On the characterisation of paired monotone metrics

Paolo Gibilisco[1]

Dipartimento di Scienze, Facoltà di Economia, Università di Chieti-Pescara "G. D'Annunzio", Viale Pindaro 42, I–65127 Pescara, Italy.

Tommaso Isola [2]

Dipartimento di Matematica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, I–00133 Roma, Italy.

Abstract

Hasegawa and Petz introduced the notion of paired monotone metrics. They also gave a characterisation theorem showing that Wigner-Yanase-Dyson metrics are the only members of the paired family. In this paper we show that the characterisation theorem holds true under hypotheses that are more general than those used in the above quoted references.

Key words and phrases: Monotone metrics, Wigner-Yanase-Dyson information.

1 Introduction

Monotone metrics are the quantum counterpart of Fisher information and are classified by Petz (1996,2002). The Wigner-Yanase-Dyson information content (see Lieb (1973), Wigner and Yanase (1963))

$$I_p(\rho, A) = -\mathrm{Tr}([\rho^p, A][\rho^{1-p}, A])$$

can be seen as a one-parameter family of monotone metrics, see Hasegawa and Petz (1997). There, Hasegawa and Petz gave a proof that the WYD-metrics are the only monotone metrics possessing a certain pairing property (in Hasegawa (2003) Hasegawa discusses how this reflects on the associated relative entropy along the lines of Lesniewski and Ruskai (1999)). This is substantially related to the pairing of the non-commutative versions of Amari embeddings

$$\rho \to \frac{\rho^p}{p} \qquad \rho \to \log(\rho).$$

The purpose of the present paper is to present a partially different proof of the characterisation theorem. In Hasegawa and Petz (1997) and Hasegawa (2003) a certain boundary behaviour is used as an hypothesis.

Here we show that the characterisation theorem holds true under more general conditions, that is without the above hypothesis (see the Remark 5.1). While we use means that are relatively less elementary (the theory of regularly varying functions) it seems that the present proof also fills some gaps appearing in the arguments of Hasegawa and Petz (1997) and Hasegawa (2003). It should be emphasized that the pairing discussed here is related to the duality of non-commutative α-connections as discussed in many papers (see Nagaoka (1995), Hasegawa (1995), Gibilisco and Isola (1999), Amari and Nagaoka (2000), Grasselli and Streater (2001), Jenčova (2001), Grasselli (2002)).

Another goal of this paper is to relate the above pairing to the duality of uniformly convex Banach spaces according to the lines of our previous works Gibilisco and Pistone (1998), Gibilisco and Isola (1999, 2001a,b):

[1]Electronic mail: gibilisc@sci.unich.it

[2]Electronic mail: isola@mat.uniroma2.it

this appears, up to now, as one of the the main tools for the infinite dimensional approach to Information Geometry.

The structure of the paper is as follows. In section 2 we present the notion of pull-back of duality pairing and discuss the case of commutative Amari embeddings. In section 3 we review the theory of monotone metrics and their pairing. In section 4 one finds the basic results on regularly varying functions that are needed in the sequel. Section 5 contains the proof of the characterisation theorem.

2 Pull-back of duality pairings

Let V, W be vector spaces over $\mathbb{R}$ (or $\mathbb{C}$). One says that there is a duality pairing if there exists a separating bilinear form

$$\langle \cdot, \cdot \rangle : V \times W \to \mathbb{R}.$$

Let $\mathcal{M}, \mathcal{N}$ be differentiable manifolds. A differentiable function $\varphi : \mathcal{M} \to \mathcal{N}$ is an immersion if its differential $D_\rho \varphi : T_\rho \mathcal{M} \to T_{\varphi(\rho)} \mathcal{N}$ is injective, for any $\rho \in \mathcal{M}$.

Definition 2.1. Suppose we have a pair of immersions (φ, χ), where $\varphi : \mathcal{M} \to \mathcal{N}$ and $\chi : \mathcal{M} \to \tilde{\mathcal{N}}$, such that a duality pairing exists between $T_{\varphi(\rho)}\mathcal{N}$ and $T_{\chi(\rho)}\tilde{\mathcal{N}}$ for any $\rho \in \mathcal{M}$. Then we may pull-back this pairing on $\mathcal{M}$ defining

$$\langle u, v \rangle_\rho^{\varphi,\chi} := \langle D_\rho \varphi(u), D_\rho \chi(u) \rangle \qquad u, v \in T_\rho \mathcal{M}.$$

The most elementary example is given by the case where $\mathcal{N} = \tilde{\mathcal{N}}$ is a riemannian manifold, $\varphi = \chi$ and the duality pairing is just given by the riemannian scalar product on $T_{\varphi(\rho)}\mathcal{M}$. This is called the pull-back metric induced by the map φ.

A first non-trivial example is the following. Let X be a uniformly convex Banach space such that the dual $\tilde{X}$ is uniformly convex. We denote by $\langle \cdot, \cdot \rangle$ the standard duality pairing between X and $\tilde{X}$. Let $J : X \to \tilde{X}$ be the duality mapping, that is J is the differential of the map $v \to \frac{1}{2}||v||^2$ (see Berger (1977), p. 373). This implies that $J(v)$ is the unique element of the dual such that

$$\langle v, J(v) \rangle = ||v||^2 = ||J(v)||^2.$$

Definition 2.2. Let $\mathcal{M}$ be a manifold. If we have a map $\varphi : \mathcal{M} \to X$ we can consider a dualised pull-back that is a bilinear form defined on the tangent space of $\mathcal{M}$ by

$$\langle A, B \rangle_\rho^\varphi := \langle A, B \rangle_\rho^{\varphi, J\circ\varphi} = \langle D_\rho \varphi(A), D_\rho (J \circ \varphi)(B) \rangle.$$

Remark 2.1. For X a Hilbert space, J is the identity, and this is again the definition of pull-back metric induced by the map φ.

Example 2.1. Let $(X, \mathcal{F}, \mu)$ be a measure space. If f is a measurable function and $q \in (1, +\infty)$ then $||f||_q := (\int |f|^q)^{\frac{1}{q}}$. Moreover $\tilde{q}$ is defined by $\frac{1}{q} + \frac{1}{\tilde{q}} = 1$. Set

$$L^q = L^q(X, \mathcal{F}, \mu) = \{f \text{ is measurable} |\ ||f||_q < \infty\}$$

Define N^q as L^q with the norm

$$||f||_{N^q} := \frac{||f||_q}{q}.$$

Obviously $\widetilde{N^q}$ (the dual of N^q) can be identified with $N^{\tilde{q}}$. Indeed if $f \in N^q$ and $g \in N^{\tilde{q}}$ define

$$T_g(f) := \int \frac{f}{q}\frac{g}{\tilde{q}}$$

One has

$$||T_g|| = \sup\frac{|T_g(f)|}{||f||_{N^q}} = \sup\frac{\int \frac{f}{q}\frac{g}{\tilde{q}}}{\frac{||f||_q}{q}} = \frac{1}{\tilde{q}}\sup\frac{\int fg}{||f||_q} = \frac{||g||_{\tilde{q}}}{\tilde{q}} = ||g||_{N^{\tilde{q}}}$$

from this easily follows that $g \to T_g$ is an isometric isomorphism between $\widetilde{N^q}$ and $N^{\tilde{q}}$. Now suppose that $\rho > 0$ is measurable and $\int \rho = 1$, namely ρ is a strictly positive density. Then $v = q\rho^{\frac{1}{q}}$ is an element of the unit sphere of N^q and it is easy to see that $J(v) = \tilde{q}\rho^{\frac{1}{\tilde{q}}}$. The family of maps $\rho \to q\rho^{\frac{1}{q}}$ are just the Amari embeddings.

Let $X = \{1, ..., n\}$ and let μ be the counting measure. In this case N^q is just $\mathbb{R}^n$ with the norm $\frac{||\cdot||_q}{q}$. Let $\mathcal{P}_n = \{v \in \mathbb{R}^n | v_i > 0, \sum v_i = 1\}$.

Proposition 2.1. *Consider the Amari embedding $\varphi : \rho \in \mathcal{P}_n \to q\rho^{\frac{1}{q}} \in N^q$ for an arbitrary $q \in (1, +\infty)$. Then the bilinear form*

$$\langle A, B\rangle_\rho^\varphi := \langle A, B\rangle_\rho^{\varphi, J\circ\varphi} = \langle D_\rho\varphi(A), D_\rho(J\circ\varphi)(B)\rangle \qquad A, B \in T_\rho\mathcal{P}_n$$

is just the Fisher information.

Proof.

$$\langle D_\rho\varphi(A), D_\rho(J\circ\varphi)(B)\rangle = \int (\rho^{\frac{1}{q}-1}A)(\rho^{\frac{1}{\tilde{q}}-1}B) = \int \frac{AB}{\rho}$$

□

The above result can be stated in much greater generality using the machinery of Pistone and Sempi (1995), Gibilisco and Isola (1999).

3 Paired monotone metrics

In the commutative case a Markov morphism is a stochastic map $T : \mathbb{R}^n \to \mathbb{R}^k$. In the noncommutative case a Markov morphism is a completely positive and trace preserving operator $T : M_n \to M_k$, where M_n denotes the space of n by n complex matrices. We shall denote by $\mathcal{D}_n$ the manifold of strictly positive elements of M_n and by $\mathcal{D}_n^1 \subset \mathcal{D}_n$ the submanifold of density matrices.

In the commutative case a monotone metric is a family of riemannian metrics $g = \{g^n\}$ on $\{\mathcal{P}_n\}$, $n \in \mathbb{N}$ such that

$$g^m_{T(\rho)}(TX, TX) \leq g^n_\rho(X, X)$$

holds for every stochastic map $T : \mathbb{R}^n \to \mathbb{R}^m$ and all $\rho \in \mathcal{P}_n$ and $X \in T_\rho\mathcal{P}_n$.

In perfect analogy, a monotone metric in the noncommutative case is a family of riemannian metrics $g = \{g^n\}$ on $\{\mathcal{D}_n^1\}$, $n \in \mathbb{N}$ such that

$$g^m_{T(\rho)}(TX, TX) \leq g^n_\rho(X, X)$$

holds for every stochastic map $T : M_n \to M_m$ and all $\rho \in \mathcal{D}_n^1$ and $X \in T_\rho \mathcal{D}_n^1$ (see Chentsov and Morotzova (1990)).

Let us recall that a function $f : (0, \infty) \to \mathbb{R}$ is called operator monotone if for any $n \in \mathbb{N}$, any $A, B \in M_n$ such that $0 \le A \le B$, the inequalities $0 \le f(A) \le f(B)$ hold. An operator monotone function is said symmetric if $f(x) := xf(x^{-1})$ and normalised if $f(1) = 1$. With such operator monotone functions f one associates the so-called Chentsov–Morotzova functions

$$c_f(x, y) := \frac{1}{yf(xy^{-1})} \qquad \text{for} \qquad x, y > 0.$$

Proposition 3.1. *For a CM-function the following is true*

(i) $c(tx, ty) = \frac{1}{t} c(x, y) \qquad \forall x, y, t > 0$

(ii) $c(x) := \lim_{y \to x} c(x, y) = \frac{1}{x}$.

Define $L_\rho(A) := \rho A$, and $R_\rho(A) := A\rho$. Since L_ρ and R_ρ commute we may define $c(L_\rho, R_\rho)$. Now we can state the fundamental theorems about monotone metrics. In what follows uniqueness and classification are stated up to scalars.

Theorem 3.1. *(Chentsov 1982) There exists a unique monotone metric on $\mathcal{P}_n$ given by the Fisher information.*

Theorem 3.2. *(Petz 1996) There exists a bijective correspondence between monotone metrics on M_n and normalised symmetric operator monotone functions. This correspondence is given by the formula*

$$g_\rho^f(A, B) := Tr(A \cdot c_f(L_\rho, R_\rho)(B)).$$

The tangent space to $\mathcal{D}_n^1$ at ρ is given by $T_\rho \mathcal{D}_n^1 \equiv \{A \in M_n : A = A^*, Tr(A) = 0\}$, and can be decomposed as $T_\rho \mathcal{D}_n^1 = (T_\rho \mathcal{D}_n^1)^c \oplus (T_\rho \mathcal{D}_n^1)^o$, where $(T_\rho \mathcal{D}_n^1)^c := \{A \in T_\rho \mathcal{D}_n^1 : [A, \rho] = 0\}$, and $(T_\rho \mathcal{D}_n^1)^o$ is the orthogonal complement of $(T_\rho \mathcal{D}_n^1)^c$, with respect to the Hilbert-Schmidt scalar product $\langle A, B \rangle_{HS} := Tr(A^* B)$. A typical element of $(T_\rho D_n)^o$ has the form $A = i[\rho, U]$, where U is self-adjoint. Each statistically monotone metric has a unique expression (up to a constant) given by $Tr(\rho^{-1} A^2)$, for $A \in (T_\rho \mathcal{D}_n^1)^c$.

Proposition 3.2. *(See Bhatia 1997) Let $A \in T_\rho \mathcal{D}_n^1$ be decomposed as $A = A^c + i[\rho, U]$ where $A^c \in (T_\rho \mathcal{D}_n^1)^c$ and $i[\rho, U] \in (T_\rho \mathcal{D}_n^1)^o$. Suppose $\varphi \in \mathcal{C}^1(0, +\infty)$. Then*

$$(D_\rho \varphi)(A) = \varphi'(\rho) A^c + i[\varphi(\rho), U].$$

Let $\varphi, \chi \in \mathcal{C}^1(0, +\infty)$. Using the functional calculus one may consider (φ, χ) as a pair of functions from $\mathcal{D}_n^1$ to M_n, for which a duality pairing is provided by the Hilbert Schmidt scalar product $\langle \cdot, \cdot \rangle_{HS}$. Therefore, according to the previous section, we can define the paired metric induced by (φ, χ) as

$$\langle A, B \rangle_\rho^{\varphi, \chi} = \mathrm{Tr}(D_\rho \varphi(A) \cdot D_\rho \chi(B)), \qquad A, B \in T_\rho \mathcal{D}_n^1.$$

Proposition 3.3. *(Hasegawa and Petz (1997), Hasegawa (2003)) Let f be operator monotone, $c = c_f$ the associated CM-function. For a pair $\varphi, \chi \in \mathcal{C}^1(0, +\infty)$, the equality*

$$\langle A, B \rangle_\rho^{\varphi, \chi} = Tr(A \cdot c(L_\rho, R_\rho)(B)).$$

implies

$$c(x, y) = \frac{\varphi(x) - \varphi(y)}{x - y} \cdot \frac{\chi(x) - \chi(y)}{x - y}. \tag{3.1}$$

Proof. It is enough to consider elements of $(T_\rho D_n)^o$. Suppose $A = i[\rho, U]$ and $B = i[\rho, V]$ where U, V are self-adjoint. Using Proposition 3.2 one has $D_\rho \varphi(A) = i[\varphi(\rho), U]$ and similarly for B. Therefore

$$\langle A, B\rangle_\rho^{\varphi,\chi} = \langle D_\rho\varphi(A), D_\rho\chi(B)\rangle = \langle i[\varphi(\rho), U], i[\chi(\rho), V]\rangle = \langle i\hat{\varphi}(L_\rho, R_\rho)[\rho, U], i\hat{\chi}(L_\rho, R_\rho)[\rho, V]\rangle,$$

where $\hat{\varphi}(x, y) := \frac{\varphi(x)-\varphi(y)}{x-y}$, and similarly for $\hat{\chi}$. On the other hand it is true that

$$\mathrm{Tr}(A \cdot c(L_\rho, R_\rho)(B)) = \mathrm{Tr}(i[\rho, U]c(L_\rho, R_\rho)(i[\rho, V])).$$

From the above equations and the arbitrariness of A, B one has the conclusion. □

Definition 3.1. In the hypotheses of Proposition 3.3, we say that $\langle\cdot,\cdot\rangle_\rho^{\varphi,\chi}$ is a *paired monotone metric.* Moreover we set

$$\mathfrak{P} := \{(\varphi, \chi)|\varphi, \chi \in \mathcal{C}^1(0, +\infty) \text{ and } \varphi, \chi \text{ induce a paired monotone metric}\}$$

In what follows we give examples of elements of $\mathfrak{P}$.

Definition 3.2.

$$f_p(x) := p(1-p)\frac{(x-1)^2}{(x^p-1)(x^{1-p}-1)} \qquad p \in \mathbb{R} \setminus \{0, 1\}$$

$$f_0(x) = f_1(x) := \frac{x-1}{\log(x)}.$$

Obviously $f_p = f_{1-p}$ and

$$f_0 = \lim_{p\to 0} f_p = \lim_{p\to 1} f_p = f_1.$$

Moreover we have that f_{-1} is the function of the RLD-metric, $f_0 = f_1$ is the function of the BKM-metric and $f_{\frac{1}{2}}$ is the function of the WY-metric.

Definition 3.3.

$$(\varphi_p(x), \chi_p(x)) = (\frac{x^p}{p}, \frac{x^{1-p}}{1-p}) \qquad p \in \mathbb{R} \setminus \{0, 1\}$$

$$(\varphi_0(x), \chi_0(x)) = (\varphi_1(x), \chi_1(x)) = (x, \log x).$$

Theorem 3.3. Hasegawa and Petz (1997), Hasegawa (2003) (φ_p, χ_p) *induce a paired monotone metric if and only if* $p \in [-1, 2]$.

Proof. The proof consists in showing that the function f_p is operator monotone iff $p \in [-1, 2]$.

After this one has immediately that

$$c_p(x, y) = \frac{1}{y f_p(\frac{x}{y})} = \frac{\varphi_p(x) - \varphi_p(y)}{x-y} \cdot \frac{\chi_p(x) - \chi_p(y)}{x-y}$$

and this ends the proof. □

Now let $p \in (0, 1)$ and set $q = \frac{1}{p}$. We use again the symbol N_q to denote M_n with the norm

$$||A||_{N^q} = q^{-1}(\mathrm{Tr}(|A|^q))^{\frac{1}{q}}$$

All the commutative construction of Example 2.1 goes through. The following Proposition is the non-commutative analogue of Proposition 2.1 (see also Hasegawa and Petz (1997), Jenčova (2001), Gibilisco and Isola (2001b), Grasselli(2002)).

Proposition 3.4. *Let $\varphi : \rho \in \mathcal{D}_n^1 \to q\rho^{\frac{1}{q}} \in N_q$ be the Amari embedding. The dualised pull-back*

$$\langle A, B\rangle_\rho^\varphi := \langle A, B\rangle_\rho^{\varphi, J\circ\varphi} = \langle D_\rho\varphi(A), D_\rho(J\circ\varphi)(B)\rangle$$

coincides with the Wigner-Yanase-Dyson information.

Proof. It is a straightforward application of Proposition 3.3. □

4 Regularly varying functions

For the content of this section see Bingham et al. (1987).

Definition 4.1. Let ℓ be a measurable positive function defined on some neighbourhood $[X, +\infty)$ of infinity and satisfying

$$\lim_{x\to+\infty} \frac{\ell(tx)}{\ell(x)} = 1 \qquad \forall t > 0;$$

then ℓ is said slowly varying.

Remark 4.1. Defining $\ell(x) = \ell(X)$ on $(0, X)$ one often considers ℓ defined on $(0, +\infty)$.

Some examples of slowly varying functions are $\ell(x) = \log(x), \log(\log(x)), \exp(\log(x)/\log(\log(x)))$.

Proposition 4.1. *If ℓ is slowly varying and $p > 0$ then*

$$\lim_{x\to+\infty} x^p\ell(x) = +\infty \qquad\qquad \lim_{x\to+\infty} \frac{\ell(x)}{x^p} = 0.$$

Definition 4.2. A measurable function $h > 0$ satisfying

$$\lim_{x\to+\infty} \frac{h(tx)}{h(x)} = t^p \qquad \forall t > 0$$

is called regularly varying of index p; we write $h \in R_p$. Therefore R_0 is the class of slowly varying functions. We set $R := \cup_{p\in\mathbb{R}} R_p$.

Remark 4.2. Obviously homogeneous functions are very particular cases of regularly varying functions.

Proposition 4.2. *Assume $h > 0$ is a measurable function, and there exists a function j such that*

$$\lim_{x\to+\infty} \frac{h(tx)}{h(x)} = j(t) \in (0, +\infty) \tag{4.1}$$

for all t in a set of positive measure. Then

(i) the equation 4.1 holds for all $t > 0$;

(ii) there exists $p \in \mathbb{R}$ such that $j(t) = t^p$, $\forall t > 0$;

(iii) $h(x) = x^p\ell(x)$ where ℓ is slowly varying.

Sometimes, as in the present paper, one is interested in the behaviour at the origin.

Definition 4.3. If h is a measurable positive function and

$$\lim_{x\to 0^+} \frac{h(tx)}{h(x)} = t^p \qquad \forall t > 0$$

then one says that h is regularly varying at the origin, in symbols $h \in R_p(0^+)$.

Let $\tilde{h}(x) := h(\frac{1}{x})$. Then $h \in R_p(0^+)$ iff $\tilde{h} \in R_{-p}$.

Corollary 4.1. $h \in R_1(0^+) \Longrightarrow \lim_{x\to 0^+} h(x) = 0$.

Proof. $h \in R_1(0^+) \Longrightarrow \tilde{h} \in R_{-1}$. Therefore there exists ℓ slowly varying s.t. $\tilde{h}(x) = x^{-1}\ell(x)$. This implies

$$\lim_{x\to 0^+} h(x) = \lim_{y\to+\infty} h\left(\frac{1}{y}\right) = \lim_{y\to+\infty} \tilde{h}(y) = \lim_{y\to+\infty} \frac{\ell(y)}{y} = 0$$

where the last equality depends on Proposition 4.1. □

5 The main result

Definition 5.1. Two elements of $\mathfrak{P}$, $(\varphi,\chi), (\tilde{\varphi},\tilde{\chi})$ are equivalent if there exist constants A_1, A_2, B_1, B_2 such that $A_1 A_2 = 1$

$$\tilde{\varphi} = A_1\varphi + B_1$$
$$\tilde{\chi} = A_2\chi + B_2.$$

Obviously equivalent elements of $\mathfrak{P}$ define the same CM-function. In what follows we consider elements of $\mathfrak{P}$ up to this equivalence relation with the traditional abuse of language.

Lemma 5.1. *Suppose that* (φ,χ) *induce a paired monotone metric. Then*

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{p}{x} \qquad \Longrightarrow \qquad (\varphi,\chi) = (\varphi_p,\chi_p).$$

Proof.

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{p}{x} \Longrightarrow \log\frac{\varphi(x)}{\varphi(x_0)} = p\log\frac{x}{x_0} \Longrightarrow \varphi(x) = \varphi(x_0)(\frac{x}{x_0})^p = Ax^p.$$

We may choose $\varphi(x) = \frac{x^p}{p}$.

Now let $c(\cdot,\cdot)$ be the associated CM-function. Going to the limit $y \to x$ in equation (3.1) and using Proposition 3.1 one has

$$\varphi'(x)\chi'(x) = c(x) = \frac{1}{x} \tag{5.1}$$

If $p = 1$ then

$$\varphi(x) = x \Longrightarrow \chi'(x) = \frac{1}{x} \Longrightarrow \chi(x) = \log(x)$$

If $p \neq 1$ then

$$\varphi(x) = x^p \Longrightarrow \chi'(x) = \frac{1}{x^p} \Longrightarrow \chi(x) = \frac{x^{1-p}}{1-p}$$

and this ends the proof. □

We are ready to prove the fundamental result of the theory.

Theorem 5.1. (Hasegawa and Petz (1997), Hasegawa (2003)) *Let* $\varphi, \chi \in \mathcal{C}^1(0, +\infty)$. *Then* (φ, χ) *induce a paired monotone metric if and only if one of the following two possibilities holds*

$$(\varphi(x), \chi(x)) = (\frac{x^p}{p}, \frac{x^{1-p}}{1-p}) \qquad p \in [-1, 2] \setminus \{0, 1\}$$

$$(\varphi(x), \chi(x)) = (x, \log(x)).$$

Proof. The "if" part is just Theorem 3.3. To prove the "only if" part we need some auxiliary functions.
Let us define

$$k(x, y) := (\varphi(x) - \varphi(y))(\chi(x) - \chi(y)) = (x - y)^2 c(x, y).$$

One has

$$k(tx, ty) = t^2(x - y)^2 c(tx, ty) = t(x - y)^2 c(x, y) = tk(x, y)$$

that is k is 1-homogeneous. Moreover set $h(x) := \varphi(x)\chi(x)$.

Equation (5.1) implies that φ, χ are strictly monotone (either both increasing or both decreasing) and therefore injective. Moreover monotonicity implies that the following limits exist

$$\varphi(0^+) := \lim_{x \to 0+} \varphi(x) \qquad\qquad \chi(0^+) := \lim_{x \to 0+} \chi(x).$$

Since we consider φ, χ up to additive constants and because we can change the sign of φ, χ, we may reduce to φ, χ increasing, and have to consider three cases

a) $\varphi(0^+) = \chi(0^+) = -\infty$,
b) $\varphi(0^+) = 0, \ \chi(0^+) = -\infty$,
c) $\varphi(0^+) = \chi(0^+) = 0$.

Case a)

Suppose $\varphi(0^+) = \chi(0^+) = -\infty$. Now let $0 < y < x$; going to the limit $y \to 0^+$ we have that

$$\begin{aligned} t = \frac{k(tx, ty)}{k(x, y)} &= \lim_{y \to 0^+} \frac{\varphi(tx) - \varphi(ty)}{\varphi(x) - \varphi(y)} \cdot \frac{\chi(tx) - \chi(ty)}{\chi(x) - \chi(y)} \\ &= \lim_{y \to 0^+} \frac{\varphi(ty)}{\varphi(y)} \cdot \frac{\chi(ty)}{\chi(y)} = \lim_{y \to 0^+} \frac{h(ty)}{h(y)}. \end{aligned}$$

This means that $h \in R_1(0^+)$ and therefore by Corollary 4.1

$$+\infty = \lim_{x \to 0^+} \varphi(x)\chi(x) = \lim_{x \to 0^+} h(x) = 0$$

that is absurd.

Case b)

Suppose $\varphi(0^+) = 0$ and $\chi(0^+) = -\infty$. Again let $0 < y < x$; going to the limit $y \to 0^+$ we have that

$$t = \lim_{y \to 0^+} \frac{\varphi(tx) - \varphi(ty)}{\varphi(x) - \varphi(y)} \cdot \frac{\chi(tx) - \chi(ty)}{\chi(x) - \chi(y)} = \frac{\varphi(tx)}{\varphi(x)} \cdot \lim_{y \to 0^+} \frac{\chi(ty)}{\chi(y)}.$$

This implies that the limit

$$\lim_{y\to 0^+} \frac{\chi(ty)}{\chi(y)}$$

exists $\forall t$. Therefore there exists a function j such that

$$j(t) = \lim_{y\to 0^+} \frac{\chi(ty)}{\chi(y)} = \frac{t\varphi(x)}{\varphi(tx)}.$$

From Proposition 4.2 one has that $-\chi \in R_\beta(0^+)$ for some $\beta \in \mathbb{R}$ namely $j(t) = t^\beta$, $\forall t > 0$. So we have $\varphi(tx) = t^{1-\beta}\varphi(x)$, that is φ is p-homogeneous, with $p := 1-\beta$, and therefore by Euler $x\varphi'(x) = p\varphi(x)$ ($p \neq 0$ because $\varphi' \neq 0$); then

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{p}{x}$$

and therefore because of Lemma 5.1,

$$(\varphi, \chi) = (\varphi_p, \chi_p).$$

Since $\chi(0^+) = -\infty$ we have $p \geq 1$. Because of Theorem 3.3 one has $p \in [1, 2]$.

Case c)

The argument for this case is that of Hasegawa and Petz (1997), Hasegawa (2003) and we report it here for the sake of completeness.

One can deduce

$$\frac{\varphi(tx)-\varphi(ty)}{t(x-y)} \cdot \frac{\chi(tx)-\chi(ty)}{t(x-y)} = \frac{1}{t}\frac{\varphi(x)-\varphi(y)}{x-y} \cdot \frac{\chi(x)-\chi(y)}{x-y}$$

Going to the limit $y \to 0^+$ one has

$$\frac{\varphi(tx)}{tx} \cdot \frac{\chi(tx)}{tx} = \frac{1}{t}\frac{\varphi(x)}{x} \cdot \frac{\chi(x)}{x}$$

so that

$$\varphi(tx)\chi(tx) = t\varphi(x)\chi(x)$$

This means that $h(x) = \varphi(x)\chi(x)$ is 1-homogeneous and $h(0^+) = 0$ so that, because of Euler, one has $xh'(x) = h(x)$. This implies that $\exists b \in \mathbb{R}$ s.t. $h(x) = bx$, $\forall x > 0$. Then

$$\varphi(x)\chi(x) = bx \qquad \varphi'(x)\chi'(x) = \frac{1}{x}.$$

As φ, χ are increasing, $b > 0$. Deriving the first equation one gets

$$\varphi'(x)\chi(x) + \varphi(x)\chi'(x) = b.$$

Since $\varphi, \varphi' \neq 0$, $\forall x > 0$ one may write

$$\chi(x) = \frac{bx}{\varphi(x)} \qquad \chi'(x) = \frac{1}{x\varphi'(x)}.$$

Substituting one gets

$$\frac{\varphi'(x)}{\varphi(x)}bx + \frac{\varphi(x)}{\varphi'(x)}\frac{1}{x} = b,$$

so that if $y(x) := \frac{\varphi'(x)}{\varphi(x)} \neq 0$ the equation becomes

$$bxy(x) + \frac{1}{xy(x)} = b$$

and finally

$$bx^2y(x)^2 - bxy(x) + 1 = 0.$$

From this it follows that
i) if $0 < b < 4$ there is no solution;
ii) if $b \geq 4$ then

$$y(x) = \frac{1}{2x}\left(1 \pm \sqrt{1 - \frac{4}{b}}\right).$$

Therefore, setting

$$p := \frac{1 + \sqrt{1 - \frac{4}{b}}}{2} \in \left[\frac{1}{2}, 1\right),$$

we have

$$\frac{\varphi'(x)}{\varphi(x)} = y(x) = \frac{p}{x}$$

or

$$\frac{\varphi'(x)}{\varphi(x)} = y(x) = \frac{1-p}{x}.$$

From Lemma 5.1 one has

$$\varphi(x) = \frac{x^p}{p} \qquad \chi(x) = \frac{x^{1-p}}{1-p}$$

or viceversa. Therefore $(\varphi, \chi) = (\varphi_p, \chi_p)$, with $0 < p < 1$, and this ends the proof. □

Corollary 5.1. *If φ, χ induce a paired monotone metric then* $\lim_{x\to 0+} \varphi(x)\chi(x) = 0$.

Remark 5.1. In Hasegawa and Petz (1997), Hasegawa (2003) Theorem 5.1 is proved under the hypothesis that $\lim_{x\to 0+} \varphi(x)\chi(x) = 0$. The present proof shows that this hypothesis can be dropped. An application of this is given in Gibilisco and Isola (2003).

6 Concluding remarks

In Hasegawa (2003) Hasegawa wanted to find a family of (non-paired) operator monotone functions that "fill the gap" between the functions

$$f_{\text{Bures}}(x) = \frac{1+x}{2} \qquad f_{\frac{1}{2}}(x) = \frac{(1+\sqrt{x})^2}{4}$$

corresponding to the SLD-metric and the WY-metric. The problem can be solved by proving the following

Proposition 6.1. *The functions*

$$f^{\nu}_{\text{power}}(x) := \left(\frac{1 + x^{\frac{1}{\nu}}}{2} \right)^{\nu} \qquad 1 \le \nu \le 2$$

are operator monotone.

To prove Proposition 6.1 Hasegawa used an argument due to Petz. We just want to remark that the above result can be proved by applying to $f_{\text{Bures}}(x) = \frac{1+x}{2}$ the following

Proposition 6.2. *Let f be operator monotone, and $\nu \in [1, \infty)$. Then $x \in (0, \infty) \to f(x^{1/\nu})^{\nu}$ is operator monotone.*

Proof. See the proof of Corollary 4.3 (i) in Ando (1979). □

For $p \in (0, 1)$ namely $q \in (1, +\infty)$ Proposition 3.4 shows that it is possible to relate the duality discussed here to the geometry of spheres in L^q spaces along the lines of Gibilisco and Pistone (1998), Gibilisco and Isola (1999, 2001a,b). The same does not apply to the cases $p \in [-1, 0]$ or $p = 1$. In Gibilisco and Pistone (1998) the Amari embedding was generalised to the sphere of an Orlicz space under very general hypothesis. We conjecture that, for $p = 0, 1$ (that is for the BKM metric) one can use non-commutative analogues of the Zygmund spaces $L^{\exp}$, $L^{x \log x}$ to produce a similar construction (see also Grasselli and Streater (2000) and references therein).

Acknowledgements

It is a pleasure to thank H.Hasegawa for useful conversations on this subject at the congress "Information Geometry and its Applications", Pescara, Italy, 2002. This research has been supported by the italian MIUR program "Quantum Probability and Infinite Dimensional Analysis" 2001-2002.

References

Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry*, American Mathematical Society and Oxford University Press.

Ando, T. (1979). Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra and its Applications*, **26**, 203–241.

Berger, M. S. (1977). *Nonlinearity and functional analysis.* Academic Press, New York.

Bhatia, R. (1997). *Matrix Analysis.* Springer-Verlag, New York.

Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular variation*, Cambridge University Press, Cambridge.

Chentsov, N. (1982). *Statistical decision rules and optimal inference*, American Mathematical Society, Providence.

Chentsov, N. and Morotzova, E. (1990). Markov invariant geometry on state manifolds (in russian), *Itogi Nauki i Tekhniki*, **36**, 69-102.

Gibilisco, P. and Isola, T. (1999). Connections on statistical manifolds of density operators by geometry of noncommutative L^p-spaces, *Infinite Dimensional Analysis, Quantum Probability & Related Topics*, **2**, 169-178.

Gibilisco, P. and Isola, T. (2001a). Monotone metrics on statistical manifolds of density matrices by geometry of noncommutative L^2-spaces, *Disordered and complex systems*, 129-140, Amer. Inst. Phys., Melville.

Gibilisco, P. and Isola, T. (2001b). A characterisation of Wigner-Yanase skew information among statistically monotone metrics, *Infinite Dimensional Analysis, Quantum Probability & Related Topics*, **4**, 553-557.

Gibilisco, P. and Isola, T. (2003). Wigner-Yanase information on quantum state space: the geometric approach, *Journal of Mathematical Physics*, **44**, 3752-3762.

Gibilisco, P. and Pistone G. (1998). Connections on non-parametric statistical manifolds by Orlicz space geometry, *Infinite Dimensional Analysis, Quantum Probability & Related Topics*, **1**, 325-347.

Grasselli, M.R. (2002). Duality, monotonicity and the Wigner-Yanase-Dyson metrics, preprint 2002, [math-ph/0212022].

Grasselli, M.R. and Streater R.F. (2000). The quantum information manifold for ϵ-boundend forms, *Reports on Mathematical Physics* , **46**, 325-335.

Grasselli, M.R. and Streater R.F. (2001). On the uniqueness of Chentsov metric in quantum information geometry, *Infinite Dimensional Analysis, Quantum Probability & Related Topics*, **4**, 173-182.

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

Hasegawa, H. (1995). Non-commutative extension of the information geometry, *Quantum Communications and Measurement*, 327-337, Plenum Press, New York.

Hasegawa, H. and Petz, D. (1997). Non-commutative extension of information geometry, II, *Quantum Communication, Computing and Measurement*, 109-118, Plenum Press, New York.

Hasegawa, H. (2003). Dual Geometry of the Wigner-Yanase-Dyson Information Content. Preprint. To appear on *Infinite Dimensional Analysis, Quantum Probability & Related Topics.*

Jenčova, A. (2001). Geometry of quantum states: dual connections and divergence functions, *Reports on Mathematical Physics*, **47**, 121-138.

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

Lieb E.H. (1973). Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Mathematics*, **11**, 267-288.

Nagaoka H. (1995). Differential geometric aspects of quantum state estimation and relative entropy, *Quantum Communications and Measurement*, 449-452, Plenum Press, New York.

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

Petz, D. (2002). Covariance and Fisher information in quantum mechanics, *Journal of Physics. A. Mathematical and General*, **35**, 929-939.

Petz, D. and Hasegawa, H. (1996). On the Riemannian metric of α-entropies of density matrices, *Letters in Mathematical Physics*, **38**, 221-225.

Pistone G. and Sempi C. (1995). An infinite-dimensional geometric structure on the space of all probability measures equivalent to a given one, *Annals of Statistics*, **23**, 1543-1561.

Wigner, E. and Yanase M. (1963). Information content of distribution, *Proceedings of the National Academy of Sciences of the United States of America*, **49**, 910-918.