector at the point ρ as illustrated in Fig.1 i.e.

$$d\rho = d^c \rho + [\rho, \Delta]; \quad d\varphi(\rho) = \varphi'(\rho) d^c \rho + [\varphi(\rho), \Delta].$$

2 General monotone metrics(extended Chentsov theorem)

for including Bures metric and for interpolating it with the WYD by the power-mean metrics:

$$K_\rho(A, B) \equiv \text{Tr} A K_\rho B \quad A, B \in \mathcal{M}_h, \text{ and for symmetric-metric cases}$$

$$K'_\rho(B, A) \equiv \frac{1}{2}(K_\rho(B, A) + K_\rho(A, B)) = K'_\rho(A, B); \text{ redenoted by } K_\rho(A, B).$$

By taking the ρ -diagonalized basis, it can be expressed in terms of a two-variable function $c(\lambda, \mu)$ on $\mathbf{R}^+ \times \mathbf{R}^+$ and a single variable one $c(\lambda) \equiv c(\lambda, \lambda)$ on $\mathbf{R}^+$ in the above bilinear form:

$$K_\rho(B, A) = \sum_i c(\lambda_i) B_{ii} A_{ii} + 2 \sum_{i < j} c(\lambda_i, \lambda_j) B_{ij}^* A_{ij} \quad \text{with}$$

$$c(x\lambda, x\mu) = x^{-1} c(\lambda, \mu) \quad (\text{homogeneity of MC-functions of order -1}), \text{ in particular } c(\lambda, \lambda) = \lambda^{-1}.$$

Petz's representation of symmetric monotone metrics by f :

$K_\rho(B, A) = \langle A, R_\rho^{-1/2} f(L_\rho R_\rho^{-1})^{-1} R_\rho^{-1/2} B \rangle$, where $f(x)$ and $c(\lambda, \mu)$ are related to each other:

$$f(x) = \frac{1}{c(x, 1)} \quad \text{or} \quad c(\lambda, \mu) = \frac{1}{\mu f(\lambda/\mu)}. \quad (5)$$

Representation of symmetry $c(\mu, \lambda) = c(\lambda, \mu) \Leftrightarrow x f(1/x) = f(x)$. Then, $c(\lambda = \mu) = 1/\lambda \rightarrow f(1) = 1$. In classical information geometry, the g -divergence function conditioned by $g(1) = 0$ and $g''(x) > 0$, ($= 0$ only for $x = 1$), was shown to be entitled all the way to yield α -divergence with any real number α ($|\alpha| = 3 + g'''(1)/g''(1)$ Amari, 1986). One of remarkable outcomes of the extended Chentsov theorem is the limited order structure of the monotone metrics so that

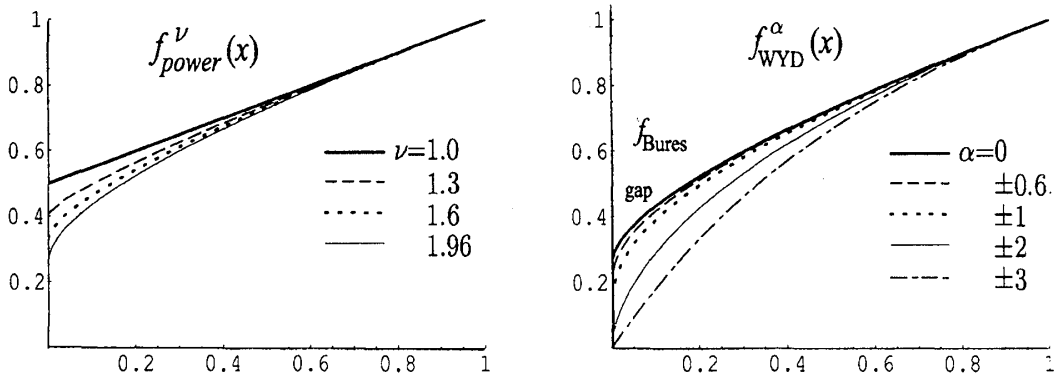
$$(\text{Bures}) \frac{1+x}{2} = f_{\max}(x) \geq f_{\text{power}}(x) \geq f_{\text{WYD}}(x) \geq f_{\min}(x) = \frac{2x}{1+x} (\text{WYD}^{|\alpha|=3}). \quad (6)$$

Specific feature of quantum information metrics There appears a region which may be called “gap region” between $f_{\max} = f_{\text{Bures}}$ and $f_{\text{WYD}}^{\alpha=0}$. This is absent in the classical manifold.

Fig.2 Order structure of monotone metrics on matrix spaces in terms of f -functions.

$$\begin{array}{cc} \text{the “gap” region} & \text{the WYD region} \\ \bar{A} & \\ f_{\text{power}}^\nu(x) = \frac{1+x^{1/\nu}}{2} & 1 \leq \nu \leq 2; \quad f_{\text{WYD}}^\alpha(x) = \frac{1-\alpha^2}{4} \frac{(1-x)^2}{(1-x^{\frac{1-\alpha}{2}})(1-x^{\frac{1+\alpha}{2}})} \quad |\alpha| \leq 3. \end{array} \quad (7)$$

($\mathcal{F}_{\text{power}} \cup \mathcal{F}_{\text{WYD}}$ forms a linear-ordered subset of $\mathcal{F}$ with coincident min-max bounds)



3 Classification of quasi-entropies by symmetry

We follow Lesniewski-Ruskai(J. Math. Phys.**40**(1999)5702-5723) to represent a convex operator function $g(x)$ in $S_g(\rho, \sigma) = \text{Tr}(\sigma - \rho)g^{(2)}(L_\sigma R_\rho^{-1})R_\rho^{-1}(\sigma - \rho)$; $g^{(2)}(x) \equiv \frac{g(x)}{(x-1)^2}$, where $g(x) = b(x-1)^2 + c \frac{(x-1)^2}{x} + \int_0^\infty \frac{(x-1)^2}{x+s} dm(s)$, two constants $b, c \geq 0$ and a measure $m \geq 0$. (8)

We define, for a $g(x)$, $g^{dual}(x) \equiv xg(x^{-1})$. Then, $((x-1)^2)^{dual} = (x-1)^2/x$ and vice versa. In general, $g^{dual}(x) = c(x-1)^2 + b \frac{(x-1)^2}{x} + \int_0^\infty \frac{(x-1)^2}{x+s} d\bar{m}(s)$, where $\bar{m}(s) \equiv sm(s^{-1})$. (9)

Let $\mathcal{G}_{sym}$ and $\mathcal{G}_{asym}$ denote the set of all quasi-entropy g -functions of x , defined for symmetric and asymmetric class, respectively, by $\mathcal{G}_{sym} = \{g; g(x) = g^{dual}(x)\}$ and $\mathcal{G}_{asym} = \{g; g(x) \neq g^{dual}(x)\}$.

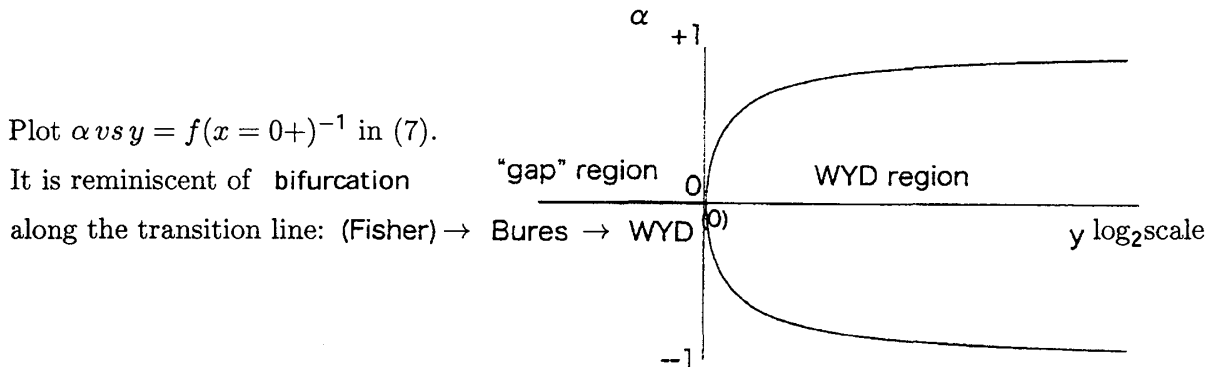
(1) Selfdual class: $S_{s,dual}(\rho, \sigma) = S_{s,dual}(\sigma, \rho)$ with a symmetric $g \in \mathcal{G}_{sym}$. In terms of the measure representations above, both $b = c$ and $\bar{m}(s) = m(s)$ hold.

(2) Non-selfdual class: $S_{ns,dual}(\rho, \sigma) \neq S_{ns,dual}(\sigma, \rho)$ with an asymmetric $g \in \mathcal{G}_{asym}$. Then, Lesniewski-Ruskai relation between Petz's $f(x)$ and $g(x)$ holds: $\frac{1}{f(x)} = \frac{g(x) + g^{dual}(x)}{2(x-1)^2}$. (10)

4 Concluding remarks

4.1 transition from the classical Fisher to quantum information metrics

Fig.3 Schematic picture: appearance of the quantum structure of information metrics.



More studies are necessary for succeeding in this idea.

4.2 universality of the WYD metrics endowed by an inclusion of topology

Lesniewski-Ruskai relation (10) says that, if a symmetric monotone metric f is identical to a member of the WYD class, then the associated quasi-entropy g -function must be equal to the $\pm\alpha$ -divergence $g_{\pm\alpha}[1]$. However, it does not assure that $g_{\pm\alpha}$'s exhaust all members of $\mathcal{G}_{asym}$. We have investigated this question, introducing a norm topology by exploiting the symmetry of f , which makes it equivalent to the space $C[0,1]$ (set of all continuous functions on $[0,1]$). Consequently, the question can be answered affirmatively. Details will be given elsewhere.

All citations in the present article are given in

[1] H. Hasegawa, Infinite Dimensional Analysis, Quantum Probability and Related Topics **6** (2003), 413-430; 数理科学「ランダム行列理論と非可換情報幾何」長谷川 洋 (現代物理と現代幾何 2002), 105-114 および「量子情報理論と幾何」古池達彦 (**488** 2004), 15-21 サイエンス社.