16

On quantum information manifolds

Anna Jenčová

16.1 Introduction

The aim of information geometry is to introduce a suitable geometrical structure on families of probability distributions or quantum states. For parametrised statistical models, such structure is based on two fundamental notions: the Fisher information and the exponential family with its dual mixed parametrisation, see for example (Amari 1985, Amari and Nagaoka 2000).

For the non-parametric situation, the solution was given by Pistone and Sempi (Pistone and Sempi 1995, Pistone and Rogantin 1999), who introduced a Banach manifold structure on the set $\mathcal{P}$ of probability distributions, equivalent to a given one. For each $\mu \in \mathcal{P}$, the authors considered the non-parametric exponential family at μ. As it turned out, this provides a C^∞-atlas on $\mathcal{P}$, with the exponential Orlicz spaces $L_\Phi(\mu)$ as the underlying Banach spaces, here Φ is the Young function of the form $\Phi(x) = \cosh(x) - 1$.

The present contribution deals with the case of quantum states: we want to introduce a similar manifold structure on the set of faithful normal states of a von Neumann algebra $\mathcal{M}$. Since there is no suitable definition of a non-commutative Orlicz space with respect to a state φ, it is not clear how to choose the Banach space for the manifold. Of course, there is a natural Banach space structure, inherited from the predual $\mathcal{M}_*$. But, as it was already pointed out in (Streater 2004), this structure is not suitable to define the geometry of states: for example, any neighbourhood of a state φ contains states such that the relative entropy with respect to φ is infinite.

In (Jenčová 2006), we suggest the following construction. We define a Luxemburg norm using a quantum Young function, similar to that in (Streater 2004) but restricted to the space of self-adjoint operators in $\mathcal{M}$. Then we take the completion under this norm. In the classical case, this norm coincides with the norm of Pistone and Sempi, restricted to bounded measurable functions. This is described in Section 16.2. In Section 16.3, we show that an equivalent Banach space can be obtained in a more natural and easier way, using some results of convex analysis. In the following sections, we use the results in (Jenčová 2006) to introduce the manifold, and discuss possible extensions.

Algebraic and Geometric Methods in Statistics, ed. Paolo Gibilisco, Eva Riccomagno, Maria Piera Rogantin and Henry P. Wynn. Published by Cambridge University Press.

Section 16.6 is devoted to channels, that is, completely positive unital maps between the algebras. We show that the structures we introduced are closely related to sufficiency of channels and a new characterisation of sufficiency is given. As it turns out, the new definition of the spaces provides a convenient way to deal with these problems.

16.2 The quantum Orlicz space

We recall the definition and some properties of the quantum exponential Orlicz space, as given in (Jenčová 2006).

16.2.1 Young functions and associated norms

Let V be a real Banach space and let V^* be its dual. We say that a function $\Phi : V \to \mathbb{R} \cup \{\infty\}$ is a *Young function*, if it satisfies:

(i) Φ 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 $\lim_{t\to\infty} \Phi(tx) = \infty$.

Since Φ is convex, its effective domain

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

is a convex set. Let us define the sets

$$\begin{aligned} C_\Phi &:= \{x \in V, \Phi(x) \leq 1\},\\ L_\Phi &:= \{x \in V, \exists s > 0,\ \text{such that } \Phi(sx) < \infty\}. \end{aligned}$$

Then L_Φ is the smallest vector space, containing $\mathrm{dom}(\Phi)$. Moreover, the Minkowski functional of C_Φ,

$$\|x\|_\Phi := \inf\{\rho > 0, x \in \rho C_\Phi\} = \inf\{\rho > 0, \Phi(\rho^{-1}x) \leq 1\}$$

defines a norm in L_Φ.

Let B_Φ be the completion of L_Φ under $\|\cdot\|_\Phi$. If the function Φ is finite valued, $\Phi : V \to \mathbb{R}$, (or, more generally, $0 \in \mathrm{int\ dom}(\Phi)$), then $L_\phi = V$ and the norm $\|\cdot\|_\Phi$ is continuous with respect to the original norm in V, so that we have the continuous inclusion $V \sqsubseteq B_\Phi$.

Let now $\Phi : V \to \mathbb{R}$ be a Young function and let the function $\Phi^* : V^* \to \mathbb{R} \cup \{\infty\}$ be the conjugate of Φ,

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

then Φ^* is a Young function as well. The associated norm satisfies

$$|v(x)| \leq 2\|x\|_\Phi \|v\|_{\Phi^*} \qquad x \in B_\Phi,\ v \in B_{\Phi^*}$$

(the Hölder inequality), so that each $v \in B_{\Phi^*}$ defines a continuous linear functional on B_Φ, in fact, it can be shown that

$$L_{\Phi^*} = B_{\Phi^*} = B_\Phi^* \sqsubseteq V^*$$

in the sense that the norm $\|\cdot\|_{\phi^*}$ is equivalent with the usual norm in B^*_Φ. Similarly, we have $L_\Phi = V \sqsubseteq B_\Phi \subseteq B^*_{\Phi^*}$.

16.2.2 Relative entropy

Let $\mathcal{M}$ be a von Neumann algebra in standard form. Let $\mathcal{M}^+_*$ be the set of normal positive linear functionals and $\mathfrak{S}_*$ be the set of normal states on $\mathcal{M}$. For ω and φ in $\mathcal{M}^+_*$, the relative entropy is defined as

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

where ξ_ω is the representing vector of ω in a natural positive cone and $\Delta_{\varphi,\xi_\omega}$ is the relative modular operator. Then S is jointly convex and weakly lower semicontinuous. We will also need the following identity

$$S(\psi_\lambda,\varphi) + \lambda S(\psi_1,\psi_\lambda) + (1-\lambda)S(\psi_2,\psi_\lambda) = \lambda S(\psi_1,\varphi) + (1-\lambda)S(\psi_2,\varphi) \quad (16.1)$$

where ψ_1, ψ_2 are normal states and $\psi_\lambda = \lambda\psi_1 + (1-\lambda)\psi_2$, $0 \leq \lambda \leq 1$. This implies that S is strictly convex in the first variable.

Let us denote

$$\begin{aligned} \mathcal{P}_\varphi &:= \{\omega \in \mathcal{M}^+_*, S(\omega,\varphi) < \infty\} \\ \mathcal{S}_\varphi &:= \{\omega \in \mathfrak{S}_*,\ S(\omega,\varphi) < \infty\} \\ K_{\varphi,C} &:= \{\omega \in \mathfrak{S}_*,\ S(\omega,\varphi) \leq C\}, \qquad C > 0. \end{aligned}$$

Then $\mathcal{P}_\varphi$ is a convex cone dense in $\mathcal{M}^+_*$ and $\mathcal{S}_\varphi$ is a convex set generating $\mathcal{P}_\varphi$. By (16.1), $\mathcal{S}_\varphi$ is a face in $\mathfrak{S}_*$. For any $C > 0$, the set $K_{\varphi,C}$ separates the elements in $\mathcal{M}$ and it is convex and compact in the $\sigma(\mathcal{M}_*,\mathcal{M})$-topology.

16.2.3 The quantum exponential Orlicz space and its dual

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

$$F_\varphi(\omega) = \begin{cases} S(\omega,\varphi) & \text{if } \omega \in \mathfrak{S}_* \\ \infty & \text{otherwise.} \end{cases}$$

Then F_φ is strictly convex and lower semicontinuous; with $\operatorname{dom}(F_\varphi) = \mathcal{S}_\varphi$. Its conjugate

$$F^*_\varphi(h) = \sup_{\omega\in\mathfrak{S}_*} \omega(h) - S(\omega,\varphi)$$

is convex and lower semicontinuous; in fact, being finite valued, it is continuous on $\mathcal{M}_s$. We have $F^{**}_\varphi = F_\varphi$ on $\mathcal{M}^*_s$.

We define the function $\Phi_\varphi : \mathcal{M}_s \to \mathbb{R}$ by

$$\Phi_\varphi(h) = \frac{\exp(F^*_\varphi(h)) + \exp(F^*_\varphi(-h))}{2} - 1.$$

Then Φ_φ is a Young function. Let us denote $\|h\|_\varphi := \|h\|_{\Phi_\varphi}$ and $B_\varphi := B_{\Phi_\varphi}$, then we call B_φ the *quantum exponential Orlicz space*.

Let $h \in \mathcal{M}_s$, $\|h\|_\varphi \leq 1$. Then

$$\cosh(\omega(h)) \leq 2e^{S(\omega,\varphi)}.$$

It follows that each $\omega \in \mathcal{S}_\varphi$ defines a continuous linear functional on B_φ. We denote by $B_{\varphi,0}$ the Banach subspace of centred elements in B_φ, that is, $h \in B_\varphi$ with $\varphi(h) = 0$. Then

$$\Phi_{\varphi,0}(h) = \frac{F_\varphi^*(h) + F_\varphi^*(-h)}{2}$$

is a Young function on $\mathcal{M}_{s,0} := \{h \in \mathcal{M}_s, \varphi(h) = 0\}$ and it defines an equivalent norm in $B_{\varphi,0}$.

Remark 16.1 Let $\mathcal{M}$ be commutative, then $\mathcal{M} = L_\infty(X, \Sigma, \mu)$ for some measure space (X, Σ, μ) with σ-finite measure μ. Then φ is a probability measure on Σ, with the density $p := d\varphi/d\mu \in L_1(X, \Sigma, \mu)$. For any Hermitian element $u \in \mathcal{M}$, $F_\varphi^*(u) = \log \int \exp(u) p d\mu$, so that

$$\Phi_\varphi(u) = \int \cosh(u) p d\mu - 1.$$

It follows that in this case, our space B_φ coincides with the closure $M_\Phi(\varphi)$ of $L_\infty(X, \Sigma, \varphi)$ in $L_\Phi(\varphi)$.

Let us now describe the dual space $B_{\varphi,0}^*$. It was proved that $B_\varphi^* = \mathcal{P}_\varphi - \mathcal{P}_\varphi$ and $B_{\varphi,0}^* = \cup_n n(K_{\varphi,1} - K_{\varphi,1})$. If we denote by $C_{\varphi,0}$ the closed unit ball in $B_{\varphi,0}^*$, then

$$C_{\varphi,0} \subseteq K_{\varphi,1} - K_{\varphi,1} \subseteq 4C_{\varphi,0} \tag{16.2}$$

so that any element in $C_{\varphi,0}$ can be written as a difference of two states in $K_{\varphi,1}$.

For v in $\mathcal{S}_\varphi - \mathcal{S}_\varphi$, let $L_v := \{\omega_1, \omega_2 \in \mathcal{S}_\varphi,\ v = \omega_1 - \omega_2\}$. We define the function $\Psi_{\varphi,0} :\ \mathcal{M}_{s,0}^* \to \mathbb{R}^+$ by

$$\Psi_{\varphi,0}(v) = \begin{cases} \inf_{L_v} S(\omega_1, \varphi) + S(\omega_2, \varphi) & \text{if } v \in \mathcal{S}_\varphi - \mathcal{S}_\varphi \\ \\ \infty & \text{otherwise.} \end{cases}$$

Then $\Psi_{\varphi,0}$ is a Young function and it was proved that

$$\Phi_{\varphi,0}^*(v) = 1/2 \Psi_{\varphi,0}(2v)$$

for $v \in \mathcal{M}_{s,0}^*$. It follows that the norm in $B_{\varphi,0}^*$ is equivalent with $\|\cdot\|_{\Psi_{\varphi,0}}$.

16.3 The spaces $A(K_\varphi)$ and $A(K_\varphi)^{**}$

In this section, we use a well-known representation of compact convex sets, see for example (Asimow and Ellis 1980) for details. We obtain a Banach space, which turns out to be equivalent to $B_{\varphi,0}$.

Let $K \subset \mathfrak{S}_*$ be a convex set, compact in the $\sigma(\mathcal{M}_*, \mathcal{M})$-topology and separating the points in $\mathcal{M}_s$. In particular, let $K_\varphi := K_{\varphi,1}$. Let $A(K)$ be the Banach space of continuous affine functions $f : K \to \mathbb{R}$, with the supremum norm. Then K can be identified with the set of states on $A(K)$, where each element $\omega \in K$ acts on $A(K)$ by evaluation $f \mapsto f(\omega)$. Moreover, the topology of K coincides with the weak*-topology of the state space.

It is clear that any self-adjoint element in $\mathcal{M}$ belongs to $A(K)$, moreover, $\mathcal{M}_s$ is a linear subspace in $A(K)$, separating the points in K and containing all the constant functions. It follows that $\mathcal{M}_s$ is norm-dense in $A(K)$.

The dual space $A(K)^*$ is the set of all elements of the form

$$p : f \mapsto a_1 f(\omega_1) - a_2 f(\omega_2)$$

for some $\omega_1, \omega_2 \in K$, $a_1, a_2 \in \mathbb{R}^+$, so that $A(K)^*$ is a real linear subspace in $\mathcal{M}_*$. The embedding of $A(K)^*$ to $\mathcal{M}_*$ is continuous and the weak*-topology on $A(K)^*$ coincides with $\sigma(\mathcal{M}_*, \mathcal{M})$ on bounded subsets. It is also easy to see that the second dual $A(K)^{**}$ is the set of all bounded affine functionals on K.

Let $L \subseteq K$ be convex and compact. For $f \in A(K)^{**}$, the restriction to L is in $A(L)^{**}$, continuous if $f \in A(K)$ and such that $\|f|_L\|_L \leq \|f\|_K$.

Lemma 16.1 *Let $a, b > 0$, then $A(K_{\varphi,a}) = A(K_{\varphi,b})$ and $A(K_{\varphi,a})^{**} = A(K_{\varphi,b})^{**}$, in the sense that the corresponding norms are equivalent.*

Proof Suppose that $a \geq b$. Since $K_{\varphi,b} \subseteq K_{\varphi,a}$, it follows that $A(K_{\varphi,a}) \subseteq A(K_{\varphi,b})$ and $A(K_{\varphi,a})^{**} \subseteq A(K_{\varphi,b})^{**}$ with $\|f\|_{\varphi,b} \leq \|f\|_{\varphi,a}$ for $f \in A(K_{\varphi,a})^{**}$. On the other hand, let $\omega \in K_{\varphi,a}$, then $\omega_t := t\omega + (1-t)\varphi \in K_{\varphi,b}$ whenever $t \leq b/a$. Then

$$\omega = a/b\omega_{b/a} - (a/b - 1)\varphi$$

so that $K_{\varphi,a}$ is contained in the closed ball with radius $(2a - b)/b$ in $A(K_{\varphi,b})^*$. It follows that any $f \in A(K_{\varphi,b})^{**}$ defines a bounded affine functional over $K_{\varphi,a}$, continuous if $f \in A(K_{\varphi,b})$ and

$$\|f\|_{\varphi,a} = \sup_{\omega \in K_{\varphi,a}} |f(\omega)| \leq \|f\|_{\varphi,b}(2a - b)/b.$$

□

We see from the above proof that $\mathcal{S}_\varphi \subset A(K_{\varphi,b})^*$ and each $K_{\varphi,a}$ is weak*-compact in $A(K_{\varphi,b})^*$. It follows that $\mathrm{dom}\,(F_\varphi) \subset A(K_{\varphi,b})^*$ and F_φ is a convex weak*-lower semicontinuous functional on $A(K_{\varphi,b})^*$.

Let us denote by $A_0(K)$ the subspace of elements $f \in A(K)$, such that $f(\varphi) = 0$. Then we have

Theorem 16.1 $A_0(K_\varphi) = B_{\varphi,0}$, *with equivalent norms.*

Proof We have by (16.2) that the norms are equivalent on $\mathcal{M}_s$. The statement follows from the fact that $\mathcal{M}_s$ is dense in both spaces. □

16.4 The perturbed states

As we have seen, $F_\varphi(\omega) = S(\omega, \varphi)$ defines a convex lower semicontinuous functional $A(K_\varphi)^* \to \mathbb{R}$. Let $f \in A(K_\varphi)^{**}$. We denote

$$c_\varphi(f) := \inf_{\omega \in \mathcal{S}_\varphi} f(\omega) + S(\omega, \varphi).$$

Then $-c_\varphi(-f)$ is the conjugate functional $F_\varphi^*(f)$, so that c_φ is concave and upper semicontinuous, with values in $\mathbb{R} \cup \{-\infty\}$.

Suppose that $c_\varphi(f)$ is finite and that there is a state $\psi \in \mathcal{S}_\varphi$, such that

$$c_\varphi(f) = f(\psi) + S(\psi, \varphi).$$

Then this state is unique, this follows from the fact that S is strictly convex in the first variable. Let us denote this state by φ^f. Note that if $f \in \mathcal{M}_s$, then φ^f exists and it is the perturbed state (Ohya and Petz 1993), so that we can see the mapping $f \mapsto \varphi^f$ as an extension of state perturbation.

In (Jenčová 2006), we defined the perturbed state for elements in $B_{\varphi,0}$; we remark that there we used the notation $c_\varphi = F_\varphi^*$ and the state was denoted by $[\varphi^h]$, $h \in B_\varphi$. It was shown that $[\varphi^h]$ is defined for all $h \in B_\varphi$ and that the map

$$B_{\varphi,0} \in h \mapsto [\varphi^h]$$

can be used to define a C^∞-atlas on the set of faithful states on $\mathcal{M}$. By Theorem 16.1, we have the same for $A_0(K_\varphi)$. We will recall the construction below, but before that, we give some results obtained for $f \in A(K_\varphi)^{**}$.

First of all, it is clear that $c_\varphi(f + c) = c_\varphi(f) + c$ for any real c and $\varphi^f = \varphi^{f+c}$ if φ^f is defined. We may therefore suppose that $f \in A_0(K_\varphi)^{**}$.

Lemma 16.2 *Let $f \in A(K_\varphi)^{**}$ be such that φ^f exists. Then for all $\omega \in \mathcal{S}_\varphi$,*

$$S(\omega, \varphi) + f(\omega) \geq S(\omega, \varphi^f) + c_\varphi(f).$$

Equality is attained on the face in $\mathcal{S}_\varphi$, generated by φ^f.

Proof The statement is proved using the identity (16.1), the same way as Lemmas 12.1 and 12.2 in (Ohya and Petz 1993). □

The previous lemma has several consequences. For example, it follows that $c_\varphi(f) \leq -S(\varphi, \varphi^f) \leq 0$ if $f \in A_0(K_\varphi)^{**}$. Further, $S(\omega, \varphi^f)$ is bounded on K_φ, so that $K_\varphi \subseteq K_{\varphi^f, C}$ for some $C > 0$. It also follows that $\mathcal{S}_\varphi \subseteq \mathcal{S}_{\varphi^f}$. In particular, $S(\varphi, \varphi^f) < \infty$ and since also $S(\varphi^f, \varphi) < \infty$, the states φ and φ^f have the same support.

Lemma 16.3 *Let $\psi = \varphi^f$ for some $f \in A(K_\varphi)^{**}$. Then we have the continuous embeddings $A(K_\psi) \sqsubseteq A(K_\varphi)$ and $A(K_\psi)^{**} \sqsubseteq A(K_\varphi)^{**}$.*

Proof Follows from $K_\varphi \subseteq K_{\psi,C}$ and Lemma 16.1. □

We will now consider the set of all states φ^f, with some $f \in A(K_\varphi)^{**}$. Let ψ be a normal state, such that $\varphi \in \mathcal{S}_\psi$. We denote

$$f_\psi(\omega) := S(\omega, \psi) - S(\omega, \varphi) - S(\varphi, \psi).$$

By identity (16.1), f_ψ is an affine functional $K_\varphi \to \mathbb{R} \cup \{\infty\}$, such that $f_\psi(\varphi) = 0$.

Theorem 16.2 *Let ψ be a normal state. Then $\psi = \varphi^f$ for some $f \in A(K_\varphi)^{**}$ if and only if $\psi \in \mathcal{S}_\varphi$ and $K_\varphi \subseteq K_{\psi,C}$ for some $C > 0$.*

Proof It is clear that $\psi \in \mathcal{S}_\varphi$ if $\psi = \varphi^f$ and we have seen that also $K_\varphi \subseteq K_{\psi,C}$. Conversely, if $K_\varphi \subseteq K_{\psi,C}$, then $f_\psi \in A_0(K_\varphi)^{**}$ and

$$f_\psi(\omega) + S(\omega, \varphi) = S(\omega, \psi) - S(\varphi, \psi) \geq -S(\varphi, \psi)$$

for all $\omega \in \mathcal{S}_\varphi$. Since equality is attained for $\omega = \psi$, $\psi = \varphi^{f_\psi}$. $\square$

Note also that, by the above proof, $c_\varphi(f_\psi) = -S(\varphi, \psi)$.

16.4.1 The subdifferential

Let $\psi \in \mathcal{S}_\varphi$. The *subdifferential at* ψ is the set of elements $f \in A_0(K_\varphi)^{**}$, such that $\psi = \varphi^f$. Let us denote the subdifferential by $\partial_\varphi(\psi)$. By Theorem 16.2, the subdifferential at ψ is non-empty if and only if $K_\varphi \subseteq K_{\psi,C}$.

Lemma 16.4 *If $\partial_\varphi(\psi) \neq \emptyset$, then it is a closed convex subset in $A_0(K_\varphi)^{**}$. Moreover, c_φ is affine over $\partial_\varphi(\psi)$.*

Proof Let $f, g \in \partial_\varphi(\psi)$ and let $g_\lambda = \lambda g + (1 - \lambda) f$, $\lambda \in (0, 1)$. Then

$$g_\lambda(\psi) + S(\psi, \varphi) = \lambda c_\varphi(g) + (1 - \lambda) c_\varphi(f).$$

Since c_φ is concave, this implies that $\psi = \varphi^{g_\lambda}$ and that $c_\varphi(g_\lambda) = \lambda c_\varphi(g) + (1 - \lambda) c_\varphi(f)$. Moreover, we can write

$$\partial_\varphi(\psi) = \{g \in A(K_\varphi)^{**},\ c_\varphi(g) - g(\psi) \geq S(\psi, \varphi)\}$$

and this set is closed, since c_φ is upper semicontinuous. $\square$

Lemma 16.5 *Let $\psi \in \mathcal{S}_\varphi$, $\partial_\varphi(\psi) \neq \emptyset$ and let $g \in A_0(K_\varphi)^{**}$. Then $g \in \partial_\varphi(\psi)$ if and only if there is some $k \in \mathbb{R}$, such that*

$$g(\omega) - f_\psi(\omega) \geq k,\ \omega \in \mathcal{S}_\varphi \quad \textit{and} \quad g(\psi) - f(\psi) = k. \tag{16.3}$$

In this case, $k = c_\varphi(g) - c_\varphi(f_\psi) \leq 0$.

Proof If $g \in \partial_\varphi(\psi)$, then (16.3) follows from Lemma 16.2 and $k \leq 0$ is obtained by putting $\omega = \varphi$. Conversely, suppose that (16.3) is true, then we have for $\omega \in \mathcal{S}_\varphi$

$$g(\omega) + S(\omega, \varphi) = g(\omega) - f_\psi(\omega) + f_\psi(\omega) + S(\omega, \varphi) \geq k + c_\varphi(f_\psi)$$

and $g(\psi)+S(\psi,\omega)=k+c_\varphi(f_\psi)$, this implies that $\psi=\varphi^g$ and $c_\varphi(g)=k+c_\varphi(f_\psi)$. □

16.4.2 The chain rule

Let $\mathcal{C}_\varphi := \{\psi \in S_\varphi,\ K_\varphi \subseteq K_{\psi,C},\ K_\psi \subseteq K_{\varphi,A}$ for some $A, C>0\}$.

Theorem 16.3 *Let $\psi \in \mathcal{C}_\varphi$. Then*

(i) $\mathcal{S}_\varphi = \mathcal{S}_\psi$,
(ii) $A(K_\varphi)=A(K_\psi)$, $A(K_\varphi)^{**}=A(K_\psi)^{**}$, *with equivalent norms,*
(iii) $\varphi \in \mathcal{C}_\psi$.

Proof Let $\psi \in \mathcal{C}_\varphi$. By Theorem 16.2, $\psi=\varphi^f$, $f\in A(K_\varphi)^{**}$ and also $\varphi=\psi^g$ for some $g\in A(K_\psi)^{**}$. Now we have (i) by Lemma 16.2 and (ii) by Lemma 16.3, (iii) is obvious. □

We also have the following chain rule.

Theorem 16.4 *Let $\psi \in \mathcal{C}_\varphi$ and let $g \in A(K_\varphi)^{**}$ be such that ψ^g exists. Then*

$$c_\psi(g)=c_\varphi(g+f)-c_\varphi(f), \qquad \psi^g=\varphi^{f+g} \tag{16.4}$$

holds for $f=f_\psi$.

Proof Suppose that ψ^g exists, then

$$g(\omega)+f_\psi(\omega)+S(\omega,\varphi)=g(\omega)+S(\omega,\psi)+c_\varphi(f_\psi)\geq c_\psi(g)+c_\varphi(f_\psi)$$

for all $\omega\in\mathcal{S}_\varphi=\mathcal{S}_\psi$ and equality is attained at $\omega=\psi^g$. This implies $c_\psi(g)=c_\varphi(g+f)-c_\varphi(f)$ and $\psi^g=\varphi^{f+g}$. □

16.5 The manifold structure

Let $\mathcal{F}$ be the set of faithful normal states on $\mathcal{M}$. Let $\varphi\in\mathcal{F}$. In this section we show that we can use the map $f\mapsto\varphi^f$ to define the manifold structure on $\mathcal{F}$. So far, it is not clear if this map is well-defined or one-to-one on $A_0(K_\varphi)^{**}$. The situation is better if we restrict to $A_0(K_\varphi)$, as Theorem 16.5 shows.

Theorem 16.5 *Let $f\in A(K_\varphi)$. Then*

(i) *φ^f exists and $\varphi^f\in\mathcal{C}_\varphi$.*
(ii) *If $g\in A(K_\varphi)$ is such that $\varphi^g=\varphi^f$, then $f-g=\varphi(f-g)$.*
(iii) *In Lemma 16.2, equality is attained for all $\omega\in\mathcal{S}_\varphi$, in particular,*

$$f-f(\varphi)=f_{\varphi^f}.$$

(iv) *The chain rule (16.4) holds for all $f,g\in A(K_\varphi)$.*

Proof We may suppose that $f \in A_0(K_\varphi)$. By the results in (Jenčová 2006) and Theorem 16.1, if $f \in A_0(K_\varphi) = B_{\varphi,0}$, then $\psi = \varphi^f$ exists, $f - f(\psi) \in A_0(K_\psi) = B_{\psi,0}$ and $\varphi = \psi^{-f}$. By Theorem 16.2, $\psi \in \mathcal{C}_\varphi$ and (i) is proved. (ii),(iii) and (iv) were proved in (Jenčová 2006). □

Proposition 16.1 is not needed in our construction. It shows that each $\psi \in \mathcal{C}_\varphi$ is faithful on $A(K_\varphi)$.

Proposition 16.1 *Let $\psi \in \mathcal{C}_\varphi$ and let $g \in A(K_\varphi)$ be positive. Then $g(\psi) = 0$ implies $g = 0$.*

Proof Let g be a positive element in $A(K_\varphi) = A(K_\psi)$, with $g(\psi) = 0$, then by Lemma 16.5, $f_\psi + g \in \partial_\varphi(\psi)$. Since ψ^g exists, we have by the chain rule that $\psi^g = \varphi^{f_\psi + g} = \psi$. Since $g \in A_0(K_\psi)$, $g = 0$. □

Let us 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 B_i, and for all i, j, $e_i(U_i \cap U_j)$ is open in B_i;
(iii) the map $e_j e_i^{-1} : e_i(U_i \cap U_j) \to e_j(U_i \cap U_j)$ is a C^p-isomorphism for all i, j.

Let now $X = \mathcal{F}$. For $\varphi \in \mathcal{F}$, let V_φ be the open unit ball in $A_0(K_\varphi)$ and let $s_\varphi : V_\varphi \to \mathcal{F}$ be the map $f \mapsto \varphi^f$. By Theorem 16.5, s_φ is a bijection onto the set $U_\varphi := s_\varphi(V_\varphi)$. Let e_φ be the map $U_\varphi \ni \psi \mapsto f_\psi \in V_\varphi$. Then we have

Theorem 16.6 *(Jenčová 2006) $\{(U_\varphi, e_\varphi), \varphi \in \mathcal{F}\}$ is a C^∞-atlas on $\mathcal{F}$.*

In the commutative case, the space corresponding to $A(K_\varphi)$ is not the exponential Orlicz space L_Φ, but the subspace M_Φ, see Remark 16.1. The corresponding commutative information manifold structure was considered in (Grasselli 2009). It follows from the theory of Orlicz spaces that (under some reasonable conditions on the base measure μ)

$$M_\Phi(\mu)^* = L_{\Phi^*}(\mu), \qquad L_{\Phi^*}(\mu)^* = L_\Phi(\mu).$$

By comparing $A(K_\varphi)$ with these results, it seems that the quantum exponential Orlicz space should be the second dual $A(K_\varphi)^{**}$, rather than $A(K_\varphi)$.

To get the counterpart of the Pistone and Sempi manifold, we would need to extend the map s_φ to the unit ball V_φ^{**} in $A_0(K_\varphi)^{**}$ and show that it is one-to-one. At present, it is not clear how to prove this. At least, we can prove that c_φ is finite on V_φ^{**}.

Lemma 16.6 *Let $f \in A_0(K_\varphi)^{**}$, $\|f\| \leq 1$. Then $0 \geq c_\varphi(f) \geq -1$ and the infimum can be taken over K_φ.*

Proof Let $\omega \in \mathcal{S}_\varphi$ be such that $S(\omega, \varphi) > 1$. Since the function $t \mapsto S(\omega_t, \varphi)$ is convex and lower semicontinuous in $(0, 1)$, it is continuous and there is some

$t \in (0,1)$ such that $S(\omega_t, \varphi) = 1$, recall that $\omega_t = t\omega + (1-t)\varphi$. By strict convexity, it follows that $1 = S(\omega_t, \varphi) < tS(\omega, \varphi)$ and $S(\omega, \varphi) > 1/t$. On the other hand, $\omega_t \in K_\varphi$ and therefore $-1 \le f(\omega_t) = tf(\omega)$. It follows that

$$f(\omega) + S(\omega, \varphi) > -1/t + 1/t = 0 = f(\varphi) + S(\varphi, \varphi) \ge c_\varphi(f).$$

From this, $c_\varphi(f) = \inf_{\omega \in K_\varphi} f(\omega) + S(\omega, \varphi) \ge -1$. □

16.6 Channels and sufficiency

Let $\mathcal{N}$ be another von Neumann algebra. A *channel* from $\mathcal{N}$ to $\mathcal{M}$ is a completely positive, unital map $\alpha : \mathcal{N} \to \mathcal{M}$. We will also require that a channel is normal, then its dual $\alpha^* : \varphi \mapsto \varphi \circ \alpha$ maps normal states on $\mathcal{M}$ to normal states on $\mathcal{N}$.

An important property of such channels is that the relative entropy is monotone under these maps:

$$S(\omega \circ \alpha, \varphi \circ \alpha) \le S(\omega, \varphi), \qquad \omega, \varphi \in \mathfrak{S}_*.$$

This implies that α^* defines a continuous affine map $K_\varphi \to K_{\varphi\circ\alpha}$. If $f_0 \in A(K_{\varphi\circ\alpha})^{**}$, then composition with α^* defines a bounded affine functional over K_φ, which we denote by $\alpha(f_0)$. Then $\alpha(f_0)$ is continuous if $f_0 \in A(K_{\varphi\circ\alpha})$ and

$$\|\alpha(f_0)\| = \sup_{\omega \in K_\varphi} |f_0(\omega \circ \alpha)| \le \sup_{\omega_0 \in K_{\varphi\circ\alpha}} |f_0(\omega_0)| = \|f_0\|$$

so that α is a contraction $A(K_{\varphi\circ\alpha})^{**} \to A(K_\varphi)^{**}$ and $A(K_{\varphi\circ\alpha}) \to A(K_\varphi)$.

Lemma 16.7 *Let $\alpha : \mathcal{N} \to \mathcal{M}$ be a channel and let $g_0 \in A(K_{\varphi\circ\alpha})^{**}$. Then $c_{\varphi\circ\alpha}(g_0) \le c_\varphi(\alpha(g_0))$.*

Proof We compute

$$c_\varphi(\alpha(g_0)) = \inf_{\omega \in \mathcal{S}_\varphi} g_0(\omega \circ \alpha) + S(\omega, \varphi) \ge \inf_{\omega \in \mathcal{S}_\varphi} g_0(\omega \circ \alpha) + S(\omega \circ \alpha, \varphi \circ \alpha) \ge c_{\varphi\circ\alpha}(g_0).$$

□

Let $\mathcal{S}$ be a set of states in $\mathfrak{S}_*(\mathcal{M})$. We say that the channel $\alpha : \mathcal{N} \to \mathcal{M}$ is *sufficient for $\mathcal{S}$* if there is a channel $\beta : \mathcal{M} \to \mathcal{N}$, such that

$$\omega \circ \alpha \circ \beta = \omega, \qquad \omega \in \mathcal{S}.$$

This definition of sufficient channels was introduced in (Petz 1986), see also (Jenčová and Petz 2006a), and several characterisations of sufficiency were given. Here we are interested in the following two characterisations. For simplicity, we will assume that the states, as well as the channel, are faithful.

Theorem 16.7 *(Petz 1986) Let $\psi \in \mathcal{S}_\varphi$. The channel α is sufficient for the pair $\{\psi, \varphi\}$ if and only if $S(\psi, \varphi) = S(\psi \circ \alpha, \varphi \circ \alpha)$.*

Theorem 16.8 *(Jenčová and Petz 2006b) Let $\psi = \varphi^f$ for some $f \in \mathcal{M}_s$. Then α is sufficient for $\{\psi, \varphi\}$ if and only if there is some $g_0 \in \mathcal{N}_s$, such that $\psi \circ \alpha = (\varphi \circ \alpha)^{g_0}$ and $f = \alpha(g_0)$.*

In this section we show how Theorem 16.8 can be extended to pairs $\{\psi, \varphi\}$ such that $\partial_\varphi(\psi) \neq \emptyset$.

So let $\psi = \varphi^f$ for some $f \in A(K_\varphi)^{**}$ and suppose that $\alpha : \mathcal{N} \to \mathcal{M}$ is a sufficient channel for the set $\{\psi, \varphi\}$. Let us denote $\varphi_0 := \varphi \circ \alpha$, $\psi_0 := \psi \circ \alpha$. Let $\beta : \mathcal{M} \to \mathcal{N}$ be the channel such that $\varphi_0 \circ \beta = \varphi$, $\psi_0 \circ \beta = \psi$. We will show that $\psi_0 = \varphi_0^{\beta(f_\psi)}$. To see this, note that for $\omega_0 \in \mathcal{S}_{\varphi_0}$,

$$\beta(f_\psi)(\omega_0) = f_\psi(\omega_0 \circ \beta) = S(\omega_0 \circ \beta, \psi) - S(\omega_0 \circ \beta, \varphi) - S(\psi, \varphi).$$

Then

$$\begin{aligned}
&\beta(f_\psi)(\omega_0) + S(\omega_0, \varphi_0) \\
&= S(\omega_0 \circ \beta, \psi) + S(\omega_0, \varphi_0) - S(\omega_0 \circ \beta, \varphi_0 \circ \beta) - S(\psi, \varphi) \geq -S(\psi, \varphi) = c_\varphi(f_\psi)
\end{aligned}$$

by positivity and monotonicity of the relative entropy, and

$$\beta(f_\psi)(\psi_0) + S(\psi_0, \varphi_0) = -S(\psi, \varphi)$$

so that $c_{\varphi_0}(\beta(f_\psi)) = c_\varphi(f_\psi)$ and $\psi_0 = \varphi_0^{\beta(f_\psi)}$.

On the other hand, this implies by Theorem 16.2 that $f_{\psi_0} \in A(K_{\varphi_0})^{**}$ and we obtain in the same way that $\psi = \varphi^{\alpha(f_{\psi_0})}$ and $c_\varphi(\alpha(f_{\psi_0})) = c_{\varphi_0}(f_{\psi_0})$.

Theorem 16.9 *Let ψ be such that $\partial_\varphi(\psi) \neq \emptyset$ and let $\alpha : \mathcal{N} \to \mathcal{M}$ be a channel. Let $\varphi_0 = \varphi \circ \alpha$, $\psi_0 = \psi \circ \alpha$ The following are equivalent*

(i) *α is sufficient for the pair $\{\varphi, \psi\}$,*
(ii) *$f_{\psi_0} \in A(K_{\varphi_0})^{**}$ and $\psi = \varphi^{\alpha(f_{\psi_0})}$,*
(iii) *$c_{\varphi_0}(f_{\psi_0}) = c_\varphi(f_\psi)$.*

Proof The implication (i) $\to$ (ii) was already proved above. Suppose (ii) holds, then

$$\begin{aligned}
c_\varphi(\alpha(f_{\psi_0})) &= \alpha(f_{\psi_0})(\psi) + S(\psi, \varphi) = \\
&= -S(\psi_0, \varphi_0) - S(\varphi_0, \psi_0) + S(\psi, \varphi).
\end{aligned}$$

By putting $\omega = \varphi$ in Lemma 16.2, we obtain $c_\varphi(\alpha(f_{\psi_0})) \leq -S(\varphi, \psi)$. Then

$$0 \leq S(\psi, \varphi) - S(\psi_0, \varphi_0) \leq S(\varphi_0, \psi_0) - S(\varphi, \psi) \leq 0.$$

It follows that $c_\varphi(f_\psi) = -S(\varphi, \psi) = -S(\varphi_0, \psi_0) = c_{\varphi_0}(f_{\psi_0})$, hence (iii) holds. The implication (iii) $\to$ (i) follows from Theorem 16.7. □

In particular, if $\psi = \varphi^f$ for $f \in A(K_\varphi)$, the above theorem can be formulated as follows.

Theorem 16.10 *Let $\psi = \varphi^f$, $f \in A(K_\varphi)$ and $\alpha : \mathcal{N} \to \mathcal{M}$ be a channel. Then α is sufficient for $\{\psi, \varphi\}$ if and only if there is some $g_0 \in A(K_{\varphi_0})$, such that $\psi_0 = \varphi_0^{g_0}$ and $f = \alpha(g_0)$.*

Proof The statement follows from Theorem 16.9 and the fact that if $\psi = \varphi^f$ for $f \in A_0(K_\varphi)$, then we must have $f = f_\psi$, by Theorem 16.5. □

Acknowledgement

Supported by the Center of Excellence EU-QUTE and SAS-Quantum Technologies.

References

Amari, S. (1985). *Differential-geometrical Methods in Statistics* (New York, Springer-Verlag).

Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry* (AMS monograph, Oxford University Press).

Asimow, L. and Ellis, A. J. (1980). *Convexity Theory and its Applications in Functional Analysis* (London, Academic Press).

Grasselli, M. R. (2009). Dual connections in nonparametric classical information geometry, *Annals of the Institute of Statistical Mathematics* (to appear) (available at arXiv:math-ph/0104031v1).

Jenčová, A. (2006). A construction of a nonparametric quantum information manifold, *Journal of Functional Analysis* **239**, 1–20.

Jenčová, A. and Petz, D. (2006a). Sufficiency in quantum statistical inference, *Communications in Mathematical Physics* **263**, 259–76.

Jenčová, A. and Petz, D. (2006b). Sufficiency in quantum statistical inference. A survey with examples, *Infinite Dimensional Analysis, Quantum Probability and Related Topics* **9**, 331–51.

Ohya, M. and Petz, D. (1993). *Quantum Entropy and Its Use* (Heidelberg, Springer-Verlag).

Petz, D. (1986). Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, *Communications in Mathematical Physics* **105**, 123–31.

Pistone, G. and Rogantin, M. P. (1999). The exponential statistical manifold, mean parameters, orthogonality and space transformations, *Bernoulli* **5**, 721–60.

Pistone, G. and Sempi, C. (1995). An infinite-dimensional geometric structure on the space of all the probability measures equivalent to the given one, *Annals of Statistics* **23**, 1543–61.

Streater, R. F. (2004). Quantum Orlicz spaces in information geometry, *Open Systems and Information Dynamics* **11**, 359–75.