QUANTUM INFORMATION GEOMETRY AND NON-COMMUTATIVE L_p-SPACES.

ANNA JENČOVÁ

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

Abstract. Let M be a von Neumann algebra. We define the non-commutative extension of information geometry by embeddings of M into non-commutative L_p-spaces. Using the geometry of uniformly convex Banach spaces and duality of the L_p and L_q spaces for $1/p+1/q=1$, we show that we can introduce the α-divergence, for $\alpha \in (-1,1)$, in a similar manner as Amari in the classical case. If restricted to the positive cone, the α-divergence belongs to the class of quasi-entropies, defined by Petz.

1. Introduction.

The classical information geometry deals with the differential geometric aspects of families of probability densities with respect to a given measure μ. The theory, developed in [1, 5], has been already extended to the nonparametric case, where the manifold is modelled on some infinite dimensional Banach space, see [22, 8].

Noncommutative version of the theory has also been proposed, mostly restricted to (invertible) density operators on finite dimensional Hilbert spaces, [11, 14, 17, 21], but there are results in infinite dimensions, [9, 12, 23, 24].

An interesting part of the classical (finite dimensional) information geometry, developed by Amari and Nagaoka [1, 2], deals with the structure of Riemannian manifolds with a pair of dual flat affine connections. For such manifolds, there is a pair (θ, η) of dual affine coordinate systems, related by Legendre transformations

$$\theta_i = \frac{\partial}{\partial \eta_i}\varphi(\eta) \quad \eta_i = \frac{\partial}{\partial \theta_i}\psi(\theta),$$

where ψ, φ are potential functions, satisfying

$$\psi(\theta) + \varphi(\eta) - \sum_i \theta_i \eta_i = 0$$

A quasi–distance, called the divergence, is then defined by

$$D(\theta_1, \theta_2) = \psi(\theta_1) + \varphi(\eta_2) - \sum_i \theta_{1i}\eta_{2i}$$

For manifolds of probability density functions, flat with respect to the $\pm\alpha$–connections, the corresponding α-divergence belongs to the class of Cziszár's f-divergences

$$S_f(p, q) = \int f(\frac{q}{p})dp$$

where f is a convex function.

The f-divergences were generalized to von Neumann algebras by Petz in [20] by means of the relative modular operator of normal positive functionals on M:

$$S_g(\phi, \psi) = (g(\Delta_{\phi,\psi})\xi_\psi, \xi_\psi)$$

where f is operator convex and ξ_ψ is the vector representative of ψ. On the other hand, Amari's construction of the α-divergence, starting from a pair of dual flat connections, was extended to the manifold of faithful positive linear functionals on a matrix algebra $\mathcal{M}_n(\mathbb{C})$, [14, 11] and it was shown that this divergence belongs to the class defined by Petz. The main purpose of the present paper is the extension of this construction to all von Neumann algebras.

One of the important results of the infinite dimensional, classical and quantum, version of information geometry is the definition of Amari-Chentsov α-connections for $\alpha \in (-1, 1)$ given in [8] and [9]. This definition uses the α-embedding of the manifold of density functions (or density operators with respect to a n.s.f. trace on a semifinite von Neumann algebra) into the unit sphere of the (noncommutative) L_p-space, with $p = 2/(1-\alpha)$. It was shown that the Amari–Nagaoka duality of the $+\alpha$ and $-\alpha$-connections is exactly the L_p-space duality and that the fact that these spaces for $1 < p < \infty$ are uniformly convex is basic for this definition.

In the present paper, the α-embedding is defined in a similar manner, but it is extended to the whole predual M_*. This embedding is used to define the manifold structure on M_*. The flat α-connections are induced from the trivial connection on L_p, in our setting the connection is defined on the tangent bundle. Its dual, living on the cotangent bundle, is the $-\alpha$-connection. Moreover, the $\pm\alpha$-embeddings define a pair of dual coordinates on M_*. Using the uniform convexity of the L_p spaces, it is shown that the dual coordinates are related by potential functions, just as in Amari's theory. From this, we can define a divergence functional on $L_p(M, \phi)$.

Via the α-embedding, the divergence in $L_p(M, \phi)$ induces a functional on $M_* \times M_*$, which is called the α-divergence. We will show that if restricted to the positive cone, the α–divergence is exactly the Petz quasi-entropy S_{g_α}, with

$$g_\alpha(t) := \frac{2}{1-\alpha} + \frac{2}{1+\alpha} t - \frac{4}{1-\alpha^2} t^{\frac{1+\alpha}{2}}.$$

We will further investigate the properties of the divergence in $L_p(M, \phi)$, especially the projection theorems. These imply some existence and uniqueness results for the α–projections, which generalize the projection theorems in [1].

2. Geometry of Banach spaces.

In this paragraph, we list some facts about convexity and smoothness of Banach spaces, see [15, 7] for details.

Let X be a Banach space and let X^* be the dual of X. For $u \in X^*$ we denote $\langle x, u \rangle = u(x)$. Let K be a closed convex subset in X with nonempty interior and let S be the boundary of K. In particular, let S_r be the sphere with radius r in X.

2.1. **Supporting hyperplanes and functionals.**

Let us first recall that a (closed) hyperplane in X is a linear manifold $x + H$ with $x \in X$ and H a closed linear subspace of codimension 1. Each hyperplane is uniquely given by x and an element $u \in X^*$ having H as the null-space. Conversely, each $u \in X^*$ defines the hyperplane

$$x + \text{Ker } u = \{y \in X, \ \langle y, u \rangle = \gamma = \langle x, u \rangle\}.$$

If $x + H$ is a real hyperplane in a complex Banach space, then H is given by a real linear continuous functional u and there is a unique $v \in X^*$, such that $u = \Re v$. Each real hyperplane determines two closed half-spaces $\{\langle x, u \rangle \leq \gamma\}$ and $\{\langle x, u \rangle \geq \gamma\}$.

A *supporting hyperplane* of K is a real hyperplane $x + H$, containing at least one point of K and such that K lies in one of the two closed half-spaces determined by $x + H$. If u is the corresponding real linear functional, then u (or $-u$) attains its maximum on K. There is at least one supporting hyperplane through every boundary point of K.

2.2. **Smoothness and strict convexity.** A point $x \in S_1$ is called a *point of smoothness* if there is exactly one supporting hyperplane passing through x, the *tangent hyperplane* at x. Equivalently, there is a unique point $v_x \in X^*$, $\|v_x\| = 1$, such that v_x attains its norm at x. The tangent hyperplane is then determined by $u = \Re v_x$ and all points in the unit ball satisfy $\Re\langle y, v_x\rangle \leq 1$. If each element in S_1 is a point of smoothness, then we say that X is *smooth*.

The norm in X is said to be *weakly (Gateaux) differentiable* at $x \in S_1$ if for each $y \in S_1$, the limit

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} =: q'(x,y) \tag{1}$$

exists. The space X is smooth if and only if its norm is weakly differentiable at each $x \in S_1$. The weak derivative is given by

$$q'(x,y) = \Re\langle y, v_x\rangle.$$

In such a case, norm is weakly differentiable at each $x \in X$ except the origin, and $q'(x,y) = \Re\langle y, v_{x/\|x\|}\rangle$, where $v_{x/\|x\|}$ is the unique point in the unit sphere S_1^* in X^*, such that $\langle x, v_{x/\|x\|}\rangle = \|x\|$.

We say that K is *strictly convex*, if every boundary point of K is and extreme point, that is, if S contains no line segment. In this case, every supporting hyperplane of K contains exactly one point of S. The space X is strictly convex if the closed unit ball K_1 is strictly convex. There is a duality between strict convexity and smoothness of a Banach space:

(i) If X^* is strictly convex then X is smooth.
(ii) If X^* is smooth then X is strictly convex.

2.3. **Uniform smoothness and uniform convexity.** We say that the norm in X is *strongly (Frèchet) differentiable* at $x \in S_1$ if the limit (1) exists uniformly for $y \in S_1$, if this is true for all $x \in S_1$, the norm is termed *strongly differentiable*. The norm is *uniformly strongly differentiable* if (1) exists uniformly in $x, y \in S_1$. Clearly, if the norm is strongly differentiable, then X is smooth and the map

$$S_1 \ni x \mapsto v_x \in S_1^*$$

is well defined. Moreover, we have

Theorem 2.1. [7, 6] *The norm is (uniformly) strongly differentiable if and only if the map $x \mapsto v_x$ is single-valued and norm to norm (uniformly) continuous from S_1 to S_1^*.*

The space X is *uniformly smooth* if for each $\epsilon > 0$ there is an $\eta(\epsilon) > 0$, such that $\|x\| \geq 1$, $\|y\| \geq 1$ and $\|x - y\| \leq \eta(\epsilon)$ always implies $\|x + y\| \geq \|x\| + \|y\| - \epsilon\|x - y\|$.

The dual notion to uniform smoothness is uniform convexity: X is *uniformly convex* if for each $0 < \epsilon \leq 2$ there is a $\delta(\epsilon) > 0$ such that if $x, y \in K_1$ and $\|x-y\| \geq \epsilon$, then $\|\frac{1}{2}(x+y)\| \leq 1-\delta(\epsilon)$. The function $\delta(\epsilon)$ is called the *module of convexity.* Every uniformly convex Banach space is strictly convex and reflexive. Moreover, the following statements are equivalent.

(i) The norm on X (X^*) is uniformly strongly differentiable.
(ii) X is uniformly smooth (uniformly convex).
(iii) X^* is uniformly convex (uniformly smooth).

Let us now define the map $F : X \setminus \{0\} \to X^* \setminus \{0\}$ by

$$F(x) = \|x\| v_{x/\|x\|}$$

Then F is a *support mapping* [7], that is,

(i) $\|x\| = 1$ implies $\|F(x)\| = 1 = \langle x, F(x)\rangle$
(ii) $\lambda \geq 0$ implies $F(\lambda x) = \lambda F(x)$

It follows from paragraph 2.2 that if X is smooth, the support mapping is unique. If we put $F(0) = 0$, then we have

Theorem 2.2. [6] *Let the Banach space X be uniformly convex and let the norm be strongly differentiable. Then F is a homeomorphism of X onto X^* (in the norm topologies).*

3. Noncommutative L_p spaces.

In this section, we recall the definition and some properties of non-commutative L_p spaces on a von Neumann algebra, following the approach in [4, 16].

Let M be a von Neumann algebra and let ϕ be a faithful normal semifinite weight. We denote N_ϕ the set of $y \in M$ satisfying $\phi(y^*y) < \infty$ and M_0 the set of all elements in $N_\phi \cap N_\phi^*$, entire analytic with respect to the modular automorphism σ_t^ϕ associated with ϕ. We also denote the GNS map by $N_\phi \ni y \mapsto \eta_\phi(y) \in H_\phi$.

Let $1 \leq p \leq \infty$. The non-commutative L_p-space with respect to ϕ is the space $L_p(M, \phi)$ of all closed operators acting on the Hilbert space H_ϕ, satisfying

$$T J_\phi \sigma^\phi_{-i/p}(y) J_\phi \supset J_\phi y J_\phi T,$$

for all $y \in M_0$, such that the L_p-norm

$$\|T\|_p = \{\sup_{x\in M_0, \|x\|\leq 1} \||T|^{p/2}\eta_\phi(x)\|\}^{2/p}$$

is finite. Then $L_p(M,\phi)$ with this norm is a Banach space. Let $1 < p < \infty$, then $L_p(M,\phi)$ is uniformly convex and uniformly strongly differentiable. The dual space $L_p^*(M,\phi)$ is $L_q(M,\phi)$, with $1/p+1/q=1$, where the duality is given by

$$\langle T, T'\rangle_\phi = \lim_{y\to 1}(\, T\eta_\phi(y)\, , T'\eta_\phi(y)\,) \tag{2}$$

where $T \in L_p(M,\phi)$, $T' \in L_q(M,\phi)$. The limit is taken in the *-strong topology with restriction $y \in M_0$, $\|y\| \le 1$. Each $T \in L_p(M,\phi)$, $1 \le p < \infty$, has a unique polar decomposition of the form

$$T = u\Delta_{\psi,\phi}^{1/p} \tag{3}$$

where $\psi \in M_*^+$, $u \in M$ is a partial isometry, such that the support projection $s(\psi) = u^*u$ and $\Delta_{\psi,\phi}$ is the relative modular operator, see Appendix C in [4] for definition and basic properties. We have

$$\|u\Delta_{\psi,\phi}^{1/p}\|_p = \psi(1)^{1/p}. \tag{4}$$

On the other hand, each operator of the form (3) is in $L_p(M,\phi)$. The positive cone $L_p^+(M,\phi)$ is the set of positive operators in $L_p(M,\phi)$ and we have

$$L_p^+(M,\phi) = \{\Delta_{\psi,\phi}^{1/p},\ \psi \in M_*^+\}$$

The identity

$$\varphi(au) = \langle u\Delta_{\varphi,\phi}, a^*\rangle_\phi \tag{5}$$

for $a \in M$ gives an isometric isomorphism of M_* and $L_1(M,\phi)$. Similarly, $L_2(M,\phi)$ is isomorphic to H_ϕ by

$$u\Delta_{\varphi,\phi}^{1/2} \mapsto u\xi_\varphi,$$

where ξ_φ is the vector representative of φ in the natural positive cone in H_ϕ.

If $\tilde\phi$ is a different n.s.f. weight, then there is an isometric isomorphism $\tau_p(\tilde\phi,\phi):\ L_p(M,\phi) \to L_p(M,\tilde\phi)$ and

$$\langle T, T'\rangle_\phi = \langle \tau_p(\tilde\phi,\phi)T, \tau_q(\tilde\phi,\phi)T'\rangle_{\tilde\phi} \tag{6}$$

holds for all $T \in L_p(M,\phi)$ and $T' \in L_q(M,\phi)$.

A bilinear form on $L_p(M,\phi)\times L_q(M,\phi)$ is defined by

$$[T, T']_\phi = \langle T, T'^*\rangle_\phi, \quad T \in L_p(M,\phi),\ T' \in L_q(M,\phi)$$

If $T_k \in L_{p_k}(M,\phi)$, $\sum_k 1/p_k = 1/r$, then the product $T = T_1...T_n$ is well defined as an element of $L_r(M,\phi)$ and

$$\|T\|_r \le \|T_1\|_{p_1}\dots\|T_n\|_{p_n}$$

If $r = 1$, then

$$(7) \qquad [T_1 \dots T_n]_\phi \;:=\; [T,1]_\phi = [T_1 \dots T_k, T_{k+1} \dots T_n]_\phi = \\ = [T_{k+1} \dots T_n T_1 \dots T_k]_\phi$$

for each $1 \le k \le n-1$ and

$$(8) \qquad |[T_1 \dots T_n]_\phi| \le \|T_1\|_{p_1} \dots \|T_n\|_{p_n}$$

Let $L_p^h(M,\phi) = \{T \in L_p(M,\phi),\ T = T^*\}$. Since the adjoint operation is a conjugate linear isometry on $L_p(M,\phi)$, $L_p^h(M,\phi)$ is a real Banach subspace of $L_p(M,\phi)$. Moreover, for $1 < p < \infty$, $L_p^h(M,\phi)$ is uniformly convex and the dual $L_p^{h*}(M,\phi) = L_q^h(M,\phi)$ for $1/p+1/q = 1$.

4. The manifold and dual affine connections.

Let M be a von Neumann algebra and let ϕ be a faithful normal semifinite weight.

For $-1 < \alpha < 1$, we define the non-commutative α-embedding by

$$\begin{aligned} \ell_\alpha^\phi:\ M_* &\to L_p(M,\phi), \quad p = \frac{2}{1-\alpha} \\ \omega &\mapsto pu\Delta_{\varphi,\phi}^{1/p} \end{aligned}$$

where $\omega(a) = \varphi(au)$, $a \in M$ is the polar decomposition of ω. It is clear from uniqueness of the polar decompositions that ℓ_α^ϕ is bijective. Moreover, it maps the hermitian (that is, $\omega(a^*) = \overline{\omega(a)}$) elements in M_* onto $L_p^h(M,\phi)$ and M_*^+ onto the positive cone $L_p^+(M,\phi)$.

If ψ is a different f.n.s. weight, then the space $L_p(M,\psi)$ is identified with $L_p(M,\phi)$ by the isometric isomorphism $\tau_p(\psi,\phi)$. The corresponding α-embeddings are related by

$$\ell_\alpha^\psi = \tau_p(\psi,\phi)\ell_\alpha^\phi$$

We denote by $\mathcal{M}_\alpha$ the set M_* with the manifold structure induced from ℓ_α^ϕ. Due to the above isomorphism, the manifold structure does not depend of the choice of ϕ. For $\omega \in M_*$, $\ell_\alpha^\phi(\omega) \in L_p(M,\phi)$ will be called the α-coordinate of ω. The $-\alpha$-coordinate is an element of the dual space $L_q(M,\phi)$, $1/p + 1/q = 1$. Moreover, for $\omega_1, \omega_2 \in M_*$ and a n.s.f. weight ψ, we have by (6)

$$(9) \qquad \langle \ell_\alpha^\psi(\omega_1), \ell_{-\alpha}^\psi(\omega_2)\rangle_\psi = \langle \tau_p(\psi,\phi)\ell_\alpha^\phi(\omega_1), \tau_q(\psi,\phi)\ell_{-\alpha}^\phi(\omega_2)\rangle_\psi = \\ = \langle \ell_\alpha^\phi(\omega_1), \ell_{-\alpha}^\phi(\omega_2)\rangle_\phi$$

In the sequel, we will just write ℓ_α instead of ℓ_α^ϕ. We will say that $\ell_\alpha(\omega)$ and $\ell_{-\alpha}(\omega)$ are dual coordinates of $\omega \in M_*$.

Remark 4.1. It is clear that it is also possible to define the manifold structure on hermitian elements in M_*, using the real Banach space $L_p^h(M,\phi)$. All the subsequent statements hold also for this case.

The trivial connection on $L_p(M,\phi)$ induces a globally flat affine connection on the tangent bundle $T\mathcal{M}_\alpha$, called the α-connection. Let us recall that there is a one-to-one correspondence between affine connections and parallel transports on $T\mathcal{M}_\alpha$. If the connection is globally flat, the parallel transport is given by a family of isomorphisms $U_{x,y}: T_x(\mathcal{M}_\alpha) \to T_y(\mathcal{M}_\alpha)$, $x, y \in \mathcal{M}_\alpha$, satisfying

(i) $U_{x,x} = Id$,

(ii) $U_{y,z}U_{x,y} = U_{x,z}$

In our case, the tangent space $T_x(\mathcal{M}_\alpha)$ can be identified with $L_p(M,\phi)$ and the map $U_{x,y}$ is the identity map for all $x, y \in \mathcal{M}_\alpha$. We define the dual connection as in [8], that is, a linear connection on the cotangent bundle $T^*\mathcal{M}_\alpha$, such that the corresponding parallel transport U^* satisfies

$$\langle v, U^*_{x,y}(w)\rangle_\phi = \langle U_{y,x}(v), w\rangle_\phi = \langle v, w\rangle_\phi$$

for $w \in (T_x(\mathcal{M}_\alpha))^* \equiv L_q(M,\phi)$ and $v \in T_y(\mathcal{M}_\alpha)$. Obviously, U^* is the trivial parallel transport in $L_q(M,\phi)$, hence the dual of the α-connection is the $-\alpha$-connection.

5. Duality.

The L_p spaces for $1 < p < \infty$ are uniformly convex and uniformly smooth, therefore we can use the results of Section 2.

Let $\omega \in M_*$. We will show how ω is related to its dual coordinates.

Proposition 5.1. *Let $\omega \in M_*$ and let $\omega(a) = \psi(au)$ be the polar decomposition. Then*

$$pq\omega(a) = \langle \ell_\alpha(\omega), a^*u^*\ell_{-\alpha}(\omega)\rangle_\phi, \quad a \in M.$$

Proof. We have $s(\Delta_{\psi,\phi}) = s(\psi) = u^*u$. From this and from (5), (7) it follows that

$$\begin{aligned}\psi(au) &= \langle u\Delta_{\psi,\phi}, a^*\rangle_\phi = [u\Delta_{\psi,\phi}u^*ua]_\phi = [u\Delta^{1/p}_{\psi,\phi}\Delta^{1/q}_{\psi,\phi}u^*ua]_\phi = \\ &= [u\Delta^{1/p}_{\psi,\phi}, \Delta^{1/q}_{\psi,\phi}u^*ua]_\phi = \frac{1}{pq}\langle \ell_\alpha(\omega), a^*u^*\ell_{-\alpha}(\omega)\rangle_\phi\end{aligned}$$

□

Let $x = \ell_\alpha(\omega)$ and $\tilde{x} = \ell_{-\alpha}(\omega)$ be the dual coordinates of $\omega \in M_*$. The map

$$x \mapsto \tilde{x} = \ell_{-\alpha}\ell_\alpha^{-1}(x)$$

is called the duality map. It is easy to see from (4) and Proposition 5.1 that for $x \in L_p(M, \phi)$ we have

$$\|\frac{\tilde{x}}{q}\|_q^q = \|\frac{x}{p}\|_p^p = \frac{1}{pq}\langle x, \tilde{x}\rangle_\phi \tag{10}$$

It follows that

$$v_{x/\|x\|_p} = \|\frac{x}{p}\|_p^{1-p}\frac{\tilde{x}}{q} \tag{11}$$

Proposition 5.2. *The duality map is a homeomorphism $L_p(M, \phi) \to L_q(M, \phi)$.*

Proof. It is immediate from (10) that the duality map is continuous at 0. Further, let F be the map defined in Section 2 and $x \neq 0$, then we have from (11)

$$F(x) = \|x\|_p v_{x/\|x\|_p} = \frac{p^p}{pq}\|x\|_p^{2-p}\tilde{x}$$

The statement now follows from Theorem 2.2.

□

Let us define the function $\Psi_p : \ L_p(M, \phi) \to R^+$ by

$$\Psi_p(x) = q\|\frac{x}{p}\|_p^p = q\varphi(1),$$

where $x = pu\Delta_{\varphi,\phi}^{1/p}$. Then we have

Proposition 5.3. *Ψ_p is strongly differentiable. The strong derivative at x is given by*

$$D\Psi_p(x)(y) = \Re\langle y, \tilde{x}\rangle_\phi, \quad y \in L_p(M, \phi)$$

where $\tilde{x}$ is the dual coordinate. If $1/p + 1/q = 1$, then

$$\Psi_q(\tilde{x}) = \Re\langle x, \tilde{x}\rangle_\phi - \Psi_p(x)$$

Proof. We have from uniform smoothness of $L_p(M, \phi)$ that the norm is strongly differentiable at all points except $x = 0$ and

$$D\|x\|_p(y) = \Re\langle y, v_{x/\|x\|_p}\rangle_\phi$$

It follows from (11) that for $x \neq 0$,

$$D\Psi_p(x)(y) = q\|\frac{x}{p}\|_p^{p-1}\Re\langle y, v_{x/\|x\|_p}\rangle_\phi = \Re\langle y, \tilde{x}\rangle_\phi$$

As $p > 1$, the function $\|\frac{x}{p}\|_p^p$ is strongly differentiable at $x = 0$ and

$$D\Psi_p(0)(y) = 0 = \Re\langle y, \tilde{0}\rangle_\phi$$

The last equality is rather obvious.

□

In the commutative case, as well as on the manifold of positive definite $n \times n$ matrices, Ψ_p is the potential function in the sense of Amari, see [1] and [14, 11]. In general, it is not twice differentiable, but the above Proposition shows that the Legendre transformations, relating the dual coordinate systems, are still valid. It will be also clear from the results of the next Section, that

$$\Psi_q(\tilde{x}) = \sup_{y \in L_p(M,\phi)} (\Re\langle y, \tilde{x}\rangle_\phi - \Psi_p(y))$$

hence Ψ_q is the conjugate of the convex function Ψ_p.

6. Divergence in $L_p(M, \phi)$.

Following [1], the function $D_p : L_p(M,\phi) \times L_p(M,\phi) \to R^+$, defined by

$$D_p(x,y) = \Psi_p(x) + \Psi_q(\tilde{y}) - \Re\langle x, \tilde{y}\rangle_\phi$$

is called the divergence. It has the following properties.

Proposition 6.1. (i) *Let $f_p(t) = p + qt^p - pqt$. Then*

$$D_p(x,y) \geq \|\frac{y}{p}\|_p^p f_p(\frac{\|x\|_p}{\|y\|_p}) \tag{12}$$

for all $x, y \in L_p(M,\phi)$, where for $y = 0$, we take the limit $\lim_{t\to 0} t^p f_p(s/t) = qs^p$ for all s. In particular, $D_p(x,y) \geq 0$ for all $x, y \in L_p(M,\phi)$ and equality is attained if and only if $x = y$.

(ii) *D_p is jointly continuous and strongly differentiable in the first variable.*

(iii) $D_p(y,x) = D_q(\tilde{x}, \tilde{y})$

(iv) $D_p(x,y) + D_p(y,z) = D_p(x,z) + \Re\langle x - y, \tilde{z} - \tilde{y}\rangle_\phi$

Proof. The statement (ii) follows from Proposition 5.2 and 5.3. (iii) and (iv) follow easily from the definition of D_p. We will now prove (i). If $y = 0$, then $D_p(x,y) = \Psi_p(x)$ and if $x = 0$, $D_p(x,y) = \Psi_q(\tilde{y})$, which is equal to the right hand side of (12).

Let now $x \neq 0$, $y \neq 0$ and let $t = \|x\|_p/\|y\|_p$. Then by (11)

$$\Re\langle x, \tilde{y}\rangle_\phi = tq\|\frac{y}{p}\|_p^{p-1}\Re\langle\frac{x}{t}, v_{y/\|y\|_p}\rangle_\phi$$

Let $\|y\|_p = r$ and let S_r be the sphere with radius r in $L_p(M,\phi)$. Then $y, \frac{x}{t} \in S_r$. From Section 2, the tangent hyperplane $y + H$ to S_r at y is given by $\Re\langle z, v_{y/r}\rangle_\phi = r$, S_r lies entirely in the half-space given by $\Re\langle z, v_{y/r}\rangle_\phi \leq r$ and y is the unique point of S_r contained in $y + H$. Hence,

$$D_p(x,y) \geq \Psi_p(x) + \Psi_q(\tilde{y}) - tpq\|\frac{y}{p}\|_p^p = \|\frac{y}{p}\|_p^p f_p(t) \geq 0,$$

where equality is attained in the first inequality if and only if $\frac{x}{t} = y$, and in the second inequality if and only if $t = 1$. $\square$

We will also need the following lemma.

Lemma 6.1. *Let $y \in L_p(M, \phi)$, $d > 0$ and let*

$$U_{y,d} := \{x \in L_p(M, \phi),\ D_p(x, y) \leq d\}$$

Then $U_{y,d}$ is convex and weakly compact.

Proof. It is easy to see that D_p is convex in the first variable, therefore the set $U_{y,d}$ is also convex. To show that it is weakly compact, it is sufficient to prove that it is weakly closed and norm bounded.

Let $\{x_n\}$ be a sequence in $U_{y,d}$, converging weakly to some $x \in L_p(M, \phi)$. Then $\|x\|_p \leq \liminf_{n\to\infty} \|x_n\|_p$. We therefore have

$$\begin{aligned} D_p(x, y) &= \Psi_q(\tilde{y}) + q\|\frac{x}{p}\|_p^p - \langle x, \tilde{y}\rangle_\phi \leq \\ &\leq \liminf_{n\to\infty}(\Psi_q(\tilde{y}) + q\|\frac{x_n}{p}\|_p^p - \langle x_n, \tilde{y}\rangle_\phi) = \\ &= \liminf_{n\to\infty} D_p(x_n, y) \leq d \end{aligned}$$

and $U_{y,d}$ is weakly closed.

Finally, for $x \in U_{y,d}$, we have by Proposition 6.1 (i),

$$\|\frac{y}{p}\|_p^p f_p(\frac{\|x\|_p}{\|y\|_p}) \leq D_p(x, y) \leq d$$

As $f_p(t)$ goes to infinity for $t \to \infty$, we see that $\|x\|_p$ is bounded. $\square$

7. D_p-PROJECTIONS.

Let C be a subset in $L_p(M, \phi)$, $y \in L_p(M, \phi)$. If there is a point $x_m \in C$, such that

$$D_p(x_m, y) = \min_{x\in C} D_p(x, y)$$

then x_m will be called a D_p-projection of y to C. In this section, we prove some uniqueness and existence results for D_p-projections.

Proposition 7.1. *Let C be a convex subset in $L_p(M, \phi)$, $y \in L_p(M, \phi)$ and $x_m \in C$. The following are equivalent.*

(i) $D_p(x_m, y) = \min_{x\in C} D_p(x, y)$

(ii) *$\tilde{y} - \tilde{x}_m$ is in the normal cone to C at x_m, that is,*

$$\Re\langle x - x_m, \tilde{y} - \tilde{x}_m\rangle_\phi \leq 0, \quad \forall x \in C$$

(iii) $D_p(x, y) \geq D_p(x, x_m) + D_p(x_m, y), \quad \forall x \in C$

If such a point exists, it is unique.

Proof. Let x_m be a point in C satisfying (i) and let $x \in C$. Then $x_t = tx + (1-t)x_m$ lies in C for all $t \in [0,1]$ and thus $D_p(x_t, y) \geq D_p(x_m, y)$ on $[0,1]$. We have from Proposition 5.3

$$0 \leq \frac{d}{dt^+} D_p(x_t, y)|_{t=0} = \Re\langle x - x_m, \tilde{x}_m - \tilde{y}\rangle_\phi$$

which is (ii). Further, from Proposition 6.1 (iv)

$$\Re\langle x - x_m, \tilde{x}_m - \tilde{y}\rangle_\phi = D_p(x, y) - D_p(x, x_m) - D_p(x_m, y),$$

hence (ii) implies (iii). Finally, let x_m satisfy (iii), then we clearly have $D_p(x_m, y) \leq D_p(x, y)$, for all $x \in C$.

To prove uniqueness, suppose that x_1 and x_2 are points in C, satisfying (iii). Then

$$D_p(x_1, y) \geq D_p(x_1, x_2) + D_p(x_2, y) \geq D_p(x_1, x_2) + D_p(x_2, x_1) + D_p(x_1, y).$$

It follows that $D_p(x_1, x_2) + D_p(x_2, x_1) \leq 0$ and hence $x_1 = x_2$. □

Proposition 7.2. *Let $C \neq \emptyset$ be a weakly closed subset in $L_p(M, \phi)$ and $y \in L_p(M, \phi)$. Then there exists a D_p-projection of y to C.*

Proof. For some $d > 0$, the set $U_{y,d}$ has a nonempty intersection with C. By Lemma 6.1, the sets $U_{y,d} \cap C$ are weakly compact. The intersection of these sets for all such d is therefore nonempty and is equal to some $U_{y,\rho} \cap C$. Then $\rho = \min_{x \in C} D_p(x, y)$ and all the points in $U_{y,\rho} \cap C$ are D_p-projections of y in C. □

By Propositions 7.1 and 7.2, if C is a nonempty convex weakly closed subset in $L_p(M, \phi)$, we can define the map $y \mapsto x_m$, which sends each point y to its unique D_p-projection in C.

Proposition 7.3. *Let $C \neq \emptyset$ be a weakly closed convex subset in $L_p(M, \phi)$. Then the D_p-projection is continuous from $L_p(M, \phi)$ with its norm topology to C with the relative weak topology.*

Proof. Let $\{y^n\}$ be a sequence in $L_p(M, \phi)$ converging in norm to y. Let x_m^n be the unique D_p-projection of y^n and x_m be the unique D_p-projection of y in C, obtained by Propositions 7.2 and 7.1. We have to prove that x_m^n converges weakly to x_m.

As the duality map is continuous, we have $\tilde{y}^n \to \tilde{y}$ in $L_q(M, \phi)$. Further, we have from joint continuity of D_p that $\lim D_p(z, y^n) = D_p(z, y)$ and $\lim D_p(y^n, z) = D_p(y, z)$ for any $z \in L_p(M, \phi)$, in particular, $\lim D_p(y^n, y) = \lim D_p(y, y^n) = 0$.

Let $k_1 > 0$ be such that $\|y\| \le k_1$ and $\|y^n\| \le k_1$ for all n. Further, let us choose an element $z \in C$. By Proposition 6.1 (i) and the definition of the D_p-projection, we have

$$f_p\Big(\frac{\|x_m^n\|}{\|y^n\|}\Big) \le \frac{D_p(z, y^n)}{\|y^n/p\|^p} \tag{13}$$

$$f_p\Big(\frac{\|x_m\|}{\|y\|}\Big) \le \frac{D_p(z, y)}{\|y/p\|^p} \tag{14}$$

As the right hand side of (13) converges to the right hand side of (14), we see that there is a constant $k_2 > 0$, such that $\|x_m\| \le k_2$ and $\|x_m^n\| \le k_2$ for all n. For sufficiently large n, we get by Proposition 6.1 (iv)

$$\begin{aligned} d_n &:= D_p(x_m^n, y^n) = \inf_{x\in C, \|x\|_p \le k_2} D_p(x, y^n) = \\ &= \inf_{x\in C, \|x\|_p \le k_2} \{D_p(x, y) + D_p(y, y^n) - \Re\langle x - y, \tilde{y}^n - \tilde{y}\rangle_\phi\} \le \\ &\le D_p(x_m, y) + D_p(y, y^n) + (k_1 + k_2)\|\tilde{y} - \tilde{y}^n\|_q \le d + \varepsilon \end{aligned}$$

where $d := D_p(x_m, y)$. Similarly,

$$\begin{aligned} D_p(x_m^n, y) &= D_p(x_m^n, y^n) + D_p(y^n, y) - \Re\langle x_m^n - y^n, \tilde{y} - \tilde{y}^n\rangle_\phi \le \\ &\le d_n + D_p(y^n, y) + (k_1 + k_2)\|\tilde{y} - \tilde{y}^n\|_q \le d + 2\varepsilon \end{aligned}$$

Hence for sufficiently large n, $x_m^n \in U_{y,d+2\varepsilon} \cap C$. These sets are nonempty and weakly compact, therefore $\{x_m^n\}$ contains a weakly convergent subsequence. On the other hand, any limit of such subsequence has to be in $U_{y,d+2\varepsilon} \cap C$ for all ε and thus also in $\bigcap_\varepsilon U_{y,d+2\varepsilon} \cap C$. By uniqueness, this intersection contains a single point x_m, it follows that x_m^n converges weakly to x_m. □

8. The α-divergence in M_*.

Let $\alpha \in (-1, 1)$ and let $p = \frac{2}{1-\alpha}$. The divergence in $L_p(M, \phi)$ defines the functional $S_\alpha : M_* \times M_* \to R^+$,

$$\begin{aligned} S_\alpha(\omega_1, \omega_2) &:= D_p(\ell_\alpha(\omega_1), \ell_\alpha(\omega_2)) = \\ &= q\varphi(1) + p\psi(1) - pq\Re\langle u\Delta_{\varphi,\phi}^{1/p}, v\Delta_{\psi,\phi}^{1/q}\rangle_\phi \end{aligned}$$

where $\omega_1(a) = \varphi(au)$ and $\omega_2(a) = \psi(av)$ are the polar decompositions. It is called the α-divergence. Let us remark that a similar definition for the α-divergence appeared also in the Discussion in [19].

It follows from (9) that S_α does not depend of ϕ. In particular, if ψ is faithful, then

$$\langle u\Delta_{\varphi,\phi}^{1/p}, v\Delta_{\psi,\phi}^{1/q}\rangle_\phi = (\Delta_{\varphi,\xi_\psi}^{1/(2p)}\xi_\psi, \Delta_{\varphi,\xi_\psi}^{1/(2p)}u^*v\xi_\psi)$$

where ξ_ψ is a vector representative of ψ. It follows that if $\varphi, \psi \in M_*^+$, ψ is faithful and $\Delta_{\varphi,\xi_\psi} = \int \lambda E_\lambda$ is the spectral decomposition, then

$$S_\alpha(\varphi,\psi) = (g_p(\Delta_{\varphi,\xi_\psi})\xi_\psi, \xi_\psi) = \int g_p(\lambda)\|E_\lambda \xi_\psi\|^2$$

where $g_p(t) = p + qt - pqt^{1/p}$. Hence, in this case the α-divergence is equal to the quasi entropy $S^1_{g_p}$, defined by Petz in [20](see also [18]). We will show that this is true on the whole of $M_*^+ \times M_*^+$.

Lemma 8.1. *Let $\varphi, \psi \in M_*^+$, $u, v \in M$ be partial isometries satisfying $u^*u = s(\varphi)$, $v^*v = s(\psi)$. Let $p, q > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\langle u\Delta^{1/p}_{\varphi,\phi}, v\Delta^{1/q}_{\psi,\phi}\rangle_\phi = (\,\Delta^{1/(2p)}_{\varphi,\xi_\psi}\xi_\psi, \Delta^{1/(2p)}_{\varphi,\xi_\psi}u^*v\xi_\psi\,) \tag{15}$$

where ξ_ψ is a vector representative of ψ.

Proof. Let $1/p \le 1/2$. We have

$$\langle u\Delta^{1/p}_{\varphi,\phi}\,, v\Delta^{1/q}_{\psi,\phi}\rangle_\phi = \lim_{y\to 1}(\,\Delta^{1/2-1/p}_{\psi,\phi}v^*u\Delta^{1/p}_{\varphi,\phi}\eta_\phi(y)\,, \Delta^{1/2}_{\psi,\phi}\eta_\phi(y)\,),$$

with $y \in M_0$, $\|y\| \le 1$. For $y \in N_\phi$,

$$\begin{aligned}(\,\Delta^{1/2-1/p}_{\psi,\phi}v^*u\Delta^{1/p}_{\varphi,\phi}\eta_\phi(y)\,, \Delta^{1/2}_{\psi,\phi}\eta_\phi(y)\,) &= \\ = (\,J_{\xi_\psi,\eta_\phi}\Delta^{1/2}_{\psi,\phi}\eta_\phi(y)\,, J_{\xi_\psi,\eta_\phi}\Delta^{1/2-1/p}_{\psi,\phi}v^*u\Delta^{1/p}_{\varphi,\phi}\eta_\phi(y)\,) &= \\ = (\,y^*\xi_\psi\,, J_{\xi_\psi,\eta_\phi}\Delta^{1/2-1/p}_{\psi,\phi}v^*u\Delta^{1/p}_{\varphi,\phi}\eta_\phi(y)\,),&\end{aligned} \tag{16}$$

here we have used that $J^*_{\xi_\psi,\eta_\phi}J_{\xi_\psi,\eta_\phi} = s(\psi) = s(\Delta_{\psi,\phi})$, the support of $\Delta_{\psi,\phi}$. Let $t \in \mathbb{R}$, then

$$\begin{gathered}J_{\xi_\psi,\eta_\phi}\Delta^{1/2-it}_{\psi,\phi}v^*u\Delta^{it}_{\varphi,\phi}\eta_\phi(y) = S_{\xi_\psi,\eta_\phi}\Delta^{-it}_{\psi,\phi}v^*u\Delta^{it}_{\varphi,\phi}\eta_\phi(y) = \\ S_{\xi_\psi,\eta_\phi}v^*\Delta^{-it}_{\psi_v,\phi}\Delta^{it}_{\varphi_u,\phi}u\eta_\phi(y) = S_{\xi_\psi,\eta_\phi}v^*(D\psi_v : D\varphi_u)_{-t}u\eta_\phi(y) = \\ y^*u^*(D\psi_v : D\varphi_u)^*_{-t}v\xi_\psi\end{gathered}$$

where $\varphi_u(a) = \varphi(u^*au)$ and $u\Delta^{it}_{\varphi,\phi}u^* = \Delta^{it}_{\varphi_u,\phi}$ by (C.8) in [4]. From this, we have

$$\begin{aligned}(\,y^*\xi_\psi\,, J_{\xi_\psi,\eta_\phi}\Delta^{1/2-it}_{\psi,\phi}v^*u\Delta^{it}_{\varphi,\phi}\eta_\phi(y)\,) &= (\,(D\psi_v : D\varphi_u)_{-t}uyy^*\xi_\psi\,, v\xi_\psi\,) = \\ &= (\,\Delta^{-it}_{\psi_v,\xi_\psi}\Delta^{it}_{\varphi_u,\xi_\psi}uyy^*\xi_\psi\,, v\xi_\psi\,) = (\,u\Delta^{it}_{\varphi,\xi_\psi}yy^*\xi_\psi\,, v\xi_\psi\,),\end{aligned}$$

where we have used (C.5) and (C.8) of [4]. It follows that for $z = it$,

$$(\,y^*\xi_\psi\,, J_{\xi_\psi,\eta_\phi}\Delta^{1/2-z}_{\psi,\phi}v^*u\Delta^{z}_{\varphi,\phi}\eta_\phi(y)\,) = (y^*\xi_\psi\,, y^*\Delta^{\bar z}_{\varphi,\xi_\psi}u^*v\xi_\psi\,). \tag{17}$$

By Lemma 3.1 in [16], both sides of (17) are holomorphic for $0 < \Re z < 1/2$ and continuous for $0 \le \Re z \le 1/2$. The equation (15) holds for $1/p \le 1/2$ by (16) and analytic continuation of (17).

Let now $1/q \le 1/2$. We have by the first part of the proof

$$\begin{aligned}\langle u\Delta^{1/p}_{\varphi,\phi}\,, v\Delta^{1/q}_{\psi,\phi}\rangle_\phi &= (\,u\xi_\varphi\,, v\Delta^{1/q}_{\xi_\psi,\xi_\varphi}\xi_\varphi\,) = (\,S_{\xi_\varphi,\xi_\psi}u^*v\xi_\psi\,, \Delta^{1/q}_{\xi_\psi,\xi_\varphi}S_{\xi_\varphi,\xi_\psi}\xi_\psi\,)\\ &= (\,J_{\xi_\psi,\xi_\varphi}\Delta^{1/q}_{\xi_\psi,\xi_\varphi}J_{\xi_\varphi,\xi_\psi}\Delta^{1/2}_{\xi_\varphi,\xi_\psi}\xi_\psi\,, \Delta^{1/2}_{\xi_\varphi,\xi_\psi}u^*v\xi_\psi\,) =\\ &= (\,\Delta^{1/p-1/2}_{\varphi,\xi_\psi}\xi_\psi, \Delta^{1/2}_{\varphi,\xi_\psi}u^*v\xi_\psi\,),\end{aligned}$$

we have used the equations (C.14) $J^*_{\eta_1,\eta_2} = J_{\eta_2,\eta_1}$ and (β5) $J_{\eta_1,\eta_2}\Delta_{\eta_1,\eta_2}J_{\eta_2,\eta_1} = \Delta^{-1}_{\eta_2,\eta_1}$ from Appendix C in [4]. □

It follows that $S_\alpha(\varphi,\psi) = S^1_{g_p}(\varphi,\psi)$ for all positive normal functionals φ and ψ. The function g_p, $1 < p < \infty$ is operator convex and it follows from the results in [20] that

(i) S_α is jointly convex on $M^+_* \times M^+_*$
(ii) S_α decreases under stochastic maps on $M^+_* \times M^+_*$
(iii) S_α is lower semicontinuous on $M*^+ \times \mathcal{F}(M^+_*)$ endowed with the product of norm topologies, where $\mathcal{F}(M^+_*)$ denotes the set of faithful elements in M^+_*.

The following properties of the α-divergence are valid on $M_* \times M_*$ and are immediate consequences of the results of Section 6.

(i) Positivity

$$S_\alpha(\varphi,\psi) \ge \|\psi\|_1 g_p(\frac{\|\varphi\|_1}{\|\psi\|_1}) \ge 0$$

and $S_\alpha(\varphi,\psi) = 0$ if and only if $\varphi = \psi$ (here $\|\cdot\|_1$ is the norm in M_*).
(ii) $S_\alpha(\varphi,\psi) = S_{-\alpha}(\psi,\varphi)$
(iii) generalized Pythagorean relation

$$S_\alpha(\varphi,\psi) + S_\alpha(\psi,\sigma) = S_\alpha(\varphi,\sigma) + \Re\langle \ell_\alpha(\varphi) - \ell_\alpha(\psi), \ell_{-\alpha}(\sigma) - \ell_{-\alpha}(\psi)\rangle_\phi$$

Let $\varphi,\psi \in M_*$ and let x_t, $t \in [0,1]$ be a curve in $L_p(M,\phi)$, given by

$$x_t := \ell_\alpha(\varphi) + t(\ell_\alpha(\psi) - \ell_\alpha(\varphi)),$$

then $\ell^{-1}_\alpha(x_t)$ is the α-geodesic in M_*, connecting φ and ψ. Notice that the Pythagorean relation (iii) is a generalization of the classical version in [1], which says that equality is attained if and only if the α-geodesic connecting ψ and φ is orthogonal to the $-\alpha$-geodesic connecting ψ and σ.

We also define the α-projection of $\varphi \in M_*$ onto a subset $C \subset M_*$ as the element in C that minimizes $S_\alpha(\cdot,\varphi)$ over C. We will say that a subset $C \subset M_*$ is α-convex if $\ell_\alpha(C)$ is convex. The next Proposition is a generalization of the results in [1, 2] and follows directly from Proposition 7.1.

Proposition 8.1. *Let $C \subset M_*$ be α-convex and let $\psi \in M_*$, $\varphi_m \in C$. The following are equivalent.*

(i) *φ_m is an α-projection of ψ in C.*

(ii) *For all $\sigma \in C$,*

$$S_\alpha(\sigma, \psi) \geq S_\alpha(\varphi_m, \psi) + S_{-\alpha}(\varphi_m, \sigma).$$

(iii) *If ψ_t is the $-\alpha$-geodesic connecting φ_m and ψ, then $\frac{d}{dt}\ell_{-\alpha}(\psi_t)$ lies in the normal cone to $\ell_\alpha(C)$ at $\ell_\alpha(\varphi_m)$.*

If such a point exists, it is unique.

The topology induced by the α-embedding from the norm, resp. the weak topology in $L_p(M, \phi)$ will be called the α-, resp. the α-weak topology. The following Proposition is also immediate from Section 7.

Proposition 8.2. *Let C be a nonempty subset in M_* and let $\psi \in M_*$.*

(i) *If C is α-weakly closed, then there exists an α-projection of ψ in C.*

(ii) *If C is α-convex, α-weakly closed, then the α-projection is a continuous map from M_* with the α-topology to C with the relative α-weak topology.*

9. The case $\alpha = 0$.

Let $\alpha = 0$, $p = q = 2$. The space $L_2(M, \phi)$ can be identified with the Hilbert space H_ϕ and the dual pairing $\langle \cdot, \cdot \rangle_\phi$ is the inner product $(\cdot, \cdot)$ in H_ϕ. Through this identification, the 0-embedding becomes the map

$$\omega \mapsto 2u\xi_\varphi$$

where $\omega(a) = \varphi(au)$ is the polar decomposition of ω and ξ_φ is the unique vector representative of φ in the natural positive cone V in H_ϕ. Hence the 0-embedding maps M_* bijectively onto H_ϕ. Up to multiplication by 2, the restriction of ℓ_0 to the positive cone M_*^+ corresponds to the identification of the positive normal functionals with elements in V proved by Araki in [3]. It has been also shown that this identification is a homeomorphism $M_*^+ \to V$. It follows that the relative 0-topology is the same as the relative L_1-topology in M_*^+.

The duality map is the identity on H_ϕ and the potential function is

$$\Psi_2(x) = \frac{1}{2}\|x\|^2$$

Therefore, the potential function is C^∞-diferentiable and

$$D^2\Psi_2(x)(y, z) = \Re(y, z) \qquad \forall x \in H_\phi$$

It follows that Ψ_2 defines a Riemannian metric in the tangent bundle $T\mathcal{M}_0$, which corresponds to the real part of the inner product, induced from the 0-embedding. In the matrix case, this metric was studied on density matrices and it was shown that it coincides with the Wigner–Yanase metric, see [10].

The D_2-divergence in H_ϕ is

$$D_2(x,y) = \frac{1}{2}\|x-y\|^2,$$

hence the D_2-projection corresponds to minimizing the Hilbert space norm. This means, in particular, that there is a unique D_2-projection onto every closed convex subset of H_ϕ.

The 0-divergence in M_* becomes

$$S_0(\omega_1,\omega_2) = 2\|u\xi_\varphi - v\xi_\psi\|^2$$

On the positive cone, the 0-divergence generalizes the classical Hellinger distance.

10. The unit sphere.

The α-embedding maps the unit sphere S in M_* onto the sphere S_p with radius p in $L_p(M,\phi)$. The duality map $x \mapsto \tilde{x}$ maps S_p onto the sphere S'_q with radius q in the dual space $L_q(M,\phi)$. From (11), we have that for $x \in S_p$,

$$\tilde{x} = qv_{x/p} \tag{18}$$

Proposition 10.1. *The duality map $S_p \ni x \mapsto \tilde{x} \in S'_q$ is uniformly continuous.*

Proof. The statement follows from (18) and Theorem 2.1. □

For each $x \in S_p$, there is a unique tangent hyperplane $x + H_x$ through x, where H_x is given by the condition

$$\Re\langle y, \tilde{x}\rangle_\phi = q\Re\langle y, v_{x/p}\rangle_\phi = 0$$

Hence there is a splitting $L_p(M,\phi) = H_x \oplus [x]$ and, similarly as in [8], there is a continuous projection $\pi_x : L_p(M,\phi) \to H_x$, given by

$$\pi_x(y) = y - \Re\langle y, v_{x/\|x\|_p}\rangle_\phi \frac{x}{p} = y - \frac{1}{pq}\Re\langle y, \tilde{x}\rangle_\phi x,$$

which is obtained by minimizing the L_p-norm.

As the norm is strongly differentiable, the unit sphere forms a differentiable submanifold $\mathcal{D}_\alpha$ in $\mathcal{M}_\alpha$. Let $\omega \in \mathcal{D}_\alpha$ and let $x \in S_p$ be its α-coordinate. The tangent space $T_x(\mathcal{D}_\alpha)$ can be identified with the tangent hyperplane H_x and π_x can be used to project the α-connection onto $T\mathcal{D}_\alpha$.

Remark 10.1. Let ψ be a faithful state with the α-coordinate x in $L_p(M,\phi)$ and let ξ_ψ be the vector representative of ψ in a natural positive cone. Then ψ has the α-coordinate $\tau_p(\psi,\phi)(x) = p\Delta_{\xi_\psi}^{1/p}$ in $L_p(M,\psi)$ and the dual coordinate $\tau_q(\psi,\phi)(\tilde{x}) = q\Delta_{\xi_\psi}^{1/q}$ in $L_q(M,\psi)$. By (9) we have for $y \in L_p(M,\phi)$,

$$\langle y, \tilde{x}\rangle_\phi = \langle \tau_p(\psi,\phi)(y), q\Delta_{\xi_\psi}^{1/q}\rangle_\psi = (y_\psi \xi_\psi, q\Delta_{\xi_\psi}^{1/q}\xi_\psi) = q\langle y_\psi, 1\rangle_\psi$$

where $y_\psi = \tau_p(\psi,\phi)(y)$. It follows that $\tau_p(\psi,\phi)$ maps $T_x(\mathcal{D}_\alpha)$ onto the subspace

$$\mathcal{F}_\psi^\alpha := \{z \in L_p(M,\psi),\ \Re\langle z, 1\rangle_\psi = 0\}$$

This corresponds to the results in [9], where the α-connection is obtained on the fiber bundle $\mathcal{F}^\alpha$ over a manifold of faithful states.

The projected connection, even in the classical and the matrix case, is no longer flat. Hence, it does not define a divergence, but nevertheless, we can use the restriction of S_α as a quasi-distance on the unit sphere. This restriction has the form

$$S_\alpha(\omega_1,\omega_2) = pq(1 - \Re\langle u\Delta_{\varphi,\phi}^{1/p}, v\Delta_{\psi,\phi}^{1/q}\rangle_\phi)$$

which corresponds to the definition of the α-divergence in [1] for probability densities and in [13] for density matrices.

Acknowledgement. The paper was finished during the stay in RIKEN Brain Science Institute, Tokyo. The author wishes to thank Professor S. Amari for providing the financial support, H. Hasegawa for arranging the visit and I. Ojima and D. Petz for valuable comments. The research was supported by the grant VEGA 1/0264/03 and Science and Technology Assistance Agency under the contract No. APVT-51-032002.

References

[1] S. Amari, Differential-geometrical methods in statistic,Lecture Notes in Statistics,28 (1985)

[2] S. Amari and H. Nagaoka, *Method of information geometry*, AMS monograph, Oxford University Press, 2000

[3] H.Araki, Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon-Nikodym theorem with a chain rule, Pacific J. Math., 50, 1974, 309–354

[4] H. Araki and T. Masuda, Positive cones and L_p-spaces for von Neumann algebras,Publ. RIMS, Kyoto Univ.,18,(1982),339–411

[5] N.N. Chentsov, Statistical decision rules and optimal inferences, Translation of Math. Monog.,53, Amer. Math. Society,Providence, 1982

[6] D. F. Cudia, The geometry of Banach spaces. Smoothness. Trans. Amer. Math. Soc. 110, 284–314 (1964)

[7] J. Diestel, *Geometry of Banach Spaces - Selected Topics* Lecture Notes in Math. 485, Springer, New York 1975

[8] P.Gibilisco and G.Pistone, Connections on non-parametric statistical manifolds by Orlicz space geometry, Inf. Dim. Analysis, Quant. Prob. and Rel. Top.,1 (1998), 325–347

[9] P.Gibilisco and T.Isola, Connections on statistical manifolds of density operators by geometry of non-commutative L_p-spaces, Inf. Dim. Analysis, Quant. Prob. and Rel. Top., 2(1999),169–178

[10] P. Gibilisco and T. Isola, A characterization of Wigner Yanase skew information among statistically monotone metrics, IDAQP, 4 (2001), 553557

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

[12] M. Graselli and R.F. Streater, The quantum information manifold for epsilon-bounded forms,Rep. Math. Phys., 46, 325-335, 2000.

[13] H.Hasegawa, α-divergence of the non-commutative information geometry, Rep. Math. Phys., 33(1993), 87–93

[14] A. Jenčová, Geometry of quantum states: dual connections and divergence functions, Rep.Math.Phys., 47 (2001), 121–138

[15] G. Köthe, Topological vector spaces 1, Springer-Verlag Berlin Heidelberg New York, 1983,

[16] T.Masuda, L_p-spaces for von Neumann algebra with reference to a faithful normal semifinite weight, Publ. RIMS, Kyoto Univ.,19 (1983), 673–727

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

[18] M. Ohya and D. Petz, Quantum Entropy and its use, Springer, Heidelberg, 1993

[19] I. Ojima, Temperature as order parameter of broken scale invariance, to appear in Publ. RIMS, Kyoto Univ.

[20] D. Petz, Quasi-entropies for states of a von Neumann algebra, Publ. RIMS, Kyoto Univ. 21, 1985, 787–800

[21] D. Petz, Information geometry of quantum states, In: Quantum Probability Communications, 10, eds. R.L. Hudson et al, World Scientific, 135–158

[22] G. Pistone, C. Sempi, An infinite dimensional geometric structure on the space of all probability measures equivalent to a given one,Ann. Statist.,23 (1995),1543–1561

[23] R. F. Streater, The Information Manifold for Relatively Bounded Potentials, Tr. Mat. Inst. Steklova, 228, 217-235, 2000.

[24] R.F. Streater, The analytic quantum information manifold, pp 603-611 In: *Stochastic Processes, Physics and Geometry: New Interplays, II*,(Leipzig 1999), Amer. Math. Soc. Providence, RI 2000