Information Geometry
https://doi.org/10.1007/s41884-023-00097-x

RESEARCH PAPER

The exponential Orlicz space in quantum information geometry

Anna Jenčová[1]

Received: 10 October 2022 / Revised: 2 January 2023 / Accepted: 5 January 2023

Abstract

We review the construction of a quantum version of the exponential statistical manifold over the set of all faithful normal positive functionals on a von Neumann algebra. The construction is based on the relative entropy approach to state perturbation. We construct a quantum version of the exponential Orlicz space and discuss the properties of this space and its dual with respect to Kosaki L_p-spaces. We show that the constructed manifold admits a canonical divergence satisfying a Pythagorean relation. We also prove that the manifold structure is invariant under sufficient channels.

Keywords Quantum exponential manifold · Quantum relative entropy · Perturbation of states · Canonical divergence

1 Introduction

One of the fundamental achievements of Information geometry is the rigorous extension from parametric statistical models to the nonparametric case by Pistone and Sempi, [1], who constructed a Banach manifold structure on the set of probability measures equivalent to a given probability measure. The manifold structure is based on an Orlicz space associated to an exponential Young function $\Phi(x) = \cosh(x) - 1$. This theory has been subsequently developed in a number of works, see e.g. [2, 3]. In this construction, the properties of the moment generating function and its conjugate, the Kullback Leibler divergence (relative entropy), play a central role.

To obtain a full quantum version of the Pistone-Sempi construction would mean to introduce an analogous Banach manifold structure on the set of faithful normal states of a general (σ-finite) von Neumann algebra. The problem is that known versions

Communicated by Jan Naudts.

✉ Anna Jenčová
jenca@mat.savba.sk

1 Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia

of quantum Orlicz spaces are either restricted to the semifinite case (e.g. [4, 5]) or are technically quite involved ([6]) and it is unclear how to introduce an exponential structure on the set of states, based on these spaces.

Another approach using perturbation of states on the algebra $B(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$, was taken in [7, 8]. Here the manifold is modelled on the space of certain (unbounded) perturbation operators, which is given the structure of a Banach space. In [8], the Banach space is constructed from the free energy functional, which can be seen as the counterpart of the classical cumulant generating function.

This idea inspired the works [9, 10], where a definition of the exponential Orlicz space over a von Neumann algebra $\mathcal{M}$ with respect to a faithful normal state ρ is proposed, using the relative entropy approach to state perturbation. This definition has the advantage that it is based directly on the relative entropy $S(\cdot\|\cdot)$ and its properties. In particular, the set $\mathcal{P}_\rho$ of all normal positive linear functionals such that $S(\omega\|\rho) < \infty$ is identified with a generating cone in the dual of the constructed exponential Orlicz space, so that a manifold structure on normal states of $\mathcal{M}$, respecting the relative entropy, can be introduced by immersion into a Banach space. Moreover, an exponential manifold structure is obtained using perturbations of the state ρ by elements of the exponential Orlicz space and the connected components of this manifold are contained in $\mathcal{P}_\rho$.

In the present paper, we review the construction of the exponential Orlicz space and its dual, as defined in [9, 10]. We present the proofs in a more streamlined and precise form. The dual space is found explicitly as an Orlicz space, using the conjugate Young function. We show the relation of the constructed spaces to the Kosaki L_p-spaces. The manifold structure is introduced over the positive cone of faithful positive linear functionals, rather than states, similarly to the approach in [11]. We define a canonical divergence on the manifold, satisfying a generalized Pythagorean relation. Finally, we prove the invariance of our structures under sufficient channels, which is the counterpart of the important invariance property of the classical information geometry.

2 The exponential Orlicz space

In this section, we review the definition of the exponential Orlicz space from [9], construct its dual as an Orlicz space and study some of the properties of these spaces.

2.1 A general construction of an Orlicz space

Let X be a real vector space. A function $\Phi : X \to [0, \infty]$ is called a Young function if it satisfies:

(i) Φ is convex,
(ii) $\Phi(x) = \Phi(-x)$ for all $x \in X$ and $\Phi(0) = 0$,
(iii) if $x \neq 0$ then $\lim_{t\to\infty} \Phi(tx) = \infty$.

For a Young function Φ, put $C_\Phi := \{x \in X, \Phi(x) \leq 1\}$ and $V_\Phi := \{x \in \Phi(x),\ \exists s > 0,\ \Phi(sx) < \infty\}$. The set C_Φ is absolutely convex and $V_\Phi = \bigcup_n nC_\Phi$ is the linear span of the effective domain $\mathrm{Dom}(\Phi) = \{x \in X,\ \Phi(x) < \infty\}$. We can define a norm in V_Φ as the Minkowski functional of C_Φ:

$$\|x\|_\Phi := \inf\{\lambda > 0,\ \Phi\left(\frac{x}{\lambda}\right) \leq 1\}, \qquad x \in V_\Phi.$$

The completion of V_Φ with respect to this norm will be denoted by B_Φ.

Let (Ω, Σ, μ) be a measure space and let X be the vector space of measurable functions $f : \Omega \to \mathbb{R}$. Let φ be a Young function on $\mathbb{R}$ and put

$$\Phi_\varphi(f) := \int_\Omega \varphi(|f|)d\mu.$$

Then B_{Φ_φ} is the classical Orlicz space $L^\varphi(\Omega, \Sigma, \mu)$ and $\|\cdot\|_{\Phi_\varphi}$ is the Luxemburg-Nakano norm, [12]. As another example, let $\mathcal{M}$ be a semifinite von Neumann algebra with a faithful normal semifinite trace τ. Let X be the vector space of self-adjoint τ-measurable operators and $\Phi_\varphi(x) = \tau(\varphi(|x|))$, then we obtain the Orlicz space $L^\varphi(\mathcal{M}, \tau)$, [4]. In [5], a version of the Orlicz space with respect to a faithful normal state ρ was also defined by a construction of this type. See also [6] in the case of a general von Neumann algebra with a faithful normal weight.

It was shown in [9, Prop. 2] that if X is a Banach space and Φ is continuous, then X is continuously embedded in B_Φ. The conjugate function Φ^* is again a Young function such that $V_{\Phi^*} = \mathrm{Dom}(\Phi^*) = B_{\Phi^*}$ and we have a continuous embedding $B_{\Phi^*} \sqsubseteq X^*$. Moreover, $B_{\Phi^*} = B_\Phi^*$, with equivalent norms.

2.2 Basic setting and notations

We briefly describe the setting of von Neumann algebras and noncommutative L_p-spaces. For a quick overview of these topics, see [13].

Let $\mathcal{M}$ be a (σ-finite) von Neumann algebra. We will denote by $\mathcal{M}^*$ the dual space of $\mathcal{M}$ and by $\mathcal{M}_*$ the predual, consisting of normal functionals in $\mathcal{M}^*$. The positive cones in these spaces will be denoted by $\mathcal{M}^+$, $(\mathcal{M}^*)^+$ and $\mathcal{M}_*^+$. An element $\rho \in \mathcal{M}_*^+$ is faithful if $\rho(a) = 0$ implies $a = 0$, for any $a \in \mathcal{M}^+$.

For $1 \leq p \leq \infty$, we denote the Haagerup L_p-space over $\mathcal{M}$ by $L_p(\mathcal{M})$ and its norm by $\|\cdot\|_p$. We will use the identification of $L_\infty(\mathcal{M})$ with $\mathcal{M}$ and $L_1(\mathcal{M})$ with $\mathcal{M}_*$. Let $h_\psi \in L_1(\mathcal{M})$ be the element corresponding to $\psi \in \mathcal{M}_*$, then we can define the trace in $L_1(\mathcal{M})$ by $\mathrm{Tr}[h_\psi] = \psi(1)$.

For $p, q, r \geq 1$ such that $1/p + 1/q = 1/r$ and $h \in L_p(\mathcal{M})$, $k \in L_q(\mathcal{M})$, we have $hk \in L_r(\mathcal{M})$ and the Hölder inequality holds:

$$\|hk\|_r \leq \|h\|_p \|k\|_q.$$

For $1 \leq p < \infty$ and $1/p + 1/q = 1$, the space $L_q(\mathcal{M})$ can be identified with the dual space $L_p(\mathcal{M})^*$, with duality given by

$$\langle h, k\rangle = \mathrm{Tr}[hk], \qquad h \in L_p(\mathcal{M}),\ k \in L_q(\mathcal{M}).$$

The space $L_2(\mathcal{M})$ is a Hilbert space with inner product

$$(h,k)=\mathrm{Tr}[h^*k],\qquad h,k\in L_2(\mathcal{M}).$$

We will use the representation of $\mathcal{M}$ on $L_2(\mathcal{M})$ by the left action $\lambda(a): h\mapsto ah$ for $a\in\mathcal{M}$ and $h\in L_2(\mathcal{M})$. The quadruple $(\lambda(\mathcal{M}), L_2(\mathcal{M}), L_2(\mathcal{M})^+, J)$, where $L_2(\mathcal{M})^+$ is the cone of positive operators in $L_2(\mathcal{M})$ and J is defined by $Jh=h^*$, is a standard form of $\mathcal{M}$ ([14, Thm. 3.6], [13, Thm. 9.29]). For more on the standard form see [15] or [13, Sect. 3]. Any positive normal functional $\varphi\in\mathcal{M}_*^+$ has a unique vector representative $h_\varphi^{1/2}$ in the cone $L_2(\mathcal{M})^+$, that is,

$$\varphi(a)=(h_\varphi^{1/2}, ah_\varphi^{1/2}),\qquad a\in\mathcal{M}.$$

Let us now fix a faithful positive normal functional $\rho\in\mathcal{M}_*^+$. The (symmetric) Kosaki L_p-space with respect to ρ [16] is defined via complex interpolation, using the continuous embedding

$$i_{\infty,\rho}:\mathcal{M}\to L_1(\mathcal{M}),\qquad a\mapsto h_\rho^{1/2}ah_\rho^{1/2}. \tag{1}$$

Let us denote the image $i_{\infty,\rho}(\mathcal{M})$ by $L_\infty(\mathcal{M},\rho)$, with the norm $\|i_{\infty,\rho}(a)\|_{\infty,\rho}=\|a\|$. The interpolation space $C_{1/p}(L_\infty(\mathcal{M},\rho), L_1(\mathcal{M}))$ [17] will be denoted by $L_p(\mathcal{M},\rho)$ and the norm by $\|\cdot\|_{p,\rho}$. The map

$$i_{p,\rho}: L_p(\mathcal{M})\to L_1(\mathcal{M}),\qquad k\mapsto h_\rho^{1/2q}kh_\rho^{1/2q} \tag{2}$$

with $1/p+1/q=1$ is an isometric isomorphism of $L_p(\mathcal{M})$ onto $L_p(\mathcal{M},\rho)$ for $1\le p\le\infty$. From the properties of complex interpolation spaces, we have for $1\le p'\le p\le\infty$ the continuous embeddings $L_p(\mathcal{M},\rho)\sqsubseteq L_{p'}(\mathcal{M},\rho)\sqsubseteq L_1(\mathcal{M})$. We have $L_q(\mathcal{M},\rho)\simeq L_p(\mathcal{M},\rho)^*$ for $1\le p<\infty$ and $1/p+1/q=1$, with duality given by

$$\langle i_{p,\rho}(k), i_{q,\rho}(l)\rangle=\mathrm{Tr}[kl],\qquad k\in L_p(\mathcal{M}),\ l\in L_q(\mathcal{M}).$$

Note also that the Kosaki L_p-spaces can be constructed as in Sect. 2.1, where $X=\mathcal{M}^s:=\{a=a^*\in\mathcal{M}\}$ and $\Phi(a)=\|h_\rho^{1/2p}ah_\rho^{1/2p}\|_p$, [18].

Let $\mathcal{N}$ be another von Neumann algebra and let $T: L_1(\mathcal{M})\to L_1(\mathcal{N})$ be a positive linear map that preserves trace. Such a map will be called a channel. The adjoint of T is a positive unital normal map $T^*:\mathcal{N}\to\mathcal{M}$.

Let ρ be a faithful element in $\mathcal{M}_*^+$. One can see ([19, Sect. 3.3]) that if $e=s(T(\rho))$ is the support projection of $T(\rho)$, then we have $T(\omega)=eT(\omega)e$ for all $\omega\in L_1(\mathcal{M})$, hence we may suppose that $T(\rho)$ is faithful by replacing $\mathcal{N}$ by $e\mathcal{N}e$.

Proposition 1 *[19] The restriction of a channel T to $L_p(\mathcal{M},\rho)$ is contraction $L_p(\mathcal{M},\rho)\to L_p(\mathcal{N},T(\rho))$, for any $1\le p\le\infty$.*

In the case $p = \infty$, there is a positive linear map $T^*_\rho : \mathcal{M} \to \mathcal{N}$, defined by

$$T(h_\rho^{1/2} a h_\rho^{1/2}) = T(\rho)^{1/2} T^*_\rho(a) T(\rho)^{1/2}, \qquad a \in \mathcal{M}.$$

The map T^*_ρ was introduced in [20] and is called the Petz dual of T (with respect to ρ). It was also proved that T^*_ρ is unital and normal, moreover, it is n-positive if and only if T is n-positive, for any n. Let $T_\rho : L_1(\mathcal{N}) \to L_1(\mathcal{M})$ be the preadjoint of T^*_ρ. Then $T_\rho \circ T(\rho) = \rho$ and the Petz dual of T_ρ is T^*.

2.3 Relative entropy and related functionals

The Araki relative entropy for $\omega, \rho \in \mathcal{M}_*^+$ [21–23] is defined using the relative modular operator $\Delta_{\rho,\omega} (= \Delta_{\rho, h_\omega^{1/2}})$ as

$$S(\omega\|\rho) = \begin{cases} -\langle \log(\Delta_{\rho,\omega}) h_\omega^{1/2}, h_\omega^{1/2} \rangle & \text{if } s(\omega) \le s(\rho) \\ \infty & \text{otherwise.} \end{cases}$$

Here $s(\rho)$ denotes the support projection of ρ. Alternatively, we have the following variational formula due to Kosaki [24]:

$$S(\omega\|\rho) = \sup_n \sup \left\{ \omega(1) \log n - \int_{1/n}^{\infty} (\omega(y(t)^* y(t)) + t^{-1} \rho(x(t)x(t)^*)) \frac{dt}{t} \right\} \quad (3)$$

here the second supremum is taken over all step functions $x : (1/n, \infty) \to L$ with finite range, $y(t) = 1 - x(t)$ and L is a subspace in $\mathcal{M}$ containing 1 which is dense in the strong*-operator topology.

The relative entropy S is a jointly convex function $S : \mathcal{M}_*^+ \times \mathcal{M}_*^+ \to \mathbb{R} \cup \{\infty\}$, lower semicontinuous with respect to the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology. Moreover, S is strictly convex in the first variable, which can be inferred from the identity [23, Prop. 5.22]

$$S(\omega\|\rho) + \sum_i S(\omega_i\|\omega) = \sum_i S(\omega_i\|\rho), \quad \omega = \sum_{i=1}^k \omega_i, \ \omega \in \mathcal{M}_*^+. \quad (4)$$

Note also that since $\omega_i \le \omega$ in (4), we have $S(\omega_i\|\omega) \le S(\omega_i\|\omega_i) < \infty$ for all i. See [23, Sec. 5] for details and a list of further important properties of S. The next statement shows the relation to the Kosaki L_p-space.

Proposition 2 *[19, 25] Let $\omega, \rho \in \mathcal{M}_*^+$ be such that $h_\omega \in L_p(\mathcal{M}, \rho)$ for some $p > 1$. Then the function $f : (1, p] \to \mathbb{R}$, defined as*

$$f(\alpha) := \frac{1}{\alpha - 1} \log \frac{\|h_\omega\|_{\alpha,\rho}^\alpha}{\omega(1)}$$

is increasing and $\lim_{\alpha \downarrow 1} f(\alpha) = \frac{1}{\omega(1)} S(\omega\|\rho)$.

Using the Kosaki variational formula (3), the relative entropy can be readily extended to a function $S : (\mathcal{M}^*)^+ \times (\mathcal{M}^*)^+ \to \mathbb{R} \cup \{\infty\}$.

Proposition 3 *Let $\rho \in \mathcal{M}_*^+$ and let $\omega \in (\mathcal{M}^*)^+$. If $\omega \notin \mathcal{M}_*^+$, then $S(\omega\|\rho) = \infty$.*

Proof [23] There is another way to define the relative entropy for elements in $(\mathcal{M}^*)^+$. Let $(\pi_u, \mathcal{H}_u)$ be the universal representation of $\mathcal{M}$ and let $\bar{\mathcal{M}} = \pi_u(\mathcal{M})'' \cong \mathcal{M}^{**}$ be the universal enveloping von Neumann algebra of $\mathcal{M}$ [26]. Then each element of the dual space $\omega \in \mathcal{M}^*$ has a unique extension to a normal functional $\bar{\omega}$ on $\bar{\mathcal{M}}$ and $\mathcal{M}^*$ is the predual of $\bar{\mathcal{M}}$. Moreover, there is a central projection $z_0 \in \bar{\mathcal{M}}$ such that $\mathcal{M}_* = \mathcal{M}^* z_0$. We can define for $\omega, \rho \in (\mathcal{M}^*)^+$ the relative entropy $\bar{S} : (\mathcal{M}^*)^+ \times (\mathcal{M}^*)^+ \to \mathbb{R}$ as

$$\bar{S}(\omega\|\rho) := S_{\bar{\mathcal{M}}}(\bar{\omega}\|\bar{\rho})$$

(here $S_{\bar{\mathcal{M}}}$ is computed with respect to the von Neumann algebra $\bar{\mathcal{M}}$). Now note that we may use $L = \pi_u(\mathcal{M})$ in the variational formula for $S_{\bar{\mathcal{M}}}$ and $L = \mathcal{M}$ for S, which implies that $\bar{S} = S$. Let $\rho \in \mathcal{M}_*^+$, $\omega \in (\mathcal{M}^*)^+$ and assume that ω is not normal. Then we must have $\bar{\omega}(1 - z_0) > 0$ but $\bar{\rho}(1 - z_0) = 0$, so that $s(\bar{\omega}) \not\leq s(\bar{\rho})$. By definition of the relative entropy, this implies that $\bar{S}(\omega\|\rho) = \infty$. $\square$

From now on, let us fix a faithful normal functional $\rho \in \mathcal{M}_*^+$. Let $\mathcal{M}^s$ denote the real vector subspace of self-adjoint elements of $\mathcal{M}$. Then $\mathcal{M}^s$ is closed in $\mathcal{M}$ and its Banach space dual is the space $(\mathcal{M}^*)^s$ of all linear functionals $\varphi \in \mathcal{M}^*$ satisfying $\varphi(a^*) = \overline{\varphi(a)}$, $a \in \mathcal{M}$. Note that we have $(\mathcal{M}^*)^s = (\mathcal{M}^*)^+ - (\mathcal{M}^*)^+$. Similarly, $\mathcal{M}_*^s = (\mathcal{M}^*)^s \cap \mathcal{M}_*$ is the predual of $\mathcal{M}^s$ and $\mathcal{M}_*^s = \mathcal{M}_*^+ - \mathcal{M}_*^+$. Let us define the function $F_\rho : (\mathcal{M}^*)^s \to \mathbb{R}$ by

$$F_\rho(\omega) := \begin{cases} S(\omega\|\rho) - \omega(1) & \text{if } \omega \in (\mathcal{M}^*)^+ \\ \infty & \text{otherwise.} \end{cases}$$

We also define the sets

$$\mathcal{S}_C := \{\omega \in (\mathcal{M}^*)^s,\ F_\rho(\omega) \leq C\},\ C \in \mathbb{R}, \qquad \mathcal{P}_\rho := \{\omega \in (\mathcal{M}^*)^s,\ F_\rho(\omega) < \infty\}.$$

In other words, $\mathcal{P}_\rho$ is the effective domain of F_ρ. Note that we have $\mathcal{S}_C \subseteq \mathcal{P}_\rho \subseteq \mathcal{M}_*^+$, by Proposition 3. The next proposition lists some important properties of the function F_ρ and these sets.

Proposition 4 *(i) $F_\rho : (\mathcal{M}^*)^s \to \mathbb{R}$ is strictly convex and lower semicontinuous in the $\sigma((\mathcal{M}^*)^s, \mathcal{M}^s)$ topology.*

(ii) We have the inequalities

$$F_\rho(\omega) \geq \omega(1)(\log \frac{\omega(1)}{\rho(1)} - 1) \geq -\rho(1).$$

The first inequality becomes an equality if and only if $\omega = \lambda\rho$ for some $\lambda \geq 0$. In particular, $F_\rho(\omega) = -\rho(1)$ if and only if $\omega = \rho$.

(iii) *For any* $C \in \mathbb{R}$, $\mathcal{S}_C$ *is convex and compact in both the* $\sigma((\mathcal{M}^*)^s, \mathcal{M}^s)$ *and the* $\sigma(\mathcal{M}_*^s, \mathcal{M}^s)$*-topology.*

(iv) *The set* $\mathcal{P}_\rho$ *is a face of the cone* $\mathcal{M}_*^+$, *containing* $L_p(\mathcal{M}, \rho)^+$ *for any* $1 < p \leq \infty$.

Proof The proof of (i)-(ii) follows from the variational formula and properties of S. For the proof of (iii), let $\omega \in \mathcal{S}_C$, then by (ii),

$$\omega(1)(\log \frac{\omega(1)}{\rho(1)} - 1) \leq F_\rho(\omega) \leq C.$$

This implies that $\omega(1) = \|\omega\|$ must be bounded over $\mathcal{S}_C$. Since $\sigma((\mathcal{M}^*)^s, \mathcal{M}^s)$ is the weak*-topology on $(\mathcal{M}^*)^s$ and $\mathcal{S}_C$ is closed by (i), this implies that $\mathcal{S}_C$ is compact. But $\mathcal{S}_C \subseteq \mathcal{M}_*^s$, so that it is also compact in the $\sigma(\mathcal{M}_*^s, \mathcal{M}^s)$-topology.

To prove the last statement (iv), let $\omega = \sum_i \omega_i$ for some $\omega_i \in \mathcal{M}_*^+$. Then by (4)

$$F_\rho(\omega) + \sum_i S(\omega_i \| \omega) = \sum_i F_\rho(\omega_i).$$

Since $S(\omega_i \| \omega) < \infty$, $\omega \in \mathcal{P}_\rho$ if and only if all $\omega_i \in \mathcal{P}_\rho$, so that $\mathcal{P}_\rho$ is a face of $\mathcal{M}_*^+$. The fact that $L_p(\mathcal{M}, \rho)^+ \subseteq \mathcal{P}_\rho$ for $1 < p \leq \infty$ follows from Proposition 2.

□

We also have the following important monotonicity property.

Proposition 5 *Let* $T : L_1(\mathcal{M}) \to L_1(\mathcal{N})$ *be a channel. Then*

$$F_{T(\rho)}(T(\omega)) \leq F_\rho(\omega), \qquad \omega \in (\mathcal{M}^*)^s.$$

Proof The statement follows from [27, Sect. 3.1 and Thm. 4.1]. □

We next study the Legendre-Fenchel conjugate of F_ρ with respect to the dual pair $((\mathcal{M}^*)^s, \mathcal{M}^s)$, see e.g. [28] or [29] for more information on Legendre-Fenchel duality of convex functions. Namely, we define the function C_ρ on $\mathcal{M}^s$ as

$$C_\rho(a) := F_\rho^*(a) = \sup_{\omega \in (\mathcal{M}^*)^s} \omega(a) - F_\rho(\omega), \qquad a \in \mathcal{M}^s. \tag{5}$$

The proof of the following result can be obtained from [23, Sec. 12]. See also [30, 31]. We collect the arguments for convenience of the reader.

Theorem 6 *The supremum in (5) is attained at a unique functional* $\rho^a \in \mathcal{M}_*^+$. *The element* ρ^a *is faithful and* $C_\rho(a) = \rho^a(1)$. *Moreover, we have the equality*

$$\omega(a) + S(\omega \| \rho^a) = S(\omega \| \rho), \qquad \omega \in \mathcal{M}_*^+ \tag{6}$$

and the chain rule

$$\rho^{a+b} = (\rho^a)^b, \quad C_\rho(a+b) = C_{\rho^a}(b), \qquad a, b \in \mathcal{M}^s. \tag{7}$$

Proof Let $a \in \mathcal{M}^s$ and let $\xi(a)$ denote the perturbed vector [32]

$$\xi(a) = \sum_{n=0}^{\infty} \int_0^{1/2} dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_n} dt_n \Delta_\rho^{t_n} a \Delta_\rho^{t_{n-1}-t_n} a \dots \Delta_\rho^{t_1-t_2} a h_\rho^{1/2}.$$

Then $\xi(a) \in L_2(\mathcal{M})^+$ and the functional $\rho^a \in \mathcal{M}_*^+$ given by $(\xi(a), \cdot\xi(a))$ is faithful. By [22, Thm. 3.10], ρ^a satisfies (6). It follows that for $\omega \in \mathcal{P}_\rho$,

$$\omega(a) - F_\rho(\omega) = -F_{\rho^a}(\omega) \leq \rho^a(1),$$

with equality if and only if $\omega = \rho^a$ (Proposition 4 (iii)). By replacing ρ by ρ^b in (6), we obtain

$$\omega(a+b) + S(\omega\|(\rho^b)^a) = \omega(b) + S(\omega\|\rho^b) = S(\omega\|\rho),$$

which implies the chain rule (7).

□

Example 1 Let $\mathcal{M} = B(\mathcal{H})$, the algebra of bounded operators on a Hilbert space $\mathcal{H}$. The functional ρ is represented as a density operator $\rho \in B(\mathcal{H})^+$ with finite trace, such that $\rho(a) = \mathrm{tr}[\rho a]$, (tr being the usual trace on $B(\mathcal{H})$). One can see that in this case,

$$\rho^a = \exp(\log \rho + a).$$

The following result is obtained from [28, Prop. 5.3 and 5.4].

Lemma 7 *The function C_ρ is Gateaux differentiable, with Gateaux derivative at $b \in \mathcal{M}^s$ given by $C'_\rho(b) = \rho^b$. For $a, b \in \mathcal{M}^s$, we have*

$$C_\rho(a) - C_\rho(b) \geq \rho^b(a-b).$$

2.4 The exponential Young function and its dual

We now introduce a conjugate pair of Young functions on the Banach spaces $\mathcal{M}^s$ and $(\mathcal{M}^*)^s$. Define

$$\Phi_\rho(a) := \frac{1}{2}(C_\rho(a) + C_\rho(-a)) - \rho(1), \qquad a \in \mathcal{M}^s.$$
$$\Psi_\rho(\psi) := \frac{1}{2} \inf_{\substack{\omega_\pm \in (\mathcal{M}^*)^+ \\ 2\psi = \omega_+ - \omega_-}} \left[F_\rho(\omega_+) + F_\rho(\omega_-)\right] + \rho(1), \qquad \psi \in (\mathcal{M}^*)^s.$$

It was proved in [9] that Φ_ρ is a strictly convex and continuous Young function $\mathcal{M}^s \to \mathbb{R}$. We now look at the properties of Ψ_ρ.

Lemma 8 *Ψ_ρ is a strictly convex and weak*-lower semicontinuous function on $(\mathcal{M}^*)^s$, with effective domain*

$$\mathrm{Dom}(\Psi_\rho) = \{\psi \in (\mathcal{M}^*)^s,\ \Psi_\rho(\psi) < \infty\} = \mathcal{P}_\rho - \mathcal{P}_\rho \subseteq \mathcal{M}_*^s.$$

Proof It is quite clear that $\Psi_\rho(\psi)$ is finite if and only if $\psi = \omega_+ - \omega_-$ for some $\omega_\pm \in \mathcal{P}_\rho$. Further, strict convexity of F_ρ implies that Ψ_ρ is strictly convex as well. For the last statement we have to show that for any $c > 0$, the set $\{\psi \in (\mathcal{M}^*)^s,\ \Psi_\rho(\psi) \le c\}$ is weak*-closed. So assume that (ψ_i) is a net in $(\mathcal{M}^*)^s$ such that $\Psi_\rho(\psi_i) \le c$ and let $\psi_i \to \psi$ in the weak*-topology. For each $\varepsilon > 0$ and for all i there are some functionals $\psi^\varepsilon_{i,\pm} \in \mathcal{M}_*^+$ such that $2\psi_i = \psi^\varepsilon_{i,+} - \psi^\varepsilon_{i,-}$ and

$$\frac{1}{2}[F_\rho(\psi^\varepsilon_{i,+}) + F_\rho(\psi^\varepsilon_{i,-})] + \rho(1) \le c + \varepsilon.$$

Using Proposition 4 (ii), we obtain that $\psi^\varepsilon_{i,\pm} \in \mathcal{S}_{K_\epsilon}$ with $K_\epsilon = 2(c+\varepsilon) - \rho(1)$. By Proposition 4 (iii), $\mathcal{S}_{K_\varepsilon}$ is weak*-compact, so that there is a subnet (ψ_j) and some $\psi^\varepsilon_\pm \in \mathcal{M}_*^+$ such that $\psi^\varepsilon_{j,\pm} \to \psi^\varepsilon_\pm$. We therefore have $\psi^\varepsilon_+ - \psi^\varepsilon_- = \lim \psi^\varepsilon_{j,+} - \psi^\varepsilon_{j,-} = 2\psi$ and by weak*-lower semicontinuity of F_ρ,

$$\begin{aligned}\Psi_\rho(\psi) &\le \frac{1}{2}[F_\rho(\psi^\varepsilon_+) + F_\rho(\psi^\varepsilon_-)] + \rho(1) \le \liminf_j \frac{1}{2}[F_\rho(\psi^\varepsilon_{j,+}) + F_\rho(\psi^\varepsilon_{j,-})]\\ &\quad + \rho(1) \le c + \epsilon.\end{aligned}$$

Since this holds for all $\varepsilon > 0$, we have $\Psi_\rho(\psi) \le c$.

□

Proposition 9 *Ψ_ρ is the Legendre-Fenchel conjugate of Φ_ρ, with respect to the dual pair $(\mathcal{M}^s, (\mathcal{M}^*)^s)$. In particular, Ψ_ρ is a Young function on $(\mathcal{M}^*)^s$.*

Proof Since F_ρ is weak*-lower semicontinuous, we see that $C^*_\rho = F^{**}_\rho = F_\rho$. Let D_ρ be given by $D_\rho(a) = C_\rho(-a)$ for $a \in \mathcal{M}^s$, then $D^*_\rho(\psi) = C^*_\rho(-\psi)$ for $\psi \in (\mathcal{M}^*)^s$. By [29, Cor. 2.3.5] and the fact that Ψ_ρ is weak*-lower semicontinuous, we obtain $\Psi_\rho = \Phi^*_\rho$, so that Ψ_ρ is a Young function on $(\mathcal{M}^*)^s$ by [9, Lemma 3.4]. □

2.5 The spaces $E_{\exp}(\mathcal{M}, \rho)$ and $L_{\log}(\mathcal{M}, \rho)$

Using the Young functions Φ_ρ and Ψ_ρ, we construct the corresponding Banach spaces B_{Φ_ρ} and B_{Ψ_ρ} as in Sect. 2.1. The following is a consequence of the results of Sect. 2.4 and [9, Prop. 2].

Proposition 10 *We have $V_{\Phi_\rho} = \mathcal{M}^s$ and $B_{\Psi_\rho} = V_{\Psi_\rho} = \mathcal{P}_\rho - \mathcal{P}_\rho$. Moreover, $B_{\Psi_\rho} = B^*_{\Phi_\rho}$ (with equivalent norms) and we have the continuous embeddings*

$$\mathcal{M}^s \sqsubseteq B_{\Phi_\rho}, \qquad B_{\Psi_\rho} \sqsubseteq \mathcal{M}^s_*.$$

Let us now look at the case when $\mathcal{M}$ is commutative. Since ρ is faithful, $\mathcal{M}$ can be identified with the space $L_\infty(\Omega, \Sigma, \rho)$ where ρ is a finite measure on (Ω, Σ). Let $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(x) = \cosh(x) - 1$ and let ψ be its conjugate, then ψ satisfies the Δ_2 condition $\psi(2u) \leq K\psi(u)$ for some $K > 0$. The exponential Orlicz space $L^\phi(\Omega, \Sigma, \rho)$ is the dual space of $L^\psi(\Omega, \Sigma, \rho)$. Since the measure ρ is finite, we have $L_\infty(X, \Sigma, \rho) \subseteq L^\phi(X, \Sigma, \rho)$ and one can see that the norm obtained from our construction coincides with the Luxemburg-Nakano norm in $L^\phi(X, \Sigma, \rho)$. Hence B_{Φ_ρ} coincides with the closure $E^\phi(X, \Sigma, \rho)$ of $L_\infty(X, \Sigma, \rho)$ in $L^\phi(X, \Sigma, \rho)$. We then have

$$L^\psi(X, \Sigma, \rho) = E^\phi(X, \Sigma, \rho)^* = B_{\Psi_\rho}$$

and $L^\phi(X, \Sigma, \rho)$ coincides with the second dual $B^{**}_{\Phi_\rho}$, see [12] for details. These facts were also pointed out in [33]. It is therefore reasonable to identify the noncommutative counterpart of L^ψ with the space B_{Ψ_ρ}, while the noncommutative exponential Orlicz space should be identified with $B^{**}_{\Phi_\rho} = B^*_{\Psi_\rho}$. Nevertheless, we will work with the more tractable space B_{Φ_ρ}, which is a strict subset of $B^{**}_{\Phi_\rho}$ in general.

Let us denote $E_{\exp}(\mathcal{M}, \rho) := B_{\Phi_\rho}$, with the norm $\|\cdot\|_{\exp,\rho} := \|\cdot\|_{\Phi_\rho}$ and $L_{\log}(\mathcal{M}, \rho) := B_{\Psi_\rho}$, with the norm $\|\cdot\|_{\log,\rho} := \|\cdot\|_{\Psi_\rho}$. In the rest of this section, we will identify $\mathcal{M}_*$ with $L_1(\mathcal{M})$, so that $\mathcal{M}^s_*$ is identified with the space $L_1(\mathcal{M})^s$ of self-adjoint elements and $\mathcal{M}^+_*$ with the cone $L_1(\mathcal{M})^+$ of positive elements in $L_1(\mathcal{M})$.

Theorem 11 *(i)* $L_{\log}(\mathcal{M}, \rho) = \mathcal{P}_\rho - \mathcal{P}_\rho$ *is dense in* $L_1(\mathcal{M})^s$ *and we have*

$$L_p(\mathcal{M}, \rho)^s \sqsubseteq L_{\log}(\mathcal{M}, \rho) \sqsubseteq L_1(\mathcal{M})^s, \qquad 1 < p \leq \infty.$$

(ii) $L_{\log}(\mathcal{M}, \rho)^+ := L_{\log}(\mathcal{M}, \rho) \cap L_1(\mathcal{M})^+ = \mathcal{P}_\rho$ *is a closed convex cone in* $L_{\log}(\mathcal{M}, \rho)$.

(iii) Let $\psi \in L_{\log}(\mathcal{M}, \rho)$. *Then* $\|\psi\|_{\log,\rho} \leq 1$ *if and only if there are some* $\omega_\pm \in \mathcal{P}_\rho$ *such that* $\psi = \frac{1}{2}(\omega_+ - \omega_-)$ *and*

$$F_\rho(\omega_+) + F_\rho(\omega_-) \leq 2 - 2\rho(1).$$

Proof By Proposition 4 (iv), we see that $\mathrm{Dom}(\Psi_\rho) = \mathcal{P}_\rho - \mathcal{P}_\rho$ contains the self-adjoint part $L_p(\mathcal{M}, \rho)^s = L_p(\mathcal{M}, \rho)^+ - L_p(\mathcal{M}, \rho)^+$ of $L_p(\mathcal{M}, \rho)$, for any $p > 1$, so $L_{\log}(\mathcal{M}, \rho)$ is dense in $L_1(\mathcal{M})^s$. This also shows that Ψ_ρ is finite valued on $L_p(\mathcal{M}, \rho)^s$. Since $L_p(\mathcal{M}, \rho)^s \sqsubseteq L_1(\mathcal{M})^s$ and Ψ_ρ is lower semicontinuous on $L_1(\mathcal{M})^s$, the restriction of Ψ_ρ defines a convex and lower semicontinuous Young function $L_p(\mathcal{M}, \rho)^s \to \mathbb{R}$, which then must be continuous by [28, Cor. 2.5]. Let B be the corresponding Banach space, then by construction, B is a closed subspace in $L_{\log}(\mathcal{M}, \rho)$ and using again [9, Prop. 2], we have

$$L_p(\mathcal{M}, \rho)^s \sqsubseteq B \sqsubseteq L_{\log}(\mathcal{M}, \rho).$$

Let now $\omega \in L_{\log}(\mathcal{M}, \rho)^+$, then there are some $\omega_\pm \in \mathcal{P}_\rho$ such that $2\omega = \omega_+ - \omega_-$. It follows that $2\omega + \omega_- = \omega_+ \in \mathcal{P}_\rho$. By Proposition 4 (iv), this implies that we must have $\omega \in \mathcal{P}_\rho$ as well. The fact that the cone is closed in $L_{\log}(\mathcal{M}, \rho)$ follows by the continuous embedding in $L_1(\mathcal{M})^s$.

Assume that $\|\psi\|_{\log,\rho} \leq 1$, equivalently, $\Psi_\rho(\psi) \leq 1$. Then for any $n \in \mathbb{N}$, there are some $\omega_{\pm,n} \in \mathcal{P}_\rho$ such that $\psi = \frac{1}{2}(\omega_{+,n} - \omega_{-,n})$ and $F_\rho(\omega_{+,n}) + F_\rho(\omega_{-,n}) \leq 2(1+1/n-\rho(1))$. It then follows that $\omega_{\pm,n} \in \mathcal{S}_C$ for some C and all n. By Proposition 4 (iii), there is some subsequence such that $\omega_{\pm,n_k} \to \omega_\pm$ in the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology. It follows that $\psi = \frac{1}{2}(\omega_+ - \omega_-)$ and by lower semicontinuity,

$$F_\rho(\omega_+) + F_\rho(\omega_-) \leq \liminf F_\rho(\omega_{+,n_k}) + F_\rho(\omega_{-,n_k}) \leq 2 - 2\rho(1).$$

The converse is obvious. □

Let us now recall the embedding $i_{\infty,\rho} : \mathcal{M}^s \to L_1(\mathcal{M})^s$, given by (1). Note that $i_{\infty,\rho}(\mathcal{M}^s) = L_\infty(\mathcal{M}, \rho)^s \sqsubseteq L_p(\mathcal{M}, \rho)^s$, for all $1 \leq p \leq \infty$.

Theorem 12 *For any $1 \leq p < \infty$, $i_{\infty,\rho}$ extends to a continuous embedding*

$$i_{\exp,\rho} : E_{\exp}(\mathcal{M}, \rho) \to L_p(\mathcal{M}, \rho)$$

and $i_{\exp,\rho}(E_{\exp}(\mathcal{M}, \rho))$ is dense in $L_p(\mathcal{M}, \rho)^s$.

Proof Let $a \in \mathcal{M}^s$, $1 \leq p < \infty$ and let $1/p + 1/q = 1$. By Theorem 11, we have $L_q(\mathcal{M}, \rho)^s \sqsubseteq L_{\log}(\mathcal{M}, \rho)$. It follows that for any $k \in L_q(\mathcal{M}, \rho)$, we have

$$\langle i_{\infty,\rho}(a), k\rangle = \mathrm{Tr}[ak] \leq \|a\|_{\exp,\rho}\|k\|_{\log,\rho}.$$

Since $\|k\|_{\log,\rho} \leq M\|k\|_{q,\rho}$ for some $M > 0$, this shows that $i_{\infty,\rho} : \mathcal{M}^s \to L_p(\mathcal{M}, \rho)^s$ is continuous with respect to the norm $\|\cdot\|_{\exp,\rho}$ in $\mathcal{M}^s$ and therefore has a unique continuous extension $i_{\exp,\rho}$. The rest follows from the fact that $i_{\infty,\rho}(\mathcal{M}^s) = L_\infty(\mathcal{M}, \rho)^s$ is dense in $L_p(\mathcal{M}, \rho)^s$ for any p.

□

To summarize, we have for $1 < p < \infty$:

$$L_\infty(\mathcal{M}, \rho) \sqsubseteq E_{\exp}(\mathcal{M}, \rho) \sqsubseteq L_p(\mathcal{M}, \rho) \sqsubseteq L_{\log}(\mathcal{M}, \rho) \sqsubseteq L_1(\mathcal{M}). \tag{8}$$

Note that we have analogous properties for the classical exponential Orlicz spaces, [2, Prop. 8].

Proposition 13 *Let $T : L_1(\mathcal{M}) \to L_1(\mathcal{N})$ be a channel. The restriction of T to $L_{\log}(\mathcal{M}, \rho)$ is a contraction $L_{\log}(\mathcal{M}, \rho) \to L_{\log}(\mathcal{N}, T(\rho))$. The adjoint T^* extends uniquely to a contraction $E_{\exp}(\mathcal{N}, T(\rho)) \to E_{\exp}(\mathcal{M}, \rho)$.*

Proof By Proposition 5, we have $F_{T(\rho)}(T(\omega)) \leq F_\rho(\omega)$. Let $\psi \in L_{\log}(\mathcal{M}, \rho)$, then

$$\Psi_{T(\rho)}(T(\psi)) \leq \frac{1}{2} \inf_{2\psi=\psi_+-\psi_-} (F_{T(\rho)}(T(\psi_+)) + F_{T(\rho)}(T(\psi_-))) + T(\rho)(1)$$

$$\leq \frac{1}{2} \inf_{2\psi=\psi_+-\psi_-} (F_\rho(\psi_+) + F_\rho(\psi_-)) + \rho(1) = \Psi_\rho(\psi).$$

By definition of $L_{\log}(\mathcal{M}, \rho)$ and duality, this implies the statement.

□

3 The quantum exponential manifold

Let $\tilde{\mathcal{F}}$ denote the set of all faithful functionals in the cone $\mathcal{M}_*^+$. In this section we will construct a Banach manifold structure on $\tilde{\mathcal{F}}$, using an extension of Theorem 6 to perturbations in $E_{\exp}(\mathcal{M}, \rho)$.

3.1 Extended perturbations

Since the effective domain of the relative entropy function F_ρ is the positive cone $\mathcal{P}_\rho = L_{\log}(\mathcal{M}, \rho)^+$, it can be regarded as a strictly convex function on $L_{\log}(\mathcal{M}, \rho)$. In this section, we will investigate the function F_ρ and its conjugate with respect to the dual pair $(L_{\log}(\mathcal{M}, \rho), E_{\exp}(\mathcal{M}, \rho))$.

We first note that F_ρ as a function on $L_{\log}(\mathcal{M}, \rho)$ is weak*-lower semicontinuous. Indeed, since $\mathcal{M}^s$ is norm dense in $E_{\exp}(\mathcal{M}, \rho)$, the weak*-topology on $L_{\log}(\mathcal{M}, \rho)$ coincides with the restriction of the $\sigma(\mathcal{M}_*^s, \mathcal{M}^s)$-topology on norm-bounded subsets. By Proposition 4 (iii), the claim will follow by the next Lemma.

Lemma 14 *For each $C \in \mathbb{R}$, $\mathcal{S}_C$ is norm-bounded in $L_{\log}(\mathcal{M}, \rho)$.*

Proof We may assume that $C \geq -\rho(1)$, otherwise $\mathcal{S}_C$ is empty. If $\omega \in \mathcal{S}_C$, then $\omega \in \mathcal{M}_*^+$ and we have (using the decomposition $2(\frac{1}{2}\omega) = \omega - 0$)

$$\Psi_\rho\left(\frac{1}{2}\omega\right) \leq \frac{1}{2} F_\rho(\omega) + \rho(1) \leq \frac{1}{2} C + \rho(1).$$

If $\Psi_\rho(\frac{1}{2}\omega) \leq 1$, then $\|\omega\|_{\log,\rho} \leq 2$, otherwise we have by [9, Lemma 3.3] that $\|\frac{1}{2}\omega\|_{\log,\rho} \leq \Psi_\rho(\frac{1}{2}\omega) \leq \frac{1}{2}C + \rho(1)$. Hence $\|\omega\|_{\log,\rho} \leq \max\{2, C + 2\rho(1)\}$. □

Let us now recall the Legendre-Fenchel conjugate function C_ρ, defined in (5). It is easily seen that C_ρ is bounded over the unit ball with respect to $\|\cdot\|_{\exp,\rho}$ in $\mathcal{M}^s$. By [28, Cor. 2.4], this implies that C_ρ is continuous (in fact, locally Lipschitz) with respect to this norm. It follows that C_ρ extends uniquely to a continuous function $E_{\exp}(\mathcal{M}, \rho) \to \mathbb{R}$, which will be again denoted by C_ρ. The next result shows that this extension is the conjugate function of F_ρ with respect to the dual pair $(L_{\log}(\mathcal{M}, \rho), E_{\exp}(\mathcal{M}, \rho))$.

Theorem 15 *For $h \in E_{\exp}(\mathcal{M}, \rho)$, we have*

$$C_\rho(h) = \sup_{\omega \in L_{\log}(\mathcal{M},\rho)} \omega(h) - F_\rho(\omega).$$

The supremum is attained at a unique functional $\rho^h \in \mathcal{P}_\rho$. Moreover, ρ^h is faithful, $C_\rho(h) = \rho^h(1)$ and the map $E_{\exp}(\mathcal{M}, \rho) \ni h \mapsto \rho^h \in \mathcal{M}_*$ is norm-to-norm continuous.

Proof Let $a_n \in \mathcal{M}^s$ be a sequence such that $\|h - a_n\|_\rho \to 0$. By putting $a = 2a_n$ and $b = a_n$ in Lemma 7, we obtain the inequality

$$C_\rho(2a_n) - C_\rho(a_n) \geq \rho^{a_n}(a_n), \qquad \forall n.$$

By continuity of C_ρ, this implies that $\{\rho^{a_n}(a_n)\}_n$ is a bounded sequence, so that also

$$\{F_\rho(\rho^{a_n}) = \rho^{a_n}(a_n) - C_\rho(a_n)\}_n$$

is bounded and therefore $\rho^{a_n} \in \mathcal{S}_K$ for some K. By Proposition 4 (iii) we may assume (by restricting to a subsequence) that there is some $\sigma \in \mathcal{S}_K$ such that $\rho^{a_n} \to \sigma$ in the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology. Since $\mathcal{S}_K$ is norm bounded in $L_{\log}(\mathcal{M}, \rho)$ (Lemma 14) and the weak*-topology coincides with the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology on $\mathcal{S}_K$, it can be seen that $\rho^{a_n}(a_n) \to \sigma(h)$. For $\omega \in \mathcal{P}_\rho$, we get by definition of ρ^{a_n} and lower semicontinuity

$$\begin{aligned}\omega(h) - F_\rho(\omega) &= \lim_n(\omega(a_n) - F_\rho(\omega)) \leq \lim_n C_\rho(a_n) = \lim_n(\rho^{a_n}(a_n) - F_\rho(\rho^{a_n}))\\ &\leq \sigma(h) - F_\rho(\sigma).\end{aligned}$$

It follows that σ is a maximizer of $\omega(h) - F_\rho(\omega)$ and by strict convexity of F_ρ such maximizer is unique. Let us denote $\rho^h := \sigma$. Note that we have $C_\rho(h) = \lim_n C_\rho(a_n) = \lim_n \rho^{a_n}(1) = \rho^h(1)$ and the above computation also implies that

$$C_\rho(h) \leq \rho^h(h) - F_\rho(\rho^h) = \sup_{\omega \in \mathcal{P}_\rho} \omega(h) - F_\rho(\omega).$$

On the other hand, we obtain using (6) and lower semicontinuity of S:

$$S(\omega\|\rho) = \lim_n(\omega(a_n) + S(\omega\|\rho^{a_n})) \geq \omega(h) + S(\omega\|\rho^h), \quad \forall \omega \in \mathcal{P}_\rho. \tag{9}$$

Putting $\omega = \rho$, we see that $S(\rho\|\rho^h)$ is finite, so that ρ^h must be faithful. Further, putting $\omega = \rho^h$ it follows that $\rho^h(h) - F_\rho(\rho^h) = \rho^h(1) = C_\rho(h)$. We also get $\lim_n S(\rho^h\|\rho^{a_n}) = 0$, so that $\rho^{a_n} \to \rho^h$ strongly in $\mathcal{M}_*$, this is easily extended to all sequences $h_n \to h$ in $E_{\exp}(\mathcal{M}, \rho)$.

□

We now extend the equalities in Theorem 6 to all elements in $E_{\exp}(\mathcal{M}, \rho)$.

Lemma 16 *Let $h \in E_{\exp}(\mathcal{M}, \rho)$.*

(i) For $a \in \mathcal{M}^s$, we have the chain rule

$$\rho^{h+a} = (\rho^h)^a, \qquad C_\rho(h + a) = C_{\rho^h}(a).$$

(ii) The norm $\|\cdot\|_{\exp,\rho^h}$ is continuous with respect to $\|\cdot\|_{\exp,\rho}$ on $\mathcal{M}^s$.

Proof Let $a_n \in \mathcal{M}^s$ be a sequence such that $\|h - a_n\|_{\exp,\rho} \to 0$. For $a \in \mathcal{M}^s$ we have $C_\rho(a_n + a) \to C_\rho(h + a)$ and $\rho^{a_n+a} \to \rho^{h+a}$ strongly, by Theorem 15. Since also $\rho^{a_n} \to \rho^h$ strongly, we have $(\rho^{a_n})^a \to (\rho^h)^a$ strongly, by [34, Thm. 1.1]. By the chain rule (7), we obtain

$$\rho^{h+a} = (\rho^h)^a, \qquad C_\rho(h + a) = \rho^{h+a}(1) = (\rho^h)^a(1) = C_{\rho^h}(a).$$

To prove (ii), note that (i) implies

$$\Phi_{\rho^h}(a) = \frac{1}{2}(C_\rho(h + a) - C_\rho(h) + C_\rho(h - a) - C_\rho(h)), \qquad a \in \mathcal{M}^s.$$

By continuity of C_ρ, this shows that there is some $\delta > 0$ such that $\Phi_{\rho^h}(a) < 1$ whenever $\|a\|_{\exp,\rho} < \delta$, this proves (ii). □

Theorem 17 *Let $h \in E_{\exp}(\mathcal{M}, \rho)$. Then*

$$\omega(h) + S(\omega\|\rho^h) = S(\omega\|\rho), \qquad \omega \in \mathcal{P}_\rho. \tag{10}$$

Moreover, $E_{\exp}(\mathcal{M}, \rho) = E_{\exp}(\mathcal{M}, \rho^h)$ (equivalent norms) and we have the chain rule

$$\rho^{h+k} = (\rho^h)^k, \qquad C_\rho(h + k) = C_{\rho^h}(k), \qquad h, k \in E_{\exp}(\mathcal{M}, \rho). \tag{11}$$

Proof Let $a_n \in \mathcal{M}^s$, $\|a_n - h\|_{\exp,\rho} \to 0$. By Lemma 16 and Theorem 15, we obtain that also $h \in E_{\exp}(\mathcal{M}, \rho^h)$ and $\|a_n - h\|_{\exp,\rho^h} \to 0$. Moreover,

$$(\rho^h)^{-h} = \lim_n (\rho^h)^{-a_n} = \lim_n \rho^{h-a_n} = \rho^0 = \rho.$$

Replacing ρ by ρ^h and h by $-h$ in (9), we obtain

$$S(\omega\|\rho^h) \geq -\omega(h) + S(\omega\|\rho).$$

Together with (9), this implies (10). Similarly, using this replacement in Lemma 16 (ii), we obtain that $E_{\exp}(\mathcal{M}, \rho) = E_{\exp}(\mathcal{M}, \rho^h)$ with equivalent norms. The chain rule (11) is now proved from (10) exactly as in the proof of Theorem 6. □

Corollary 18 *With respect to the dual pair $(L_{\log}(\mathcal{M}, \rho), E_{\exp}(\mathcal{M}, \rho))$, we have $C_\rho = F_\rho^*$ and $F_\rho = C_\rho^*$. Moreover, C_ρ is strictly convex and Gateaux differentiable on $E_{\exp}(\mathcal{M}, \rho)$, with the Gateaux derivative $C'_\rho(h) = \rho^h$, and $h \mapsto \rho^h$ defines an injective and norm-to-weak*-continuous map $E_{\exp}(\mathcal{M}, \rho) \to L_{\log}(\mathcal{M}, \rho)$.*

Proof The first part is clear from Theorem 15 and weak*-lower semicontinuity of F_ρ. Differentiability of C_ρ is then obtained from [28, Prop. 5.3]. Injectivity of the map $h \mapsto \rho^h$ follows by Theorem 17, this also implies strict convexity of C_ρ (e.g. as in the proof of [9, Thm. 7.3]). For continuity, see e.g. [29]. □

3.2 Exponential families in $\mathcal{M}_*^+$

Let $E \subseteq E_{\exp}(\mathcal{M}, \rho)$ be a closed subspace. The set

$$\mathcal{E}_\rho(E) := \{\rho^h,\ h \in E\}$$

will be called an exponential family (at ρ). The set $\mathcal{E}_\rho := \mathcal{E}_\rho(E_{\exp}(\mathcal{M}, \rho))$ will be called the full exponential family (at ρ). For the following characterization of elements of $\mathcal{E}_\rho$, note that by (4)

$$\omega \mapsto S(\omega\|\rho) - S(\omega\|\sigma)$$

defines an affine map $h_{\sigma,\rho} : \mathcal{P}_\rho \to [-\infty, \infty)$ such that $h_{\sigma,\rho}(0) = 0$.

Corollary 19 *Let $\sigma \in \mathcal{M}_*^+$. Then $\sigma = \rho^h$ for some $h \in E_{\exp}(\mathcal{M}, \rho)$ if and only if there is some $C > -\rho(1)$ such that $h_{\sigma,\rho}$ is bounded and $\sigma(\mathcal{M}_*, \mathcal{M})$-continuous on the set $\mathcal{S}_C$. In this case h coincides with $h_{\sigma,\rho}$ on $\mathcal{P}_\rho$.*

Proof Assume that $\sigma = \rho^h$ for $h \in E_{\exp}(\mathcal{M}, \rho)$, then by Theorem 17, we see that $h(\omega) = h_{\sigma,\rho}(\omega)$ for $\omega \in \mathcal{P}_\rho$. Since the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology coincides with the weak*-topology on $\mathcal{S}_C$, the assertion follows from $L_{\log}(\mathcal{M}, \rho) = E_{\exp}(\mathcal{M}, \rho)^*$.

Assume conversely that $h_{\sigma,\rho}$ has the stated properties on $\mathcal{S}_C$ for some $C > -\rho(1)$. Then the same is true for any $C' \in \mathbb{R}$, since for $C' > C$, there is some $t \in [0, 1]$ such that $F_\rho(t\omega + (1-t)\rho) \le tC' - (1-t)\rho(1) \le C$ for any $\omega \in \mathcal{S}_{C'}$.

Now note that by Theorem 11 (iii) and Proposition 4 (ii), the unit ball in $L_{\log}(\mathcal{M}, \rho)$ is a subset of $\mathcal{S}_C - \mathcal{S}_C$ for $C = 2\rho(1) - 1$, so that $h_{\sigma,\rho}$ extends to a bounded linear map on $L_{\log}(\mathcal{M}, \rho)$, moreover, since the weak*-topology coincides with the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology on bounded subsets in $L_{\log}(\mathcal{M}, \rho)$, this extension is weak*-continuous and hence defines an element $h \in E_{\exp}(\mathcal{M}, \rho)$. For $\omega \in \mathcal{P}_\rho$, we get

$$\omega(h) - F_\rho(\omega) = -F_\sigma(\omega) \le \sigma(1) = \sigma(h) - F_\rho(\sigma),$$

so that $\sigma = \rho^h$.

□

We are now ready to introduce a Banach manifold structure on $\tilde{\mathcal{F}}$ using the parametrization $h \mapsto \rho^h$, similarly as in [9] for the set of faithful states. For $\rho \in \tilde{\mathcal{F}}$, let V_ρ be the open unit ball in $E_{\exp}(\mathcal{M}, \rho)$ and $s_\rho : V_\rho \to \tilde{\mathcal{F}}$ the map $h \mapsto \rho^h$. We construct a C^∞-atlas on $\tilde{\mathcal{F}}$ as

$$\{(U_\rho, e_\rho),\ \rho \in \tilde{\mathcal{F}}\}$$

where $U_\rho = s_\rho(V_\rho)$ and $e_\rho = s_\rho^{-1}|_{U_\rho}$. To show that this is indeed a C^∞-atlas, it is enough to notice that if $U_{\rho_1} \cap U_{\rho_2} \neq \emptyset$, then we must have $\rho_1 = \rho_2^k$ for some $k = E_{\exp}(\mathcal{M}, \rho_2)$, and

$$e_{\phi_1}(U_{\rho_1} \cap U_{\rho_2}) = \{h_1 \in E_{\exp}(\mathcal{M}, \rho_1),\ \|h_1\|_{\exp,\rho_1} < 1,\ \|h_1 + k\|_{\exp,\rho_2} < 1\}.$$

The proof is finished similarly as in [9], using the equivalence of the two norms $\|\cdot\|_{\exp,\rho_1}$ and $\|\cdot\|_{\exp,\rho_2}$. It is also clear that the connected components of $\tilde{\mathcal{F}}$ are exactly the full exponential families $\mathcal{E}_\rho$, $\rho \in \tilde{\mathcal{F}}$.

3.3 The canonical divergence

Using Corollary 18, we can introduce a canonical divergence in the connected component $\mathcal{E}_\rho$, $\rho \in \tilde{\mathcal{F}}$, as the Bregman divergence associated with C_ρ:

$$D_\rho(h\|k) := C_\rho(h) - C_\rho(k) - \langle C'_\rho(k), h-k\rangle, \qquad h, k \in E_{\exp}(\mathcal{M}, \rho).$$

Theorem 20 *Let $\rho \in \tilde{\mathcal{F}}$, $h, k \in E_{\exp}(\mathcal{M}, \rho)$. Then*

(i) $D_\rho(h\|k) = S(\rho^k\|\rho^h) - (\rho^k - \rho^h)(1)$.
(ii) $D_\rho(h\|k) \geq 0$, *with equality if and only if* $h = k$.
(iii) The function $D_\rho : E_{\exp}(\mathcal{M}, \rho) \times E_{\exp}(\mathcal{M}, \rho) \to \mathbb{R}$ *is jointly continuous, and it is strictly convex and Gateaux differentiable in the first variable.*
(iv) D_ρ *satisfies the generalized Pythagorean relation*

$$D(h\|k) + D(k\|l) = D(h\|l) + (\rho^k - \rho^l)(k-h), \quad h, k, l \in E_{\exp}(\mathcal{M}, \rho).$$

Proof The statement (i) is obtained from Theorem 17. The rest follows by the properties of the Bregman divergence. More explicitly, (ii) can be seen from [28, Prop. 5.4] and strict convexity of C_ρ. To prove joint continuity, let h_n and k_n be two sequences such that $h_n \to h$, $k_n \to k$ in $E_{\exp}(\mathcal{M}, \rho)$. By Corollary 18, we have $C'_\rho(k) = \rho^k$. Since $k \mapsto \rho^k$ is norm to weak*-continuous, it follows that ρ^{k_n} is a norm-bounded sequence in $L_{\log}(\mathcal{M}, \rho)$, this implies that $\rho^{k_n}(h_n - k_n) \to \rho^k(h-k)$, we then have $D_\rho(\rho^{k_n}\|\rho^{h_n}) \to D_\rho(\rho^h\|\rho^k)$ by continuity of C_ρ. The rest of (iii) is straightforward from properties of C_ρ. The Pythagorean relation (iv) is clear from the definition. $\square$

3.4 Sufficient channels and invariance

Let $\mathcal{E}$ be a subset of nonzero elements in $\mathcal{M}_*^+$ and let $T : L_1(\mathcal{M}) \to L_1(\mathcal{N})$ be a channel. We say that T is sufficient for $\mathcal{E}$ if there is a channel $S : L_1(\mathcal{N}) \to L_1(\mathcal{M})$ such that

$$S \circ T(\sigma) = \sigma, \qquad \forall \sigma \in \mathcal{E}.$$

In this situation, S will be called a recovery channel for T on $\mathcal{E}$.

The notion of a sufficient channel was introduced by Petz [20, 35] in the situation when $\mathcal{E}$ is a set of states. Since the channels are trace preserving, the extension to positive functionals is straightforward.

Theorem 21 *([20, 35]) Let T be a 2-positive channel and assume that there is some faithful element $\rho \in \mathcal{E}$ such that $\mathcal{E} \subseteq \mathcal{P}_\rho$. The following are equivalent.*

(i) T *is sufficient for* $\mathcal{E}$*;*
(ii) $S(T(\sigma)\|T(\rho)) = S(\sigma\|\rho)$*;*
(iii) The Petz dual T_ρ *of* T *with respect to* ρ *is a recovery channel for* T *on* $\mathcal{E}$*.*

We will study the case when $\mathcal{E} = \mathcal{E}_\rho(E)$ is an exponential family at some $\rho \in \tilde{\mathcal{F}}$. Then the conditions of the above theorem are fulfilled.

Theorem 22 *Let* $\rho \in \tilde{\mathcal{F}}$*,* $h \in E_{\exp}(\mathcal{M}, \rho)$*. Let* $T : L_1(\mathcal{M}) \to L_1(\mathcal{N})$ *be a 2-positive channel and let* T_ρ *be the Petz dual of* T *with respect to* ρ*. The following are equivalent.*

(i) T *is sufficient with respect to* $\{\rho, \rho^h\}$*.*
(ii) $T(\rho^h) = T(\rho)^{h_0}$ *for some* $h_0 \in E_{\exp}(\mathcal{N}, T(\rho))$ *and* $h = T^*(h_0)$
(iii) $T^* \circ T^*_\rho(h) = h$*.*

Proof Since T^*_ρ defines a map $E_{\exp}(\mathcal{M}, \rho) \to E_{\exp}(\mathcal{N}, T(\rho))$, we have for $\omega_0 \in \mathcal{P}_{T(\rho)}$,

$$\begin{aligned} \omega_0(T^*_\rho(h)) - F_{T(\rho)}(\omega_0) &= F_\rho(T_\rho(\omega_0)) - F_{\rho^h}(T_\rho(\omega_0)) - F_{T(\rho)}(\omega_0) \\ &\leq -F_{\rho^h}(T_\rho(\omega_0)) \leq \rho^h(1) = C_\rho(h). \end{aligned} \tag{12}$$

Here we have used Theorem 17, monotonicity of relative entropy together with the fact that $\rho = T_\rho \circ T(\rho)$, and Proposition 4 (ii). Assume (i), then by Theorem 21 (ii) and (iii) we get

$$T(\rho^h)(T^*_\rho(h)) - F_{T(\rho)}(T(\rho^h)) = \rho^h(h) - F_\rho(\rho^h) = C_\rho(h).$$

It follows that the maximum in (12) is attained at $\omega_0 = T(\rho^h)$, so that $T(\rho^h) = T(\rho)^{h_0}$, with $h_0 = T^*_\rho(h)$. By a similar computation, we obtain

$$\omega(T^*(h_0)) - F_\rho(\omega) \leq C_{T(\rho)}(h_0), \qquad \forall \omega \in \mathcal{P}_\rho$$

and equality is attained for $\omega = \rho^h$. Hence $\rho^h = \rho^{T^*(h_0)}$, so that $h = T^*(h_0) = T^*(T^*_\rho(h))$ by injectivity of the map $h \mapsto \rho^h$. This proves that (i) implies both (ii) and (iii). Assume (ii), then we have

$$\begin{aligned} C_\rho(h) = \rho^h(1) = T(\rho^h)(1) = T(\rho)^{h_0}(1) = C_{T(\rho)}(h_0) = T(\rho^h)(h_0) - F_{T(\rho)}(T(\rho^h)) \\ = \rho^h(h) - F_{T(\rho)}(T(\rho^h)) \geq \rho^h(h) - F_\rho(\rho^h) = C_\rho(h). \end{aligned}$$

This implies (i) by Theorem 21 (ii). Finally, from (iii) and $T_\rho \circ T(\rho) = \rho$, we have

$$\begin{aligned} C_\rho(h) &\geq \sup_{\omega_0 \in \mathcal{P}_{T(\rho)}} T_\rho(\omega_0)(h) - F_\rho(T_\rho(\omega_0)) \geq \sup_{\omega_0} \omega_0(T^*_\rho(h)) - F_{T(\rho)}(\omega_0) \\ &\geq T(\rho^h)(T^*_\rho(h)) - F_{T(\rho)}(T(\rho^h)) \geq \rho^h(T^* \circ T^*_\rho(h)) - F_\rho(\rho^h) = C_\rho(h). \end{aligned}$$

This shows that $F_\rho(\rho^h) = F_{T(\rho)}(T(\rho^h))$, which implies (i) by Theorem 21. $\square$

Corollary 23 *Let $\rho \in \tilde{\mathcal{F}}$ and let $\mathcal{E} = \{\rho^h,\ h \in E_0\}$ for some subset $E_0 \subseteq E_{\exp}(\mathcal{M}, \rho)$. Let $E \subseteq E_{\exp}(\mathcal{M}, \rho)$ be the closed linear span of E_0. Let $T : L_1(\mathcal{M}) \to L_1(\mathcal{N})$ be a 2-positive channel sufficient with respect to $\mathcal{E}$. Then*

(i) T is sufficient for the exponential family $\mathcal{E}_\rho(E)$.
(ii) $T^*_\rho|_E$ is an isometric isomorphism of E onto $T^*_\rho(E)$ and we have $T(\mathcal{E}_\rho(E)) = \mathcal{E}_{T(\rho)}(T^*_\rho(E))$, and $T(\rho^h) = T(\rho)^{T^*_\rho(h)}$, for $h \in E$.

4 Conclusions

We have constructed an exponential manifold structure over the set $\tilde{\mathcal{F}}$ of faithful positive functionals on a von Neumann algebra, which in the commutative case coincides with a restriction of the Pistone-Sempi construction. The manifold is based on the Araki relative entropy and its conjugate C_ρ, playing the role of the moment generating function from the classical theory. We showed the relation of the obtained structures to Kosaki L_p spaces and proved an invariance property of the exponential manifold. Note that the function C_ρ was only proved to be Gateaux differentiable, so we do not get the full power of the Pistone-Sempi construction. Nevertheless, the manifold admits a canonical divergence satisfying a generalized Pythagorean relation.

Acknowledgements The research was supported by the grants VEGA 1/0142/20 and the Slovak Research and Development Agency grant APVV-20-0069.

Data Availibility Statement Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

Declarations

Conflict of interest There is no conflict of interests.

References

1. Pistone, G., Sempi, C.: An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. Ann. Stat. (1995). https://doi.org/10.1214/aos/1176324311
2. Pistone, G., Rogantin, M.: The exponential statistical manifold: Mean parameters, orthogonality, and space transformation. Bernoulli **5**(4), 721–760 (1999). https://doi.org/10.1214/aos/1176324311
3. Cena, A., Pistone, G.: Exponential statistical manifold. Ann. I. Stat. Math. **59**(1), 27–56 (2007). https://doi.org/10.1007/s10463-006-0096-y
4. Kunze, W.: Noncommutative Orlicz spaces and generalized Arens algebras. Math. Nachr. **147**(1), 123–138 (1990). https://doi.org/10.1002/mana.19901470114
5. Al-Rashed, M., Zegarlinski, B.: Noncommutative Orlicz spaces associated to a state. Studia Math. **180**, 199–209 (2007). https://doi.org/10.4064/sm180-3-1
6. Labuschagne, L.E.: A crossed product approach to Orlicz spaces. P. Lond. Math. Soc. **107**(5), 965–1003 (2013). https://doi.org/10.1112/plms/pdt006
7. Grasselli, M., Streater, R.F.: The quantum information manifold for ε-bounded forms. Rep. Math. Phys. **46**(3), 325–335 (2000). https://doi.org/10.1016/S0034-4877(00)90003-X
8. Streater, R.F.: Quantum Orlicz spaces in information geometry. Open Syst. Inf. Dyn. **11**(4), 359–375 (2004). https://doi.org/10.1007/s11080-004-6626-2

9. Jenčová, A.: A construction of a nonparametric quantum information manifold. J. Funct. Anal. **239**(1), 1–20 (2006). https://doi.org/10.1016/j.jfa.2006.02.007
10. Jenčová, A.: On quantum information manifolds. In: Algebraic and Geometric Methods in Statistics. Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511642401
11. Ay, N., Jost, J., Vân Lê, H., Schwachhöfer, L.: Information Geometry, vol. 64. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56478-4
12. Rao, M.M., Ren, Z.D.: Theory of Orlicz Spaces. M. Dekker, New York (1991)
13. Hiai, F.: Lectures on Selected Topics in Von Neumann Algebras. EMS Press, Berlin (2021). https://doi.org/10.4171/ELM/32
14. Terp, M.: L_p Spaces Associated with Von Neumann Algebras. Notes. Copenhagen University, Copenhagen (1981)
15. Takesaki, M.: Theory of Operator Algebras II. Encyclopaedia of Mathematical Sciences, vol. 125. Springer, Berlin (2003). https://doi.org/10.1007/978-3-662-10451-4
16. Kosaki, H.: Applications of the complex interpolation method to a von Neumann algebra: non-commutative L_p-spaces. J. Funct. Anal. **56**, 26–78 (1984). https://doi.org/10.1016/0022-1236(84)90025-9
17. Bergh, J., Löfström, J.: Interpolation Spaces: An Introduction. Springer, Berlin (1976)
18. Zolotarev, A.A.: L_p-spaces with respect to a state on a von Neumann algebra, and interpolation. Izv. VUZ Matematika **26**, 41–50 (1982)
19. Jenčová, A.: Rényi relative entropies and noncommutative L_p-spaces. Ann. Henri Poincaré **19**, 2513–2542 (2018). https://doi.org/10.1007/s00023-018-0683-5
20. Petz, D.: Sufficiency of channels over von Neumann algebras. Q. J. Math. **39**(1), 97–108 (1988)
21. Araki, H.: Relative entropy of states of von Neumann algebras I. Publ. Res. I. Math. Sci. **11**(3), 809–833 (1976)
22. Araki, H.: Relative entropy of states of von Neumann algebras II. Publ. Res. I. Math. Sci. **13**(1), 173–192 (1977)
23. Ohya, M., Petz, D.: Quantum Entropy and Its Use. Lecture Notes in Computer Science. Springer, Berlin (1993)
24. Kosaki, H.: Relative entropy of states: a variational expression. J. Oper. Theory **16**, 335–348 (1986)
25. Hiai, F.: Quantum f-Divergences in Von Neumann Algebras: Reversibility of Quantum Operations. Mathematical Physics Studies. Springer, Singapore (2021). https://doi.org/10.1007/978-981-33-4199-9
26. Takesaki, M.: Theory of Operator Algebras I. Encyclopaedia of Mathematical Sciences, vol. 124. Springer, Berlin (2002). https://doi.org/10.1007/978-1-4612-6188-9
27. Jenčová, A.: Rényi relative entropies and noncommutative L_p-spaces II. Ann. Henri Poincaré **22**, 3235–3254 (2021). https://doi.org/10.1007/s00023-021-01074-9
28. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems, vol. 28. SIAM, New York (1999)
29. Zalinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)
30. Petz, D.: A variational expression for the relative entropy. Commun. Math. Phys. **114**(2), 345–349 (1988)
31. Donald, M.J.: Relative hamiltonians which are not bounded from above. J. Funct. Anal. **91**(1), 143–173 (1990). https://doi.org/10.1016/0022-1236(90)90050-U
32. Araki, H.: Relative Hamiltonian for faithful normal states of a von Neumann algebra. Publ. Res. I. Math. Sci. **9**(1), 165–209 (1973)
33. Grasselli, M.R.: Dual connections in nonparametric classical information geometry. Ann. I. Stat. Math. **62**(5), 873–896 (2010). https://doi.org/10.1007/s10463-008-0191-3
34. Donald, M.J.: Continuity and relative Hamiltonians. Commun. Math. Phys. **136**(3), 625–632 (1991)
35. Petz, D.: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. Commun. Math. Phys. **105**(1), 123–131 (1986)