A CONSTRUCTION OF A NONPARAMETRIC QUANTUM INFORMATION MANIFOLD.

ANNA JENČOVÁ

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

Abstract. We present a construction of a Banach manifold on the set of faithful normal states of a von Neumann algebra, where the underlying Banach space is a quantum analogue of an Orlicz space. On the manifold, we introduce the exponential and mixture connections as dual pair of affine connections.

1. Introduction

An information manifold is a family of states of some classical or quantum system, endowed with a differentiable manifold structure. For finitely parametrized families of probability distributions, the geometry of such manifolds and its applications in parameter estimation is already well understood, see for example the books [1, 2].

The non-parametric version was introduced by Pistone and Sempi [9, 10, 11], based on the theory of Orlicz spaces. For the quantum version, some proposals for infinite dimensional manifold structure can be found in [6, 10, 11].

The aim of this paper is to introduce a differentiable manifold structure on the set of faithful states of a quantum system, represented by a von Neumann algebra \( \mathcal{M} \). Moreover, we want this manifold to be a quantum counterpart of the Pistone and Sempi construction.

We use an approach similar to Grasselli [5] in the commutative case: we define an Orlicz norm on the space of self-adjoint operators in \( \mathcal{M} \) and take the completion under this norm to be the underlying Banach space for the manifold. The norm is defined by a quantum Young function, as in [12]. The definition of a Young function on a Banach space, together with some results on the associated norms, can be found in Section 3. For a faithful state \( \varphi \), the quantum Orlicz space \( B_\varphi \) and its centered version \( B_{\varphi,0} \) are introduced in Section 4. The definition of the related Young function is based on the relative entropy approach to state perturbation. We treat the dual spaces in Section 6. The main result is contained in Section 8 where the manifold structure is introduced and, moreover, the exponential and mixture connections are defined as a pair of dual affine connections on each connected component of the manifold.
2. Preliminaries.

We recall some properties of relative entropy and perturbed states, that will be needed later. See [5] for details.

Let $\mathcal{M}$ be a von Neumann algebra in standard form. For $\omega$ and $\varphi$ in $\mathcal{M}_+^*$, the relative entropy is defined as

$$ S(\omega, \varphi) = \begin{cases} -\langle \log(\Delta_{\varphi, \xi_\omega}), \xi_\omega \rangle & \text{if } \text{supp } \omega \leq \text{supp } \varphi \\ \infty & \text{otherwise} \end{cases} $$

where $\xi_\omega$ is the representing vector of $\omega$ in a natural positive cone and $\Delta_{\varphi, \xi_\omega}$ is the relative modular operator. Then $S$ is jointly convex and weakly lower semicontinuous. Let us denote $\mathcal{P}_\varphi := \{ \omega \in \mathcal{M}_+^*, S(\omega, \varphi) < \infty \}$, then $\mathcal{P}_\varphi$ is a convex cone. We will need the following Donald’s identity

$$ (1) \quad S(\psi, \varphi) + \sum_i S(\psi_i, \psi) = \sum_i S(\psi_i, \varphi) $$

where $\psi_i \in \mathcal{M}_+^*$, $i = 1, \ldots, n$, and $\psi = \sum_i \psi_i$. Since $S(\psi_i, \psi)$ is always finite, it follows from this identity that $\sum_i \psi_i \in \mathcal{P}_\varphi$ if and only if $\psi_i \in \mathcal{P}_\varphi$ for all $i$.

Let $\mathfrak{S}_+$ be the set of normal states on $\mathcal{M}$ and let $\mathcal{S}_\varphi := \{ \omega \in \mathfrak{S}_+, S(\omega, \varphi) < \infty \}$. Then $\mathcal{S}_\varphi$ is a convex set and generates $\mathcal{P}_\varphi$. From (1), we get

$$ (2) \quad S(\psi_\lambda, \varphi) + \lambda S(\psi_1, \varphi) + (1 - \lambda) S(\psi_2, \varphi) = \lambda S(\psi_1, \varphi) + (1 - \lambda) S(\psi_2, \varphi) $$

where $\psi_1, \psi_2$ are normal states and $\psi_\lambda = \lambda \psi_1 + (1 - \lambda) \psi_2$, $0 \leq \lambda \leq 1$. As above, it follows that $\psi_\lambda \in \mathcal{S}_\varphi$ if and only if both $\psi_1, \psi_2 \in \mathcal{S}_\varphi$, in other words, $\mathcal{S}_\varphi$ is a face in $\mathfrak{S}_+$. For $C > 0$, we define the set $\mathcal{S}_C := \{ \omega, S(\omega, \varphi) \leq C \}$. Then $\mathcal{S}_C$ is convex and compact in the $\sigma(\mathcal{M}_+, \mathcal{M})$ topology.

Let us suppose that $\varphi$ is a faithful normal state on $\mathcal{M}$ and let $h$ be a self-adjoint element in $\mathcal{M}$. The perturbed state $[\varphi^h]$ is defined as the unique maximizer of

$$ (3) \quad \sup_{\omega \in \mathfrak{S}_+} \{ \omega(h) - S(\omega, \varphi) \} $$

Then $[\varphi^h]$ is a faithful normal state and $S([\varphi^h], \varphi)$ is finite. Let $c_\varphi(h)$ be the supremum in (3), that is

$$ (4) \quad c_\varphi(h) = [\varphi^h](h) - S([\varphi^h], \varphi) $$

It is known that

$$ (5) \quad \varphi(h) \leq c_\varphi(h) \leq \log \varphi(e^h) $$

Moreover, we have

$$ (6) \quad \omega(h) - S(\omega, \varphi) = c_\varphi(h) - S(\omega, [\varphi^h]) $$

for any self-adjoint $h \in \mathcal{M}$ and $\omega \in \mathfrak{S}_+$. Let $h, k$ be self-adjoint elements in $\mathcal{M}$, then the chain rule $[\varphi^{h+k}] = [[\varphi^h]^k]$ and

$$ (7) \quad c_\varphi(h + k) = c_{[\varphi^h]}(k) + c_\varphi(h) $$

holds. Let now $\xi_\varphi$ be the vector representative of $\varphi$ and let $\varphi^h \in M_+^*$ be the functional induced by the perturbed vector

$$ \xi^h_\varphi := e^{\varphi^h_{(\log \Delta + h)}} \xi_\varphi = e^{\varphi^h_{(\log \Delta + h)}} \Delta^{1/2}_{[\varphi^h], \varphi} \xi_\varphi $$

Then $c_\varphi(h) = \log \varphi^h(1)$ and $[\varphi^h] = \varphi^h / \varphi^h(1)$. Moreover, if $\varphi^h = \varphi^k$, then $h = k$. 

3. Young functions on Banach spaces and the associated norms.

Let $V$ be a real Banach space and let $V^*$ be its dual, with the duality pairing $\langle v, x \rangle = v(x)$. Recall that any convex lower semicontinuous function $V \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous with respect to the $\sigma(V, V^*)$-topology.

**The Young function.** We will say that a function $\Phi : V \to \mathbb{R} \cup \{\infty\}$ is a Young function, if it satisfies:

(i) $\Phi$ is convex and lower semicontinuous,
(ii) $\Phi(x) \geq 0$ for all $x \in V$ and $\Phi(0) = 0$
(iii) $\Phi(x) = \Phi(-x)$ for all $x \in V$
(iv) if $x \neq 0$, then $\lim_{t \to \infty} \Phi(tx) = \infty$

**Lemma 3.1.** Let $\Phi$ be a Young function. Let us define the sets

$$C_\Phi := \{ x \in V, \Phi(x) \leq 1 \}$$

$$L_\Phi := \{ x \in V, \exists s > 0, \text{ such that } \Phi(sx) < \infty \}.$$ 

Then $C_\Phi$ is absolutely convex and $L_\Phi = \bigcup_n nC_\Phi$. In particular, $L_\Phi$ is a (real) vector space.

**Proof.** Let $x, y \in C_\Phi$ and let $\alpha, \beta \in \mathbb{R}$, such that $|\alpha| + |\beta| \leq 1$. Put $\gamma = 1 - |\alpha| - |\beta|$, then

$$\Phi(\alpha x + \beta y) = \Phi(|\alpha| \text{sgn}(\alpha)x + |\beta| \text{sgn}(\beta)y + \gamma 0) \leq |\alpha|\Phi(x) + |\beta|\Phi(y) \leq 1$$

hence $\alpha x + \beta y \in C_\Phi$ and $C_\Phi$ is absolutely convex.

Let now $x \in L_\Phi$ and let $s > 0$ be such that $\Phi(sx) = K < \infty$. Choose $m \in \mathbb{N}$ such that $m \geq \max\{1/s, K/s\}$, then by convexity

$$\Phi(\frac{1}{m}x) = \Phi(\frac{1}{ms}sx) \leq \frac{1}{ms}\Phi(sx) = \frac{K}{ms} \leq 1$$

and $x \in mC_\Phi$. Since obviously $nC_\Phi \subset L_\Phi$ for all $n$, we have $L_\Phi = \bigcup_n nC_\Phi$, which clearly implies that $L_\Phi$ is a vector space. 

Let us recall that the effective domain

$$\text{dom}(\Phi) = \{ x \in V, \Phi(x) < \infty \}$$

is a convex set. Any convex lower semicontinuous function is continuous in the interior of its effective domain. [3]. Clearly, $L_\Phi$ is the smallest vector space containing $\text{dom}(\Phi)$.

In the space $L_\Phi$, we now introduce the Minkowski functional of $C_\Phi$,

$$\|x\|_\Phi := \inf \{ \rho > 0, x \in \rho C_\Phi \}.$$ 

Since $C_\Phi$ is absolutely convex and absorbing, $\| \cdot \|_\Phi$ is a seminorm. Moreover, $\|x\|_\Phi = 0$ means that $\Phi(tx) \leq 1$ for all $t > 0$. By the property (iv), this implies that $x = 0$. It follows that $\| \cdot \|_\Phi$ defines a norm in $L_\Phi$. Let us denote by $B_\Phi$ the completion of $L_\Phi$ under this norm.

**Lemma 3.2.** Let $x \in L_\Phi$. Then $\|x\|_\Phi \leq 1$ if and only if $\Phi(x) \leq 1$. 

Lemma 3.4. Let $v \in C_\Phi$ and $\Phi(x) \leq 1$. Let now $\|x\|_{\Phi} = 1$ and let $t_n < 1$ be a sequence converging to 1. Then $\Phi(t_n x) \leq 1$ for all $n$ and, by lower \textit{semi}continuity, $\Phi(x) \leq \liminf \Phi(t_n x) \leq 1$. Hence $\|x\|_{\Phi} \leq 1$ implies $\Phi(x) \leq 1$. On the other hand, if $\Phi(x) \leq 1$, then $x \in C_\Phi$ and clearly $\|x\|_{\Phi} \leq 1$. □

Lemma 3.3. Let $x \in L_\Phi$. Then $\|x\|_{\Phi} \leq 1$ implies $\Phi(x) \leq \|x\|_{\Phi}$ and $\|x\|_{\Phi} > 1$ implies $\Phi(x) > \|x\|_{\Phi}$. Moreover, if $\Phi$ is finite valued, then $\|x\|_{\Phi} = 1$ if and only if $\Phi(x) = 1$.

Proof. Let $\|x\|_{\Phi} \leq 1$. By convexity of $\Phi$ and Lemma 3.2

$$\Phi(x) = \Phi(\|x\|_{\Phi} \frac{x}{\|x\|_{\Phi}}) \leq \|x\|_{\Phi} \Phi(\frac{x}{\|x\|_{\Phi}}) \leq \|x\|_{\Phi}$$

Let now $\|x\|_{\Phi} > 1$, then $\Phi(x) > 1$. If $\Phi(x) = \infty$, then the assertion is obviously true. Let us suppose that $\Phi(x)$ is finite. The function $t \mapsto \Phi(tx)$ is convex and bounded on $0 < 1 >$, hence continuous on $(0, 1)$. It follows that $\Phi(tx) = 1$ for some $t$ in this interval and clearly $t = \frac{1}{\|x\|_{\Phi}}$. We have

$$1 = \Phi(tx) \leq t\Phi(x)$$

and hence $\|x\|_{\Phi} \leq \Phi(x)$. This also proves that last statement. □

The conjugate function. Let $V^*$ be the dual space. Let the function $\Phi^* : V^* \to \mathbb{R} \cup \{\infty\}$ be the conjugate of $\Phi$,

$$\Phi^*(v) = \sup_{x \in V} \{v(x) - \Phi(x)\} = \sup_{x \in \text{Dom}(\Phi)} \{v(x) - \Phi(x)\}$$

The function $\Phi^*$ is convex, lower \textit{semi}continuous and positive, $\Phi^*(v) = \Phi^*(-v)$ and $\Phi^*(0) = 0$. But, in general, $\Phi^*$ is not a Young function: consider the case when $\Phi(0) = 0$ and $\Phi(x) = \infty$ for all $x \neq 0$, then $\Phi$ is a Young function, but its conjugate is identically equal to 0 on $V^*$ and the condition (iv) is not satisfied.

Let $(\text{dom}(\Phi))^\perp$ be the orthogonal subspace to $\text{dom}(\Phi)$ in $V^*$, that is

$$(\text{dom}(\Phi))^\perp := \{v \in V^*, \ v(x) = 0 \text{ for all } x \in \text{dom}(\Phi)\}$$

Then $(\text{dom}(\Phi))^\perp$ is a closed subspace in $V^*$. Let $\tilde{V}$ be the quotient space $\tilde{V} = V^*/(\text{dom}(\Phi))^\perp$. If $u$ and $v$ are elements in the same equivalence class, then

$$\Phi^*(v) = \sup_{x \in \text{dom}(\Phi)} \{v(x) - \Phi(x)\} = \sup_{x \in \text{dom}(\Phi)} \{u(x) - \Phi(x)\} = \Phi^*(u)$$

and $\Phi^*$ is well defined as a function on $\tilde{V}$.

Lemma 3.4. $\Phi^* : \tilde{V} \to \mathbb{R} \cup \{\infty\}$ is a Young function.

Proof. It is easy to see that $\Phi^*$ satisfies (i), (ii) and (iii) from the definition of a Young function. Moreover, it follows from the definition of the conjugate function that

$$|v(x)| \leq \Phi(x) + \Phi^*(v), \text{ for all } x \in V, v \in \tilde{V}$$

Let $v \in \tilde{V}$, $v \neq 0$. Then there is an element $x \in \text{dom}(\Phi)$ such that $v(x) \neq 0$. It follows that $\Phi^*(tv) \geq |tv(x)| - \Phi(x)$ for all $t$ and (iv) is satisfied. □

We will define $C_{\Phi^*}$, $L_{\Phi^*}$, $\| \cdot \|_{\Phi^*}$ and $B_{\Phi^*}$ in the same way as for $\Phi$. 
Lemma 3.5. (Hölder inequality).

$$|v(x)| \leq 2\|x\|_\Phi \|v\|_{\Phi^*} \quad \text{for all } x \in B_\Phi, v \in B_{\Phi^*}.$$

Proof. Let $x \in C_\Phi$, $v \in C_{\Phi^*}$, then by $\text{Lemma 3.2}$

$$|v(x)| \leq \Phi(x) + \Phi^*(v) \leq 2 \sup_{v \in \tilde{V}} \{v(x) - \Phi^*(v)\} \leq 2 \sup_{v \in \tilde{V}^*} \{v(x) - \Phi^*(v)\} = \Phi^{**}(x).$$

Similarly, if $x \in L_\Phi$, $v \in L_{\Phi^*}$. By Lemma 3.2 $\frac{x}{\|x\|_\Phi} \in C_\Phi$, $\frac{v}{\|v\|_{\Phi^*}} \in C_{\Phi^*}$ and therefore $|v(x)| \leq 2\|x\|_\Phi \|v\|_{\Phi^*}$. Clearly, the inequality extends to $x \in B_\Phi, v \in B_{\Phi^*}$. □

The second conjugate. If $E$ is a Banach space and $H \subset E$ is a closed subspace, then the dual of the quotient space $(E/H)^*$ can be identified with $H^\perp$. It follows that $\tilde{V}^* \cap V = (\text{dom}(\Phi))^\perp$, which is nothing else than the closure of $L_\Phi$ in $V$. Let us denote this space by $\tilde{V}$.

As before, we can find the conjugate function to $\Phi^*: \tilde{V} \to \mathbb{R} \cup \{+\infty\}$ with respect to the pair $(\tilde{V}, \tilde{V}^*)$. Note that for $x \in \tilde{V}$, we have

$$\sup_{v \in \tilde{V}^*} \{v(x) - \Phi^*(v)\} = \sup_{v \in \tilde{V}^*} \{v(x) - \Phi^*(v)\} = \Phi^{**}(x),$$

where $\Phi^{**}$ is the second conjugate to $\Phi: V \to \mathbb{R} \cup \{+\infty\}$. Since $\Phi$ is convex and lower semicontinuous, $\Phi^{**}(x) = \Phi(x)$ on $V$, $\tilde{V}$. It follows in particular that the restriction of $\Phi^{**}$ to $\tilde{V}$ is a Young function.

It is clear from Hölder inequality that any $x \in L_\Phi$ defines a bounded linear functional on $B_{\Phi^*}$. Let $\|x\|_{\Phi^*}$ be its norm in $B_{\Phi^*}$, then by Lemma 3.2

$$\|x\|_{\Phi^*} = \sup\{|v(x)|, \Phi^*(v) \leq 1\}. $$

Similarly, if $v \in L_{\Phi^*}$, then $v \in B_\Phi$ and we have

$$\|v\|_\Phi = \sup\{|v(x)|, \Phi(x) \leq 1\}.$$ 

Lemma 3.6. For $x \in L_\Phi$, we have $\|x\|_\Phi \leq \|x\|_{\Phi^*} \leq 2\|x\|_\Phi$. Similarly, if $v \in L_{\Phi^*}$, then $\|v\|_{\Phi^*} \leq \|v\|_\Phi \leq 2\|v\|_{\Phi^*}$.

Proof. Let $v \in L_{\Phi^*}$. By Hölder inequality, $\|v\|_{\Phi^*} \leq 2\|v\|_{\Phi^*}$. Let now $\|v\|_\Phi = 1$, then for $x \in C_\Phi$ we have

$$v(x) - \Phi(x) \leq 1.$$

On the other hand, for $x \in \text{dom}(\Phi)$, such that $\Phi(x) > 1$, we get from Lemma 3.2

$$v(x) - \Phi(x) \leq v(x) - \|x\|_\Phi \leq 0.$$

It follows that $\Phi^*(v) \leq 1$ and $v \in C_{\Phi^*}$, hence $\|v\|_{\Phi^*} \leq 1$. Therefore, $\|v\|_{\Phi^*} \leq \|v\|_\Phi$ for all $v \in L_{\Phi^*}$. The proof for $x \in L_\Phi$ is the same, using the fact that $\Phi$ is the conjugate of $\Phi^*$.

Proposition 3.1. $B_{\Phi^*} \subseteq B_\Phi$ and $L_{\Phi^*} = \tilde{V} \cap B_{\Phi^*}$. Similarly, $B_\Phi \subseteq B_{\Phi^*}$ and $L_\Phi = \tilde{V} \cap B_{\Phi^*}$.

Proof. As we have seen, $L_{\Phi^*}$ is a vector subspace in $B_{\Phi^*}$ and the norms in $L_{\Phi^*}$ and $B_{\Phi^*}$ are equivalent, hence $B_{\Phi^*} \subseteq B_{\Phi^*}$. Let now $v \in \tilde{V} \cap B_{\Phi^*}$ be such that $\|v\|_{\Phi^*} = 1$. Then $\Phi^*(v) \leq 1$, exactly as in the proof of Lemma 3.2. It follows that for all $v \in \tilde{V} \cap B_{\Phi^*}$, $\Phi^*(v/\|v\|_{\Phi^*}) \leq 1 < \infty$ and $v \in L_{\Phi^*}$. Again, the proof for $L_{\Phi^*}$ and $B_\Phi$ is the same. □
Let $\Phi$ be a Young function such that $0$ is an interior point in $\text{dom}(\Phi)$. Then the function $\Phi$ is continuous in $0$, therefore there is an open set $U$ containing $0$ such that $U \subset C_\Phi$. It follows that $C_\Phi$ is a neighborhood of $0$ in $V$, hence it is absorbing in $V$

\[(9) \quad V = \bigcup_n nC_\Phi = L_\Phi \quad \text{(as sets)}\]

Since $C_\Phi$ is a convex body (that is, $0$ is a topological interior point), its Minkowski functional $\| \cdot \|_\Phi$ is continuous with respect to the original norm \cite{11}, p. 182. It follows that we have the continuous inclusion $V \subseteq B_\Phi$. Further, since $\text{dom}(\Phi)$ has non-empty interior, $(\text{dom}(\Phi))^\perp = \{0\}$ and $\bar{V} = V^*$. Clearly also $\bar{V} = V$.

**Proposition 3.2.** Let $0 \in \text{int} \, \text{dom}(\Phi)$. Then $V \subseteq B_\Phi \subseteq B_{\Phi^*} = B_{\Phi^*} \subseteq V^*$.

**Proof.** By \cite{8}, each $x \in V$ is in $L_\Phi$, and by continuity, $\|x\|_\Phi \leq K\|x\|$, for some fixed $K > 0$. Let $v \in B_{\Phi^*}$, then

\[|\varphi(x)| \leq \|v\|_\Phi \|x\|_\Phi \leq K\|v\|_\Phi \|x\| \quad \text{for } x \in V\]

hence $v \in V^* = \bar{V}$ and $\|v\|^* \leq K\|v\|_\Phi$. The statement now follows from Proposition 3.1. $\square$

4. The spaces $B_\varphi$ and $B_{\varphi,0}$.

Let $M_\varphi$ be the real Banach subspace of self-adjoint elements in $M$, then the dual $M_{\varphi}^*$ is the subspace of hermitian (not necessarily normal) functionals in $M^*$. We define the functional $F_{\varphi} : M_{\varphi}^* \to \mathbb{R} \cup \{\infty\}$ by

\[F_{\varphi}(\omega) = \begin{cases} S(\omega, \varphi) & \text{if } \omega \in \mathfrak{G}_\varphi \\ \infty & \text{otherwise} \end{cases}\]

Then $F_{\varphi}$ is convex and lower semicontinuous, with $\text{dom}(F_{\varphi}) = S_\varphi$. It follows from \cite{11} that $F_{\varphi}$ is strictly convex. Its conjugate $F_{\varphi}^*$ is

\[F_{\varphi}^*(h) = \sup_{\omega \in \mathfrak{G}_\varphi} \{\omega(h) - F_{\varphi}(\omega)\} = c_\varphi(h), \quad h \in M_{\varphi}\]

Hence $c_\varphi$ is convex and lower semicontinuous, in fact, since finite valued, it is continuous on $M_{\varphi}$. We have $c_{\varphi}^* = F_{\varphi}^{**} = F_{\varphi}$ on $M_{\varphi}^*$. Note also that

\[(10) \quad c_\varphi(h + \lambda) = c_\varphi(h) + \lambda, \quad \forall \lambda \in \mathbb{R}\]

We define another convex and lower semicontinuous functional on $M_{\varphi}^*$, namely,

\[\tilde{F}_{\varphi}(\omega) = \begin{cases} S(\omega, \varphi) - \omega(1) & \text{if } \omega \in M_{\varphi}^+ \\ \infty & \text{otherwise} \end{cases}\]

Then the conjugate functional is

\[\tilde{F}_{\varphi}^*(h) = \sup_{\omega \in M_{\varphi}^*} \{\omega(h) - S(\omega, \varphi) + \omega(1)\} = \sup_{\omega \in \mathfrak{G}_\varphi, \lambda \in \mathbb{R}^+} \{\lambda\omega(h) - S(\lambda\omega, \varphi) + \lambda\} =
\]

\[= \sup_{\omega \in \mathfrak{G}_\varphi, \lambda \in \mathbb{R}^+} \{\lambda(\omega(h) - S(\omega, \varphi)) - \lambda \log \lambda + \lambda\} =
\]

\[= \sup_{\lambda \in \mathbb{R}^+} \{\lambda(c_\varphi(h) + 1) - \lambda \log \lambda\} = e^{c_\varphi(h)} = \varphi^h(1)\]
Again, \( h \mapsto \varphi^h(1) \) is convex and continuous and \( \bar{F}_\varphi^* = \bar{F}_\varphi \). Next, we define a Young function on \( \mathcal{M}_s \).

Let \( \Phi_\varphi : \mathcal{M}_s \to \mathbb{R}^+ \) be defined by
\[
\Phi_\varphi(h) = \frac{\varphi^h(1) + \varphi^{-h}(1)}{2} - 1
\]

**Lemma 4.1.** \( \Phi_\varphi \) is a Young function.

**Proof.** The property (i) from the definition of a Young function follows from the properties of \( h \mapsto \varphi^h(1) \). Since \( \varphi^h(1) = e^{c_\varphi(h)} \geq e^{\omega(h) - S(\omega, \varphi)} \) for all normal states \( \omega \), we have
\[
\Phi_\varphi(h) \geq \cosh(\omega(h))e^{-S(\omega, \varphi)} - 1
\]
In particular,
\[
\Phi_\varphi(h) \geq \cosh(\varphi(h)) - 1 \geq 0 \quad \text{for all } h
\]
Since obviously \( \Phi_\varphi(0) = 0 \), (ii) follows. Let now \( h \) be such that \( \omega(h) = 0 \) for all \( \omega \in \mathcal{S}_\varphi \), then by definition, \( c_\varphi(h) = 0 \) and \( \varphi = \varphi^h \), hence \( h = 0 \). Therefore if \( h \neq 0 \), then there is a state \( \omega \in \mathcal{S}_\varphi \) such that \( \omega(h) \neq 0 \) and then \( \lim_{t \to \infty} \cosh(t\omega(h)) = \infty \), this implies (iv). Property (iii) is obviously satisfied. \( \square \)

Let \( C_\varphi := C_{\Phi_\varphi}, B_\varphi := B_{\Phi_\varphi} \) and \( \| \cdot \|_\varphi := \| \cdot \|_{\Phi_\varphi} \). Since \( \text{dom} \Phi_\varphi = \mathcal{M}_s \), we have by Proposition 3.2 that \( \mathcal{M}_s \subseteq B_\varphi \). If \( \Phi_\varphi^* \) is the conjugate of \( \Phi_\varphi \), then \( B_{\Phi_\varphi}^* = B_{\Phi_\varphi} \subseteq \mathcal{M}_s^* \).

Let now \( h \in \mathcal{M}_s \), such that \( \| h \|_\varphi = t > 0 \), that is,
\[
\Phi_\varphi(h) \leq \frac{t}{1}
\]
If \( \omega \) is a state, then by (11),
\[
\cosh(\omega(h) t) \leq 2e^{S(\omega, \varphi)}
\]
If \( \omega \in \mathcal{S}_\varphi \), then \( |\omega(h)| \leq ct \), where \( c > 0 \) is some constant depending on \( S(\omega, \varphi) \). It follows that each \( \omega \in \mathcal{S}_\varphi \) extends to a continuous linear functional on \( B_\varphi \). Moreover, for \( C > 0, \mathcal{S}_C \) is an equicontinuous subset in \( B_\varphi^* \).

Let \( \mathcal{M}_{s,0} \subseteq \mathcal{M}_s \) be the subspace of elements satisfying \( \varphi(h) = 0 \). Then by putting \( \omega = \varphi \) in (11), we get
\[
c_\varphi(h) = S(\varphi, [\varphi^h]) \geq 0
\]
Let us define
\[
\Phi_{\varphi,0}(h) = \frac{c_\varphi(h) + c_\varphi(-h)}{2}, \quad h \in \mathcal{M}_{\varphi,0}
\]
Then it is easy to check that \( \Phi_{\varphi,0} \) is a Young function on \( \mathcal{M}_{\varphi,0} \). We have

**Lemma 4.2.** Let \( h \in \mathcal{M}_{s,0} \). Then
\[
\Phi_{\varphi,0}(h) \leq \Phi_{\varphi}(h) \leq e^{2\Phi_{\varphi,0}} - 1
\]

**Proof.** The first inequality follows from \( a \leq e^a - 1 \) for \( a \geq 0 \), the second follows from \( x + y \leq 2xy \) for \( x, y \geq 1 \). \( \square \)

Let us construct the Banach space \( B_{\Phi_{\varphi,0}} := B_{\varphi,0} \) and let \( \| \cdot \|_{\varphi,0} := \| \cdot \|_{\Phi_{\varphi,0}} \).
Proposition 4.1. The norms \( \| \cdot \|_{\varphi,0} \) and \( \| \cdot \|_{\varphi} \) are equivalent on \( M_{s,0} \).

Proof. Let us denote \( C_{\varphi,0} := C_{\Phi_{\varphi,0}} \). We show that

\[
\frac{1}{2} \log 2 C_{\varphi,0} \subseteq C_{\varphi} \cap M_{s,0} \subseteq C_{\varphi,0}
\]

Let \( h \in C_{\varphi,0} \) and \( t = \frac{1}{2} \log 2 \). Then by convexity, \( \Phi_{\varphi,0}(th) \leq t = \frac{1}{2} \log 2 \) and hence

\[
\Phi_{\varphi}(th) \leq e^{2\Phi_{\varphi,0}(th)} - 1 \leq 1,
\]

which implies \( tC_{\varphi,0} \subseteq C_{\varphi} \cap M_{s,0} \). The other inclusion follows from the first inequality in Lemma 4.2. It follows from (14) that for \( h \in M_{s,0} \),

\[
\|h\|_{\varphi,0} \leq \|h\|_{\varphi} \leq \frac{2}{\log 2} \|h\|_{\varphi,0}
\]

□

Note that since \( \varphi \in S_{\varphi} \), \( \varphi \) extends to a bounded linear functional on \( B_{\varphi} \). Clearly, the completion of \( M_{s,0} \) under the norm \( \| \cdot \|_{\varphi} \) is the Banach subspace \( \{ h \in B_{\varphi}, \varphi(h) = 0 \} \). It follows from the above Proposition that \( B_{\varphi,0} \) can be identified with the subspace of centered elements in \( B_{\varphi} \).

5. Extension of \( c_{\varphi} \).

Since \( S_{\varphi} \subset B_{\varphi}^* \subset M_{s}^* \), the restriction of \( F_{\varphi} \) is a strictly convex lower semicontinuous functional on \( B_{\varphi}^* \), with effective domain \( S_{\varphi} \). Its conjugate \( F_{\varphi}^* \) is a lower semicontinuous extension of \( c_{\varphi} \) to \( B_{\varphi} \), moreover, \( F_{\varphi}^{**} = F_{\varphi} \). We show that this extension has again values in \( \mathbb{R} \) and is continuous.

Lemma 5.1. Let the sequence \( \{ h_n \}_{n} \subset M_{s} \) be Cauchy in the norm \( \| \cdot \|_{\varphi} \). Then the sequences \( \{ c_{\varphi}(h_n) \}_{n} \) and \( \{ S(\varphi h_n, \varphi) \}_{n} \) are bounded.

Proof. By (5), we have for all \( n \)

\[
\varphi(h_n) \leq c_{\varphi}(h_n)
\]

Since \( \varphi(h_n) \) converges, \( c_{\varphi}(h_n) \) is bounded from below. Further, let \( n_0 \) be such that \( \|h_n - h_{n_0}\|_{\varphi} < 1 \) for all \( n \geq n_0 \), then

\[
\omega(h_n) - S(\omega, \varphi) \leq c_{\varphi}(h_n - h_{n_0}) \leq \|h_{n_0}\| + \log 2
\]

for all such \( n \) and \( \omega \in S_{\varphi} \). It follows that \( \{ c_{\varphi}(h_n) \}_{n} \) is bounded.

If \( \{ h_n \}_{n} \) is Cauchy, then the sequence \( \{ th_n \}_{n} \) is also Cauchy for all \( t \in \mathbb{R} \) and there are constants \( A_t, B_t \), such that

\[
A_t \leq c_{\varphi}(th_n) \leq B_t, \quad \forall n
\]

On the other hand , we have

\[
\frac{d}{dt} c_{\varphi}(th_n)|_{t=1} = [\varphi h_n](h_n)
\]

By convexity,

\[
c_{\varphi}(th_n) \geq c_{\varphi}(h_n) + (t - 1) \frac{d}{dt} c_{\varphi}(th_n)|_{t=1} \geq A_1 + (t-1)[\varphi h_n](h_n)
\]
For arbitrary fixed $t > 1$, we get
\[ [\varphi^{h_n}](h_n) \leq \frac{B_t - A_1}{t - 1}, \quad \forall n \]
Boundedness of $S([\varphi^{h_n}], \varphi)$ now follows from
\[ 0 \leq S([\varphi^{h_n}], \varphi) = [\varphi^{h_n}](h_n) - c_\varphi(h_n). \]

\[ \square \]

**Theorem 5.1.** Let $\{h_n\}_n$ be a sequence in $\mathcal{M}_s$, converging to some $h$ in $B_{\varphi}$. Then
\begin{equation}
\lim_n c_\varphi(h_n) = \sup_{\omega \in S_{\varphi}} \{ \omega(h) - S(\omega, \varphi) \}
\end{equation}
and there is a unique state $\psi \in S_{\varphi}$ such that the supremum is attained. The state $\psi$ is faithful. Moreover, \( \lim_n S([\varphi^{h_n}], \varphi) = S(\psi, \varphi) \), \( \lim_n [\varphi^{h_n}(h_n)] = \psi(h) \) and \( \lim_n S(\psi, [\varphi^{h_n}]) = 0 \). In particular, $[\varphi^{h_n}]$ converges to $\psi$ in norm.

Proof. This proof is similar to the proof of Theorem 12.3. in [3].

By Lemma 5.1 there is some $C > 0$ such that $[\varphi^{h_n}] \in S_{\varphi, C}$ for all $n$. The set $S_{\varphi, C}$ is weakly relatively compact in $\mathfrak{S}$, and hence there is subsequence $[\varphi^{h_{n_k}}]$ converging weakly to a state $\psi \in S_{\varphi, C}$. We will show that $[\varphi^{h_{n_k}}](h_{n_k})$ converges to $\psi(h)$.

Since $S_{\varphi}$ is an equicontinuous subset in $B_{\varphi}^*$, $\omega(h_n)$ converges to $\omega(h)$ uniformly for all $\omega \in S_{\varphi, C}$. This implies
\[ ||[\varphi^{h_{n_k}}](h_{n_k}) - [\varphi^{h_{n_k}}](h)|| < \varepsilon \]
for sufficiently large $n_k$. We further have
\begin{align*}
||[\varphi^{h_{n_k}}](h) - \psi(h)|| & \leq ||[\varphi^{h_{n_k}}](h) - [\varphi^{h_{n_k}}](h_m)|| \\
& \quad + \ ||[\varphi^{h_{n_k}}](h_m) - \psi(h_m)|| + ||\psi(h_m) - \psi(h)|| < \varepsilon
\end{align*}
for sufficiently large $m$ and $n_k$. Putting both inequalities together, we get $[\varphi^{h_{n_k}}](h_{n_k}) \rightarrow \psi(h)$.

Let $\omega \in S_{\varphi}$. By definition,
\[ [\varphi^{h_{n_k}}](h_{n_k}) - S([\varphi^{h_{n_k}}], \varphi) = c_\varphi(h_{n_k}) \geq \omega(h_{n_k}) - S(\omega, \varphi) \]
By weak lower semicontinuity of the relative entropy, we get
\begin{equation}
\psi(h) - S(\psi, \varphi) \geq \limsup c_\varphi(h_{n_k}) \geq \omega(h) - S(\omega, \varphi)
\end{equation}
and thus $\psi$ is a maximizer of (15). On the other hand,
\[ \psi(h_{n_k}) - S(\psi, \varphi) \leq [\varphi^{h_{n_k}}](h_{n_k}) - S([\varphi^{h_{n_k}}], \varphi) = c_\varphi(h_{n_k}) \]
From this and (16), it follows that $\psi(h) - S(\psi, \varphi) = \lim c_\varphi(h_{n_k})$. It also follows that
\[ \limsup S([\varphi^{h_{n_k}}], \varphi) \leq S(\psi, \varphi) \]
and this, together with lower semicontinuity implies that $S([\varphi^{h_{n_k}}], \varphi)$ converges to $S(\psi, \varphi)$. 


To show that such $\psi$ is unique, suppose that $\psi'$ is another maximizer, then for $\psi_\lambda := \lambda \psi + (1 - \lambda)\psi'$, $0 \leq \lambda \leq 1$, we have

$$\psi(h) - S(\psi, \varphi) \geq \psi_\lambda(h) - S(\psi_\lambda, \varphi) = \lambda S(\psi, \varphi) - (1 - \lambda) S(\psi', \varphi) = \psi(h) - S(\psi, \varphi)$$

hence $\psi_\lambda$ is a maximizer as well and, moreover,

$$S(\psi_\lambda, \varphi) = \lambda S(\psi, \varphi) + (1 - \lambda) S(\psi', \varphi)$$

By strict convexity, this implies that $\psi = \psi'$. It also follows that the whole sequence $[\varphi^{h_n}]$ converges weakly to $\psi$.

Using (9), we have

$$S(\varphi, \psi) \leq \liminf_n S(\varphi, [\varphi^{h_n}]) = \lim_n c_\varphi(h_n) - \varphi(h) < \infty$$

This implies that $\supp \varphi \leq \supp \psi$ and $\psi$ is faithful. Finally, by taking the limit in the equality,

$$\psi(h_n) - S(\psi, \varphi) = c_\varphi(h_n) - S(\psi, [\varphi^{h_n}])$$

we get $\lim_n S(\psi, [\varphi^{h_n}]) \to 0$. □

**Corollary 5.1.** Let $h_n$ be a sequence in $B_\varphi$, then $h_n \to 0$ if and only if $c_\varphi(th_n) \to 0$ for all $t \in \mathbb{R}$.

**Proof.** Let $h_n$ be such that $c_\varphi(th_n) = \log \varphi^{th_n}(1)$ converges to 0, then $\varphi^{th_n}(1)$ converges to 1, for all $t \in \mathbb{R}$. Therefore, for each $\varepsilon > 0$, $\Phi_\varphi(h_n) < 1$ for large enough $n$, that is, $\|h_n\|_\varphi \to 0$. The converse follows from Theorem 5.1. □

In particular, if $h_n \in M_s$ is a sequence converging strongly to $h$, then $h_n$ converges to $h$ in $\|\cdot\|_\varphi$, see [3].

6. The dual spaces.

The dual space $M_{*0}$ is obtained as the quotient space $M_s^*/\{\varphi\}$. Each equivalence class in $M_{*0}$ can be identified with its unique element $v$ satisfying $v(1) = 0$. By Proposition 4.2, we have $B_{\varphi,0} = B_{\varphi^*} \subseteq M_{*0}$. By Proposition 4.1, $B_{\varphi,0}$ is the same as $B_{\varphi}^*/\{\varphi\}$.

**Lemma 6.1.** Let $\tilde{c}_\varphi$ be the restriction of $c_\varphi$ to $B_{\varphi,0}$. Then the conjugate functional is $c_\varphi^*(v) = F_\varphi(v + \varphi)$.

**Proof.** Let $v \in B_{\varphi}^*$, $v(1) = 0$. Then by (10),

$$F_\varphi(v + \varphi) = \sup_{h \in B_{\varphi}} \{v(h) + \varphi(h) - c_\varphi(h)\} = \sup_{h \in B_{\varphi}} \{v(h - \varphi(h)) - \tilde{c}_\varphi(h - \varphi(h))\} = \tilde{c}_\varphi^*(v).$$

□

Let $V$ be a Banach space and $V^*$ its dual. For any subset $D \subset V$, let $D^\circ$ be the polar of $D$ in $V^*$, that is, $D^\circ = \{v \in V^*, \ v(h) \leq 1, \ \forall h \in D\}$. We will need the following lemma.
Lemma 6.2. Let $F : V \to \mathbb{R}^+$ be a convex functional such that $F(0) = 0$ and let $F^*$ be its conjugate. Let $D = \{ x \in V, \ F(x) \leq 1 \}$ and $D^* = \{ v \in V^*, \ F^*(v) \leq 1 \}$. Then

$$\frac{1}{2} D^* \subseteq D^\circ \subseteq D^*$$

Proof. If $v \in D^*$, then $v(x) \leq F(x) + F^*(v) \leq 2$ for all $x \in D$ and therefore $\frac{1}{2} v \in D^\circ$. Let $v \in D^\circ$, then

$$v(x) - 1 \leq 0 \leq F(x) \quad \text{for } x \in D$$

If $F(x) > 1$, then by continuity there is some $t \in (0,1)$ such that $F(tx) = 1$. Since $tx \in D$, $v(tx) \leq 1$, moreover, by convexity, $1 = F(tx) \leq tF(x)$. Consequently,

$$v(x) - 1 \leq \frac{1}{t} - 1 \leq F(x)$$

It follows that $F^*(v) \leq 1$ and $v \in D^*$.

Proposition 6.1. Let $v$ be an element in $K^\circ_{\varphi, 0}$. Then there are states $\omega_1, \omega_2$, satisfying $S(\omega_1, \varphi) + S(\omega_2, \varphi) \leq 1$, such that $v = \omega_1 - \omega_2$.

Proof. Since $\tilde{e}_\varphi$ is continuous on $B_{\varphi, 0}$, the set $D := \{ h \in B_{\varphi, 0}, \ \tilde{e}_\varphi(h) \leq 1 \}$ is closed. Let us endow the dual pair $B_{\varphi, 0}$ and $B^*_{\varphi, 0}$ with the $\sigma(B_{\varphi, 0}, B^*_{\varphi, 0})$ and $\sigma(B^*_{\varphi, 0}, B_{\varphi, 0})$ topology, respectively. As $D$ is convex, it is closed also in this weaker topology. The set $D \cap -D$ is absolutely convex and closed, moreover,

$$(17) \quad D \cap -D \subseteq K_{\varphi, 0} \subseteq 2(D \cap -D),$$

as can be easily checked. Then

$$\frac{1}{2}(D \cap -D)^\circ \subseteq K^\circ_{\varphi, 0} \subseteq (D \cap -D)^\circ$$

By [4], $(D \cap -D)^\circ$ is the closed convex cover of $D^\circ \cup -D^\circ$, which is the same as the closed absolutely convex cover of $D^\circ$. Moreover, since $D^\circ$ is the polar of a neighborhood of $0$, it is compact [6]. Therefore its absolutely convex cover is also compact, hence closed. It follows that $(D \cap -D)^\circ$ is the absolutely convex cover of $D^\circ$.

By Lemma 6.1 and 6.2

$$\frac{1}{2}(S_1 - \varphi) \subseteq D^\circ \subseteq S_1 - \varphi$$

and this implies

$$(18) \quad \frac{1}{4} \text{abs conv } (S_1 - \varphi) \subseteq K^\circ_{\varphi, 0} \subseteq \text{abs conv } (S_1 - \varphi)$$

Let now $v \in \text{abs conv } (S_1 - \varphi)$, then there are elements $\varphi_1, \ldots, \varphi_n \in S_1$, and real numbers $\lambda_1, \ldots, \lambda_n$, $\sum_n |\lambda_n| = 1$, such that $v = \sum_n \lambda_n (\varphi_n - \varphi)$. Let $m \leq n$ be such that $\lambda_i > 0$ for $i \leq m$ and $\lambda_i < 0$ for $i > m$. Then $v = \omega_1 - \omega_2$, with

$$\omega_1 = \sum_{i=1}^m \lambda_i \varphi_i + (1 - \lambda) \varphi, \quad \omega_2 = \sum_{i=m+1}^n |\lambda_i| \varphi_i + \lambda \varphi$$
where $\lambda = \sum_{i=1}^{m} \lambda_i$, moreover, $S(\omega_1, \varphi) \leq \sum_{i=1}^{m} \lambda_i S(\varphi_i, \varphi) \leq \lambda$, and similarly, $S(\omega_2, \varphi) \leq 1 - \lambda$.

**Theorem 6.1.**

(i) $B^*_\varphi = \mathcal{P}_\varphi - \mathcal{P}_\varphi$ and $B^*_\varphi \cap M^+_\varphi = \mathcal{P}_\varphi$.

(ii) $B^*_\varphi, 0 = \cup_{n}(S_1 - S_1)$.

**Proof.** (i) Let $\omega \in B^*_\varphi$ and let $v = \omega - \omega(1)\varphi$. Then $v$ can be seen as an element in $B^*_\varphi, 0$. Let $\|v\|_{P, 0}^* = t$, then by Proposition 6.1 there are $\omega_1, \omega_2 \in S_1$, such that $\frac{1}{\varphi} = \omega_1 - \omega_2$, that is, $\omega = t\omega_1 + \omega(1)\varphi - t\omega_2$. Since $\omega_1, \omega_2, \varphi \in \mathcal{P}_\varphi$ and $\mathcal{P}_\varphi$ is a convex cone, it follows that $B^*_\varphi \subseteq \mathcal{P}_\varphi - \mathcal{P}_\varphi$. On the other hand, we have already shown that if $\omega \in \mathcal{S}_\varphi$, then $\omega \in B^*_\varphi$ and hence $\mathcal{P}_\varphi - \mathcal{P}_\varphi \subseteq B^*_\varphi$. Let $\omega \in B^*_\varphi \cap M^+_\varphi$, then we get $\omega + t\omega_2 = t\omega_1 + \omega(1)\varphi$. It follows that $\omega + t\omega_2 \in \mathcal{P}_\varphi$, and Donald’s identity implies that $\omega$ must be in $\mathcal{P}_\varphi$.

(ii) By Proposition 6.1

$$K^0_{\varphi, 0} \subseteq (S_1 - S_1) \subseteq 4K^0_{\varphi, 0}.$$ The equality now follows from the fact that the closed unit ball is absorbing in $B^*_\varphi, 0$.

In the rest of this section, we find an equivalent norm on $B^*_\varphi, 0$.

We define a function $f: \mathcal{S}_* \times \mathcal{S}_* \rightarrow \mathbb{R}^+$ by

$$f(\omega_1, \omega_2) = S(\omega_1, \varphi) + S(\omega_2, \varphi).$$

Clearly, $f$ is weakly lower semicontinuous and strictly convex. Further, let $v \in \mathcal{S}_* - \mathcal{S}_*$ and let $L_v = \{(\omega_1, \omega_2) \in \mathcal{S}_* \times \mathcal{S}_*, \omega_1 - \omega_2 = v\}$. Then $L_v$ is a weakly closed subset in $M_\varphi \times M_\varphi$.

**Lemma 6.3.** Let $v \in \mathcal{S}_\varphi - \mathcal{S}_\varphi$. Then the function $f$ attains its minimum over $L_v$ at a unique point $(v_+, v_-) \in L_v$.

**Proof.** By assumptions, $v = \omega_1 - \omega_2$ for some $\omega_1, \omega_2 \in \mathcal{S}_\varphi$. Let $C > 0$ be such that $\omega_1, \omega_2 \in \mathcal{S}_C$, then the infimum is taken over the set $L_v \cap \mathcal{S}_C \times \mathcal{S}_C$. Since $L_v$ is weakly closed and $\mathcal{S}_C$ is weakly compact, the intersection is weakly compact and $f$ attains its minimum on it. Uniqueness follows by strict convexity of $f$.

Let us now define the functional $\Psi_{\varphi, 0}: \mathcal{M}_{\varphi, 0}^* \rightarrow \mathbb{R}^+$ by

$$\Psi_{\varphi, 0}(v) = \begin{cases} f(v_+, v_-) & \text{if } v \in \mathcal{S}_\varphi - \mathcal{S}_\varphi \\ \infty & \text{otherwise} \end{cases}$$

**Lemma 6.4.** $\Psi_{\varphi, 0}$ is a Young function.

**Proof.** It is easy to check that $\Psi_{\varphi, 0}$ is convex, positive, $\Psi_{\varphi, 0}(v) = \Psi_{\varphi, 0}(-v)$ and that $\Psi_{\varphi, 0}(v) = 0$ if and only if $v = 0$. We will show that $\Psi_{\varphi, 0}$ is lower semicontinuous.

To do this, we have to prove that for any $C > 0$, the set of all $v$ satisfying $\Psi_{\varphi, 0}(v) \leq C$ is closed. Let $v_n$ be a sequence of elements in this set, converging to some $v$. Let $v_n = v_{n+} - v_{n-}$ be the corresponding decompositions, then $v_{n+}, v_{n-} \in \mathcal{S}_C$ for all $n$, hence there are elements $v_{n+}'$, $v_{n-}'$ in $\mathcal{S}_C$ and a subsequence $v_{n_k}$ such that $v_{n_k} \rightarrow v_{n+}'$ and $v_{n_k} \rightarrow v_{n-}'$ weakly. It follows that $v = v_{n+}' - v_{n-}'$ and $\Psi_{\varphi, 0}(v) \leq S(v_{n+}', \varphi) + S(v_{n-}', \varphi) \leq \liminf S(v_{n_k+}, \varphi) + S(v_{n_k-}, \varphi) \leq C$. The equality now follows from the fact that the closed unit ball is absorbing in $B^*_\varphi, 0$.
Suppose that \( v \neq 0 \), then \( \Psi_{\varphi,0}(v) > 0 \). If \( t > 1 \), then by convexity, \( t\Psi_{\varphi,0}(v) \leq \Psi_{\varphi,0}(tv) \), hence \( \lim_{t \to \infty} \Psi_{\varphi,0}(tv) = \infty \). \( \square \)

Let us find the corresponding Banach space. Note that

\[
C_{\Psi_{\varphi,0}} = \{ \omega_1 - \omega_2 : \omega_1, \omega_2 \in \Theta_{S_0}, S(\omega_1, \varphi) + S(\omega_2, \varphi) \leq 1 \}.
\]

By Proposition 6.1, this implies that \( K_{\Psi,0}^\ast \subseteq C_{\Psi_{\varphi,0}} \subseteq S_1 - S_1 \) and by Theorem 6.1 (ii), \( B_{\varphi,0}^\ast \subseteq L_{\Psi_{\varphi,0}} \subseteq B_{\varphi,0}^\ast \).

**Proposition 6.2.** \( \| \cdot \|_{\Psi_{\varphi,0}} \) defines an equivalent norm in \( B_{\varphi,0}^\ast \).

**Proof.** Let \( \Psi_{\varphi,0}^\ast : M_* \to \mathbb{R} \) be the conjugate functional, then

\[
\Psi_{\varphi,0}^\ast(h) = \sup_{v \in M_{\varphi,0}^\ast} v(h) - \Psi_{\varphi,0}(v) = \sup_{v \in S_{\varphi}^\ast - S_{\varphi}} \sup_{(\omega_1, \omega_2) \in L_v} \omega_1(h) - \omega_2(h) - f(\omega_1, \omega_2) = \sup_{\omega_1, \omega_2 \in S_{\varphi}} \omega_1(h) - S(\omega_1, \varphi) + S(\omega_2, \varphi) = 2\Phi_{\varphi,0}(h)
\]

It follows that \( \Psi_{\varphi,0}(v) = \Psi_{\varphi,0}^\ast(v) = 2\Phi_{\varphi,0}(\frac{1}{2}v) \). Since the norms \( \| \cdot \|_{\ast}^\ast \) and \( \| \cdot \|_{\Psi_{\varphi,0}} \)
are equivalent, this finishes the proof. \( \square \)

7. The chain rule.

**Proposition 7.1.** Let \( h \in B_{\varphi}, k \in M_* \). Then \( [\varphi^{h+k}] = ([\varphi^h]^k], c_\varphi(h + k) = c_\varphi(h)|c_\varphi^k(k) + c_\varphi(k) \) and for all normal states \( \omega \) the equality

\[
\omega(k) - S(\omega, [\varphi^h]) = c_\varphi(h + k) - c_\varphi(h) - S(\omega, [\varphi^{h+k}])
\]

holds.

**Proof.** Let \( h_n \in M_* \) be such that \( h_n \to h \) in \( B_{\varphi} \). By the chain rule (7), we have \([\varphi^{h_n+k}] = ([\varphi^{h_n}]^k] \) and \( c_\varphi(h_n + k) = c_\varphi(h_n)|c_\varphi^k(k) + c_\varphi(h_n) \). By Theorem 5.1, \( c_\varphi(h_n) \to c_\varphi(h), c_\varphi(h_n + k) \to c_\varphi(h + k) \) and \([\varphi^{h_n}] \to [\varphi^h], [\varphi^{h_n+k}] \to [\varphi^{h+k}] \)
strongly. Now we can proceed exactly as in the proof of Theorem 12.10 in [5] to obtain (14). By putting \( \omega = [\varphi^{h+k}] \) in this equality, we get

\[
[\varphi^{h+k}](k) + S([\varphi^{h+k}], [\varphi^h]) = c_\varphi(h + k) - c_\varphi(h) \geq \omega(k) - S(\omega, [\varphi^h]),
\]

for all \( \omega \), which implies the statement of the proposition. \( \square \)

**Theorem 7.1.** Let \( h \in B_{\varphi} \). Then \( B_{[\varphi^h]} = B_{\varphi} \) and \( S_{[\varphi^h]} = S_{\varphi} \).

**Proof.** Let \( k \in M_* \) and let \( \varepsilon > 0 \). By Proposition 7.1

\[
c_\varphi^k(k) = c_\varphi(h + k) - c_\varphi(h).
\]

Since \( c_\varphi \) is continuous on \( B_{\varphi} \), there is a \( \delta > 0 \) such that

\[
|c_\varphi(h + k) - c_\varphi(h)| < \log 2
\]

if \( \|k\|_{\varphi} < \delta \). It follows that \( \|k\|_{\varphi} < \varepsilon \) whenever \( \|k\|_{\varphi} < \delta \varepsilon \) and this implies \( B_{[\varphi^h]} \subseteq B_{[\varphi^h]} \). In particular, \( h \in B_{[\varphi^h]} \).

Let \( h_n \) be a sequence converging to \( h \) in \( B_{\varphi} \), then by (6)

\[
\omega(h_n) - S(\omega, \varphi) = c_\varphi(h_n) - S(\omega, [\varphi^{h_n}])
\]
By Theorem \textbf{5.1} and lower semicontinuity,
\[
\omega(h) - S(\omega, \varphi) \leq c_\varphi(h) - S(\omega, [\varphi^h])
\]
This implies \(S_\varphi \subseteq S_{[\varphi^h]}\).

Further, \(h_n\) converges to \(h\) in \(B_{[\varphi^h]}\) and by Theorem \textbf{6.1} and Proposition \textbf{6.1}
\[
[[\varphi^h]^{-h}] = \lim_n [[\varphi^h]^{-h_n}] = \lim_n [\varphi^{h-h_n}] = \varphi.
\]
By the first part of the proof, \(B_{[\varphi^h]} = B_\varphi\) and \(S_\varphi = S_{[\varphi^h]}\).

\textbf{Theorem 7.2.} Let \(h, k \in B_\varphi\). Then the chain rule \(c_\varphi(h + k) = c_{[\varphi^h]}(k) + c_\varphi(h), \)
\([[[\varphi^h]^k] = [\varphi^{h+k}]\) holds.

\textit{Proof.} Let \(k_n \in \mathcal{M}_s\) be a sequence converging to \(k\) in \(B_\varphi = B_{[\varphi^h]}\). Then
\[
[[[\varphi^h]^k] = \lim_n [[\varphi^h]^{k_n}] = \lim_n [\varphi^{h+k_n}] = [\varphi^{h+k}].
\]
and by Proposition \textbf{6.1}
\[
c_\varphi(h + k) = \lim_n c_{[\varphi^h]}(k_n) + c_\varphi(h) = c_{[\varphi^h]}(k) + c_\varphi(h)
\]
holds.

\textit{Proof.} By \textbf{6.1} and lower semicontinuity, we have
\[
\omega(h) - S(\omega, \varphi) \leq c_\varphi(h) - S(\omega, [\varphi^h])
\]
Since, by the chain rule, \(\varphi = [[[\varphi^h]^{-h}]\) and \(c_{[\varphi^h]}(-h) = -c_\varphi(h),\) we also have
\[
\omega(-h) - S(\omega, [\varphi^h]) \leq c_{[\varphi^h]}(-h) - S(\omega, \varphi) = -c_\varphi(h) - S(\omega, \varphi)
\]
which implies the opposite inequality.

\textbf{Corollary 7.1.} Let \(h \in B_\varphi\) and let \(\omega\) be a normal state. Then the equality
\[
\omega(h) - S(\omega, \varphi) = c_\varphi(h) - S(\omega, [\varphi^h])
\]
holds.

\textit{Proof.} By \textbf{6.1} and lower semicontinuity, we have
\[
\omega(h) - S(\omega, \varphi) \leq c_\varphi(h) - S(\omega, [\varphi^h])
\]
Since, by the chain rule, \(\varphi = [[[\varphi^h]^{-h}]\) and \(c_{[\varphi^h]}(-h) = -c_\varphi(h),\) we also have
\[
\omega(-h) - S(\omega, [\varphi^h]) \leq c_{[\varphi^h]}(-h) - S(\omega, \varphi) = -c_\varphi(h) - S(\omega, \varphi)
\]
which implies the opposite inequality.

\textbf{Corollary 7.2.} Let \([\varphi^h] = [\varphi^h]\) for some \(h, k \in B_\varphi\). Then \(h - k = \varphi(h - k).\)

\textit{Proof.} Let us suppose that \(h \in B_\varphi\) is such that \([\varphi^h] = \varphi\). Then \([\varphi^n h] = \varphi\) for all \(n \in \mathbb{N}\). It follows that \(c_\varphi(nh) = n c_\varphi(h) = n c_{[\varphi^h]}(h)\) for all \(n\) and for \(0 \leq t \leq 1,\) we have by \textbf{5.1} and convexity of \(c_\varphi\) that
\[
t c_\varphi(h) = \varphi(th) \leq c_\varphi(th) \leq t c_\varphi(h)
\]
It follows that \(c_\varphi(th) = t c_\varphi(h) = t \varphi(h)\) for all \(t \geq 0.\) Since also \([\varphi^{-h}] = [[[\varphi^h]^{-h}]\] = \varphi, we have \(c_\varphi(-th) = t c_\varphi(-h) = -t \varphi(h)\) for \(t \geq 0.\)

It is easy to see that \(c_\varphi(k - \lambda) = c_\varphi(k) - \lambda\) for all \(k \in B_\varphi\) and \(\lambda \in \mathbb{R}\). Let \(\lambda = \varphi(h),\) then it follows that
\[
c_\varphi(t(h - \lambda)) = 0 = c_\varphi(t(-h + \lambda))
\]
for all \(t \geq 0.\) This implies \(\|h - \lambda\|_\varphi = 0\) and hence \(h = \lambda.\)

Let now \([\varphi^h] = [\varphi^k],\) then \([[[\varphi^k]^{-h}] = [\varphi^{k-h}] = \varphi\) and \(h - k = \lambda = \varphi(h - k).\) \(\square\)

Note that the function \(\tilde{c}_\varphi : B_{\varphi,0} \to \mathbb{R}\) corresponds to the cumulant generating functional in the commutative case. Let us list some of its properties.
Theorem 7.3. The function \( \bar{c}_\varphi \) has the following properties.

(i) \( \bar{c}_\varphi \) is positive, strictly convex and continuous, \( \bar{c}_\varphi(0) = 0 \).
(ii) \( \bar{c}_\varphi \) is Gateaux differentiable, with \( \bar{c}_\varphi'(h) = [\varphi^h] - \varphi \).
(iii) Let \( h,k \in B_{\varphi,0} \) and \( 0 < \lambda < 1 \) be such that
\[ \bar{c}_\varphi(\lambda h + (1 - \lambda)k) = \lambda \bar{c}_\varphi(h) + (1 - \lambda)\bar{c}_\varphi(k). \]

This implies that the maximum in both expressions on the right hand side is attained at the same point. Therefore \( [\varphi^h] = [\varphi^k] \), hence \( h - k = \varphi(h - k) = 0 \).

(ii) By Theorem 7.3 \( [\varphi^h] - \varphi \) is the unique element in \( B_{\varphi,0}^* \), such that
\[ ([\varphi^h] - \varphi)(h) = \bar{c}_\varphi(h) + \bar{c}_\varphi^*(\varphi^h - \varphi). \]

By 7.2, this implies that \( \bar{c}_\varphi \) is Gateaux differentiable in \( h \) with derivative \( \bar{c}_\varphi'(h) = [\varphi^h] - \varphi \).

(iii) Let \( h_n \to h \in B_{\varphi} \), then \( [\varphi^{h_n}] \) converges strongly to \( [\varphi^h] \) and \( S([\varphi^{h_n}], \varphi) \to S([\varphi^h], \varphi) \). It follows that \( [\varphi^{h_n}](k) \to [\varphi^h](k) \) for each \( k \in M_\varphi \) and moreover, the set \( \{[\varphi^{h_n}], n \in \mathbb{N} \} \) is equicontinuous in \( B_{\varphi}^* \). This implies that \( [\varphi^{h_n}](k) \to [\varphi^h](k) \) for all \( k \in B_{\varphi} \). The map is one-to-one by Corollary 7.2.

\[ \square \]

8. A manifold structure on faithful states.

Recall that a \( C^p \)-atlas on a set \( X \) is a family of pairs \( \{(U_i, e_i)\} \), such that

(i) \( U_i \subset X \) for all \( i \) and \( \cup U_i = X \).
(ii) For all \( i \), \( e_i \) is a bijection of \( U_i \) onto an open subset \( e_i(U_i) \) in some Banach space \( \mathcal{B}_i \), and for \( i, j \), \( e_i(U_i \cap U_j) \) is open in \( \mathcal{B}_j \).
(iii) The map \( e_j e_i^{-1} : e_j(U_i \cap U_j) \to e_i(U_i \cap U_j) \) is a \( C^p \)-isomorphism for all \( i, j \).

Let \( \mathcal{F}_\varphi \) be the set of faithful normal states on \( M \). For \( \varphi \in \mathcal{F}_\varphi \), let \( V_\varphi \) be the open unit ball in \( B_{\varphi,0} \) and let \( s_\varphi : V_\varphi \to \mathcal{F}_\varphi \) be the map \( h \mapsto [\varphi^h] \). By Corollary 7.2 \( s_\varphi \) is a bijection onto the set \( s_\varphi(V_\varphi) =: U_\varphi \subset \mathcal{S}_\varphi \). Let \( e_\varphi \) be the restriction of \( s_\varphi^{-1} \) to \( U_\varphi \). Then we have

Theorem 8.1. \( \{(U_\varphi, e_\varphi), \varphi \in \mathcal{F}_\varphi \} \) is a \( C^{\infty} \)-atlas on \( \mathcal{F}_\varphi \).
Proof. The property (i) and the first part of (ii) of the definition of the $C^p$ atlas are obviously satisfied. Let $\varphi_1, \varphi_2 \in F_s$ be such that $U_{\varphi_1} \cap U_{\varphi_2} \neq \emptyset$. We prove that $c_{\varphi_1}(U_{\varphi_1} \cap U_{\varphi_2})$ is open in $B_{\varphi_1,0}$.

Let $h_1 \in c_{\varphi_1}(U_{\varphi_1} \cap U_{\varphi_2})$. Then there is some $h_2 \in B_{\varphi_2,0}$, such that $[\varphi_1^{h_1}] = [\varphi_2^{h_2}]$. By Theorem 7.1, $B_{\varphi_2} = B_{\varphi_2^{h_1}} = B_{\varphi_2^{h_2}} = B_{\varphi_2}$ and by the chain rule, $\varphi_1 = [\varphi_2^k]$, where $k = h_2 - h_1 + \varphi_2(h_1) \in B_{\varphi_2,0}$. Clearly, the map $B_{\varphi_1,0} \to B_{\varphi_2,0}$, given by $h \mapsto h - \varphi_2(h)$ is continuous.

Let $\varepsilon > 0$ be such that $h_2 + h_2' \in V_{\varphi_2}$ whenever $\|h_2\|_{\varphi_2} < \varepsilon$ and let us choose $\delta > 0$ such that $h_1 + h_1' \in V_{\varphi_1}$ and $\|h_1' - \varphi_2(h_1')\|_{\varphi_2} < \varepsilon$ for $\|h_1'\|_{\varphi_1} < \delta$. For such $h_1'$, we have

$$s_{\varphi_1}(h_1 + h_1') = [\varphi_1^{h_1 + h_1'}] = [\varphi_2^{h_2 + h_1 + h_1' - \varphi_2(h_1')}] = \frac{h_2 + h_1 + h_1' - \varphi_2(h_1')}{h_2} \in U_{\varphi_1} \cap U_{\varphi_2}$$

This proves that $s_{\varphi_1}^{-1}(U_{\varphi_1} \cap U_{\varphi_2})$ is open in $B_{\varphi_1,0}$. It is also clear that the map

$$s_{\varphi_2}^{-1} : s_{\varphi_1}^{-1}(U_{\varphi_1} \cap U_{\varphi_2}) \to B_{\varphi_2,0}$$

$$h \mapsto k + h - \varphi_2(h)$$

is $C^\infty$, which proves (iii).

It is not difficult to see that for $\varphi \in F_s$, the set $F_\varphi := \{[\varphi^h], h \in B_{\varphi,0}\}$ is a connected component of the manifold. Let us now define a family of mappings

$$U_{\varphi_1,\varphi_2}^{(e)} : B_{\varphi_1,0} \ni h \mapsto h - \varphi_2(h) \in B_{\varphi_2,0}, \quad \varphi_1, \varphi_2 \in F_\varphi$$

It is clear that this defines a parallel transport on the tangent bundle of $F_\varphi$ and the associated globally flat affine connection is the exponential connection, 4.

Let us recall that the dual connection is defined on the cotangent bundle $T^*F_\varphi$ by means of the parallel transport $\{(U_{\varphi_1,\varphi_2}^{(e)})^*, \varphi_1, \varphi_2 \in F_\varphi\}$, where

$$((U_{\varphi_1,\varphi_2}^{(e)})^*)^*v, h) = (v, U_{\varphi_1,\varphi_2}^{(e)}h), \quad v \in B_{\varphi_1,0}^*, h \in B_{\varphi_2,0},$$

and the duality is given by $\langle v, h \rangle = v(h)$. Since $v(h - \varphi_1(h)) = v(h)$ for all $\varphi_1$, the dual parallel transport is

$$U_{\varphi_1,\varphi_2}^{(m)} : B_{\varphi_1,0}^* \ni v \mapsto v \in B_{\varphi_2,0}^*, \quad \varphi_1, \varphi_2 \in F_\varphi$$

which corresponds to the mixture connection.

References

[1] S. Amari and H. Nagaoka, Method of information geometry, AMS monograph, Oxford University Press, 2000
[2] S. Amari, O.E. Barndorff-Nielsen, R.E. Kass, S.L. Lauritzen, C.R. Rao: Differential geometry in statistical inference, IMS Lecture notes-Monograph series 10, Mayward, 1987
[3] I. Ekeland, R. Temam, Convex Analysis and Variational Problems, North Holland, 1976.
[4] P. Gibrilisco 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
[5] M. R. Grasselli, Dual connections in nonparametric classical information geometry, to appear in the Annals of the Institute for Statistical Mathematics, 2005
[6] M. Grasselli and R.F. Streater, The quantum information manifold for epsilon-bounded forms, Rep. Math. Phys., 46, 325-335, 2000. math-ph/9910031
[7] G. Köthe, Topological vector spaces I, Springer-Verlag Berlin Heidelberg New York, 1983,
[8] M. Ohya and D. Petz, Quantum Entropy and its use, Springer, Heidelberg, 1993
[9] 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), 1542–1561
[10] R. F. Streater, The Information Manifold for Relatively Bounded Potentials, Tr. Mat. Inst. Steklova, 228, 217-235, 2000.

[11] 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, math-ph/9910036

[12] R. F. Streater, Quantum Orlicz spaces in information geometry, math-ph/0407046