Quantum statistical manifolds: the linear growth case

Jan Naudts

Departement Fysica, Universiteit Antwerpen,
Universiteitsplein 1, 2610 Antwerpen, Belgium
E-mail: jan.naudts@uantwerpen.ac.be.

Abstract

A deformed exponential family of states on a von Neumann algebra is constructed for one specific deformation which makes the exponential function asymptotically linear. Difficulties arising due to non-commutativity are highlighted.

1 Introduction

In a recent publication Montrucchio and Pistone [1] treat a special case of a parameter-free deformed exponential family. The present paper shows that part of this work can be transposed to a non-commutative setting in a rather straightforward manner. Both the commutative and the non-commutative versions can be useful as an inspiration for the development of a more general theory of parameter-free information geometry. It is not the ambition of the present paper to develop such a theory, but only to clarify the kind of difficulties which one encounters in the non-commutative setting.

Amari [2, 3] introduced parametrized models of Information Geometry. Parameter-free families were introduced by Pistone and Sempi [4]. See also [5, 6]. The generalization of Information Geometry to a non-commutative context is of interest because of its applications in Quantum Mechanics. However, it is not so trivial. In the simplest context the random variables of probability theory are replaced by $n \times n$-dimensional matrices, the probability distribution is replaced by a density matrix. Then use can be made of the property known as cyclic permutation under the trace. This property
restores part of the commutativity, needed to mimic the proofs of the commutative case. A more general context involves Tomita-Takesaki theory. See for instance the recent book of Petz [7]. The aim of the present paper is to go beyond the traditional setting by not longer focusing on tracial states. Technically, this is achieved by a regularization of the exponential function in the way first applied by Newton [8] and picked up by Montrucchio and Pistone [1]. The deformation of the exponential function is used to construct deformed exponential families of probability functions and, in the present paper, of state vectors in a Hilbert space.

The interest in deformed exponential families started with the q-statistics of Tsallis [9]. A further generalization was given by the author [10, 11, 12]. The latter formalism is used here and is explained below in Section 3 for the special case of linear growth. Non-commutative deformed exponential families have been considered up to now in the simple context of finite-dimensional matrices. See Section 12.

2 Non-commutative context

A statistical manifold is a differentiable manifold $\mathcal{M}$ together with a Riemannian metric $g$ and a pair of dually flat connections [3]. The manifold $\mathcal{M}$ consists of probability distributions on a given measure space $\mathcal{X}$, $d\mathcal{x}$.

In the most simple non-commutative setting the elements of $\mathcal{M}$ are density matrices instead of probability distributions. These are symmetric matrices with non-negative eigenvalues and with trace 1. In the present work the more general $C^*$-algebraic context is chosen. A state $\omega$ on a $C^*$-algebra $\mathcal{A}$ is a linear map $A \in \mathcal{A} \mapsto \omega(A) \in \mathbb{C}$ which satisfies the positivity condition: $A \geq 0$ implies $\omega(A) \geq 0$ and the normalization condition $\omega(1) = 1$ (for convenience, it is assumed that the identity $1$ belongs to $\mathcal{A}$). Note that any density matrix $\rho$ determines a state $\omega$ of the $C^*$-algebra of all square matrices $A$ of given dimension by the relation $\omega(A) = \text{Tr} \rho A$.

Given a state $\omega$ of the $C^*$-algebra $\mathcal{A}$ there exists a $^*$-representation $\pi_\omega$ of $\mathcal{A}$ as bounded linear operators on a Hilbert space $\mathcal{H}_\omega$, together with an element $\Omega_\omega$ of $\mathcal{H}$ such that

$$\omega(A) = (\pi(A)\Omega, \Omega) \quad \text{for all} \quad A \in \mathcal{A} \quad (1)$$

and $\pi_\omega(A)$ is dense in $\mathcal{H}_\omega$. This representation is unique up to unitary equivalence. Its is known as the Gelfand-Naimark-Segal (GNS) representation induced by the state $\omega$.

Let us make the simplifying assumptions that the $C^*$-algebra $\mathcal{A}$ is a von Neuman algebra of operators on a fixed Hilbert space $\mathcal{H}$ and that there is
given a fixed faithful normal state \( \omega \), which will be used as starting point of the construction following later on. Because \( \omega \) is faithful there exists an element \( \Omega \) of \( \mathcal{H} \) such that the trivial representation, defined by \( \pi_\omega(A) = A \) for all \( A \in \mathcal{A} \), is the GNS representation induced by \( \omega \). This simplifying assumption is similar to the assumption made in [1] that the probability distributions of the statistical manifold are absolutely continuous w.r.t. a given probability distribution. This condition can be relaxed — see Section 3.3 of [6].

3 The deformed logarithmic and exponential functions

The specific deformed logarithmic and exponential functions introduced below have been first used by Newton [8], without considering them as deformed functions. The approach was then picked up by Montrucchio and Pistone [1].

Fix the function \( \phi(u) = u/(1 + u) \). It is strictly positive and increasing on the interval \((0, +\infty)\). It determines a deformed logarithm \([11]\) by

\[
\log_\phi(v) = \int_1^v \frac{1}{\phi(u)} \, du = v - 1 + \log v.
\]

It is concave strictly increasing on \((0, +\infty)\). The inverse function is denoted \( \exp_\phi \). It is defined on all of the real line. It is convex strictly increasing. Special values are \( \log_\phi(1) = 0 \) and \( \exp_\phi(0) = 1 \).

Useful properties are

\[
\begin{align*}
\log_\phi(uv) & = \log_\phi(u) + \log_\phi(v) + (u - 1)(v - 1), \\
\exp_\phi(u) & = 1 + u - \log \exp_\phi(u), \\
d\exp_\phi(u)/du & = \phi(\exp_\phi(u)) = \frac{\exp_\phi(u)}{1 + \exp_\phi(u)}, \\
|\exp_\phi(u) - \exp_\phi(v)| & \leq |u - v|.
\end{align*}
\]

The inequality follows from

\[
\exp_\phi(u) - \exp_\phi(v) \leq u - v \quad \text{if } u \geq v.
\]

The latter also implies that

\[
\exp_\phi(u) \begin{cases} 
1 + u & \text{if } u \geq 0, \\
1 + u & \text{if } u \leq 0.
\end{cases}
\]
From \( \log(1 + u) \leq u \) follows

\[
\log_\phi(1 + u) = u + \log(1 + u) \leq 2u.
\]

This implies

\[
\exp_\phi(u) \geq 1 + \frac{u}{2} \quad \text{for all } u.
\] (7)

The following result is needed later on.

**Proposition 3.1** One has

1) For all \( u > 0 \) is

\[
[\log(1 + u)]^2 \leq [\log 2]^2 + (u - 1)^2 + [u - 1 + \log(1 + u)] \log u.
\] (8)

2) There exists a constant \( C \) such that

\[
[\log \phi(u)]^2 \leq C + (\log_\phi(u))^2 \quad \text{for all } u > 0.
\] (9)

**Proof**

1) The derivative of r.h.s. - l.h.s. equals

\[
2(u - 1) + \frac{1}{u} \left[ u - 1 + \log(1 + u) \right] + \left[ 1 + \frac{1}{1 + u} \right] \log u - \frac{2}{1 + u} \log(1 + u).
\]

It is a strictly increasing function which vanishes at \( u = 1 \). Hence, r.h.s. - l.h.s. of (8) is minimal at \( u = 1 \). One finally verifies that r.h.s. = l.h.s. holds at \( u = 1 \).

2) From 1) follows that

\[
[\log \phi(u)]^2 = [\log(1 + u)]^2 + [\log u]^2 - 2 [\log(1 + u)] [\log u] \leq
\]

\[
[\log 2]^2 + (u - 1)^2 + [u - 1 + \log(1 + u)] \log u + [\log u]^2 - 2 [\log(1 + u)] \log u
\]

\[
= [\log 2]^2 + (\log_\phi(u))^2 + R(u)
\]

with

\[
R(u) = -[u - 1 + \log(1 + u)] \log u.
\]
The function $R(u)$ is continuous and tends to $-\infty$ both when $u$ tends to 0 and when $u$ tends to $+\infty$. It is positive on the interval $[u_0, 1]$, with $u_0$ the solution of $\log(1 + u_0) = 1 - u_0$. In this interval it has a unique maximum, outside it is negative. Hence the function $R(u)$ is bounded above. This implies the existence of a constant $C$ such that (9) holds.

□

A function $f(u)$ is operator-monotone if $A \leq B$ implies $f(A) \leq f(B)$ for any pair $A, B$ of self-adjoint operators. The logarithm is operator-monotone. The deformed logarithm $\log_\phi(v) = v - 1 + \log v$ is the sum of operator-monotone functions. Hence it is also operator-monotone. Finally, an operator-monotone function is also automatically operator-concave. This implies that

$$\log_\phi(\lambda A + (1 - \lambda)B) \geq \lambda \log_\phi(A) + (1 - \lambda) \log_\phi(B)$$

for any pair of positive bounded operators $A$ and $B$ and for any $\lambda \in [0, 1]$. See Section 11.6 of [7]. In the Appendix an example is given of a function which is increasing and concave but not operator-monotone. This particular function gives useful results in the commutative case, results which do not follow in the present non-commutative context.

4 Construction

Now follows the construction of a special class of self-adjoint operators $X$ all satisfying $X > 0$ and $||X^{1/2}\Omega|| = 1$.

Proposition 4.1 Let be given a self-adjoint operator $H$ on the Hilbert space $\mathcal{H}$. Then one has

1) Any $\Psi$ in the domain of $H$ also belongs to the domain of $[\exp_\phi(H - \beta)]^{1/2}$ for any real number $\beta$;

2) The map $\beta \to \exp_\phi(H - \beta)$ is strongly continuous.

3) The map $\beta \to ||[\exp_\phi(H - \beta)]^{1/2} \Psi||$ is strictly decreasing for any $\Psi$ in the domain of $H$.

If in addition $\Omega$ belongs to the domain of $H$ then one has also

4) The map $\beta \to ||[\exp_\phi(H - \beta)]^{1/2} \Omega||$ is one-to one from $(-\infty, +\infty)$ to $(0, +\infty)$.

5) There exists a unique non-negative number, denoted $\alpha(H)$, for which $||[\exp_\phi(H - \alpha(H))]^{1/2} \Omega|| = 1$. 

5
6) For any real number $c$ is $\alpha(H + c) = \alpha(H) + c$.

Proof

1) From (5) follows that

$$\exp_\phi(\lambda - \beta) \leq \exp_\phi(-\beta) + |\lambda|.$$  

This implies

$$\exp_\phi(H - \beta) \leq \exp_\phi(-\beta) + |H|$$

and

$$\||[\exp_\phi(H - \beta)]^{1/2}\Psi||^2 \leq \exp_\phi(-\beta)||\Psi||^2 + (|H|\Psi, \Psi) < +\infty.$$  

This shows that $\Psi$ belongs to the domain of $[\exp_\phi(H - \beta)]^{1/2}$.

2) Let

$$H = \int \lambda dE_\lambda$$

denote the spectral decomposition of the operator $H$. From (5) follows

$$\exp_\phi(H - \beta_1) - \exp_\phi(H - \beta_2) = \int [\exp_\phi(\lambda - \beta_1) - \exp_\phi(\lambda - \beta_2)] dE_\lambda$$

$$\leq \int |\beta_1 - \beta_2| dE_\lambda$$

$$= |\beta_1 - \beta_2|.$$  

This implies strong continuity of the map

$$\beta \to \exp_\phi(H - \beta).$$

3) Assume $\gamma > \beta$. From the convexity of $\exp_\phi$ follows that

$$\exp_\phi(\lambda - \beta) \geq \exp_\phi(\lambda - \gamma) + (\gamma - \beta)\phi(\exp_\phi(\lambda - \gamma)).$$

This implies

$$\||[\exp_\phi(H - \beta)]^{1/2}\Psi||^2$$
\[\begin{align*}
&= \int \exp_\phi(\lambda - \beta) \, d(E_\lambda \Psi, \Psi) \\
\geq & \int \left[ \exp_\phi(\lambda - \gamma) + (\gamma - \beta) \phi(\exp_\phi(\lambda - \gamma)) \right] \, d(E_\lambda \Psi, \Psi) \\
= & \left\| \left[ \exp_\phi(H - \gamma) \right]^{1/2} \Psi \right\|^2 + (\gamma - \beta) \left\| \phi(\exp_\phi(H - \gamma)) \right\|^{1/2} \Psi \right\|^2.
\end{align*}\]

Because \( \phi \) is strictly positive zero cannot be an eigenvalue of \( \phi(\exp_\phi(H - \gamma)) \). Hence, \( \left\| \phi(\exp_\phi(H - \gamma)) \right\|^{1/2} \Psi \neq 0 \) and \( \gamma > \beta \) implies the strict inequality
\[\left\| \exp_\phi(H - \beta) \right\|^{1/2} \Psi \right\|^2 > \left\| \exp_\phi(H - \gamma) \right\|^{1/2} \Psi \right\|^2.\]

4) Introduce the notation \( X_\beta(H) \equiv \exp_\phi(H - \beta) \). Because the phi-logarithmic function is concave one has
\[-\beta = ((H - \beta)\Omega, \Omega) = (\log_\phi(X_\beta(H)\Omega, \Omega)) \leq \log_\phi\left( \left\| (X_\beta(H))^{1/2} \Omega \right\|^2.\]

Hence, if \( \beta \) tends to \( -\infty \) then \( \log_\phi\left( \left\| (X_\beta(H))^{1/2} \Omega \right\| \) tends to \( +\infty \). This implies that \( \left\| (X_\beta(H))^{1/2} \Omega \right\| \) tends to \( +\infty \).

On the other hand, if \( \beta \) tends to \( +\infty \) then \( \left\| (X_\beta(H))^{1/2} \Omega \right\| \) tends to zero. This follows from the following argument.

Fix \( \epsilon > 0 \). Because \( \Omega \) is in the domain of \( H \) there exists \( \lambda_\epsilon \) such that
\[\left\| \int_{\lambda_\epsilon}^{+\infty} dE_\lambda \right\| < \epsilon.\]

Without restriction take \( \lambda_\epsilon \geq 1 \). Next choose \( \beta_\epsilon \) large enough so that
\[\exp_\phi(\lambda_\epsilon - \beta_\epsilon) < \epsilon.\]

Then for all \( \beta > \beta_\epsilon \) is
\[\left\| (X_\beta(H))^{1/2} \Omega \right\|^2 = \int_{-\infty}^{+\infty} \int_{\lambda_\epsilon}^{+\infty} \exp_\phi(\lambda - \beta) \, d(E_\lambda \Omega, \Omega) + \int_{\lambda_\epsilon}^{+\infty} \exp_\phi(\lambda - \beta) \, d(E_\lambda \Omega, \Omega) \leq \epsilon \int_{-\infty}^{+\infty} d(E_\lambda \Omega, \Omega) + \int_{\lambda_\epsilon}^{+\infty} \left[ \exp_\phi(-\beta) + \lambda \right] \, d(E_\lambda \Omega, \Omega).\]
\[
\begin{align*}
\leq \epsilon + 2 \int_{\lambda_0}^{+\infty} \lambda \, d(E_\lambda \Omega, \Omega) \\
\leq 3\epsilon.
\end{align*}
\]

This finishes the proof that \( ||[X_\beta(H)]^{1/2}\Omega|| \) tends to 0 as \( \beta \) tends to \(+\infty\).

Because \( \beta \to ||[X_\beta(H)]^{1/2}\Omega|| \) is strictly decreasing and continuous one concludes that the map is one-to one.

5) The existence of a unique real number \( \alpha(H) \) is an immediate consequence of item 4). Convexity of the deformed exponential implies

\[
||[\exp(\beta)H]^{1/2}\Omega||^2 = \int \exp(\lambda) \, d(E_\lambda \Omega, \Omega) \\
\geq \exp \left( \int d(E_\lambda \Omega, \Omega) \lambda \right) \\
= \exp \left( \langle H\Omega, \Omega \rangle \right) \\
= \exp(0) \\
= 1.
\]

Because the map \( \beta \to ||[\exp(\beta)H]^{1/2}\Omega|| \) is strictly decreasing and has a value \( \geq 1 \) at \( \beta = 0 \) one concludes that it takes the value 1 at \( \alpha(H) \geq 0 \).

6) Note that \( \alpha(H + c) \leq \alpha(H) + c \) if and only if

\[
|| \left[ \exp(\beta) \left( (H + c) - (\alpha(H) + c) \right) \right]^{1/2} \Omega || \leq 1. \tag{11}
\]

In the latter expression the real numbers \( c \) cancel and equality follows. One concludes that \( \alpha(H + c) \leq \alpha(H) + c \) holds always. Replace \( H \) by \( H - c \) to conclude that also \( \alpha(H) \leq \alpha(H - c) + c \). Change the sign of \( c \) to find \( \alpha(H + c) \geq \alpha(H) + c \). The two results together yield \( \alpha(H + c) = \alpha(H) + c \).

In the commutative case the function \( K(u) \), which is the analogue of the function \( \alpha(H) \), is shown to be convex. See Proposition 1 of [1]. No such result is expected here because unbounded self-adjoint operators do not form an affine space. In addition, the convexity proof is based on the convexity of the deformed exponential function. However, a convex increasing function cannot be operator-monotone because operator-monotone functions are automatically concave. It is therefore not immediately clear how to prove convexity properties for these functions of operators.
5 Properties of the normalization function

Proposition 5.1 Let $H$ be a self-adjoint operator such that $\Omega$ belongs to its domain. One has

1) The function $t \mapsto \alpha(tH)$ is convex.

2) One has

$$\alpha((1 + \epsilon)H) \geq -\log \phi \left( \sqrt{1 + a_t^2} - a_t \right) \quad \text{with} \quad a_t = \frac{t}{2}(H\Omega, \Omega). \quad (12)$$

3) For real $t$

$$t(H\Omega, \Omega) \geq -2 \quad \text{implies} \quad \alpha(tH) \geq t(H\Omega, \Omega). \quad (13)$$

4) The function $t \mapsto \alpha(tH)$ is differentiable at $t = 0$, with vanishing derivative.

Proof

1) Let

$$H = \int \lambda dE_\lambda$$

be the spectral decomposition of $H$. Let

$$\beta = \mu \alpha(t_1H) + (1 - \mu)\alpha(t_2H).$$

One has

$$||[\exp_\phi((\mu t_1 + (1 - \mu)t_2)H - \beta)]^{1/2} \Omega||^2$$

$$= \int d(E_\lambda \Omega, \Omega) \exp_\phi((\mu t_1 + (1 - \mu)t_2)\lambda - \beta)$$

$$\leq \mu \int d(E_\lambda \Omega, \Omega) \exp_\phi(t_1\lambda - \alpha(t_1H))$$

$$+ (1 - \mu) \int d(E_\lambda \Omega, \Omega) \exp_\phi(t_2\lambda - \alpha(t_2H))$$

$$= \mu||[\exp_\phi(t_1H - \alpha(t_1H))]^{1/2} \Omega||^2 + (1 - \mu)||[\exp_\phi(t_2H - \alpha(t_2H))]^{1/2} \Omega||^2$$

$$= 1.$$

This implies that $\beta \geq \alpha((\mu t_1 + (1 - \mu)t_2)H)$. One concludes that $\mu \alpha(t_1H) + (1 - \mu)\alpha(t_2H) \geq \alpha((\mu t_1 + (1 - \mu)t_2)H)$, i.e. $t \mapsto \alpha(tH)$ is convex.
2) Because the deformed exponential is convex one has

\[
||\exp_\phi((1 + \epsilon)H - \beta)|^{1/2}\Omega||^2 \\
= \int d(E\lambda\Omega, \Omega) \exp_\phi((1 + \epsilon)\lambda - \beta) \\
\geq \int d(E\lambda\Omega, \Omega) \{\exp_\phi(\lambda - \beta) + \epsilon\lambda\phi(\exp_\phi(\lambda - \beta))\} \\
= ||\exp_\phi(H - \beta)|^{1/2}\Omega||^2 \\
+ \epsilon (H[\exp_\phi(H - \beta)]^{1/2}\Omega, [\exp_\phi(H - \beta)]^{1/2}\Omega).
\]

Hence, if the r.h.s. of the latter expression is larger than 1 then \(\beta \leq \alpha((1 + \epsilon)H)\) follows. This condition is satisfied when

\[
\beta \leq -\log_\phi \left(\sqrt{1 + a_t^2} - a_t\right).
\]

This implies (12).

3) The second derivative of the function

\[
f(a) = -\log_\phi \left(\sqrt{1 + a^2} - a\right)
\]

equals

\[
f''(a) = -\frac{1 + a}{(1 + a^2)^{3/2}}.
\]

Hence, the function \(f(a)\) is concave when \(1 + a > 0\). This implies that

\[
f(a) \leq 2a \quad \text{when} \quad a \geq -1.
\]

Because \(t \mapsto \alpha(tH)\) is convex, \(t \mapsto f(a_t)\) is concave on \([-1, +\infty)\), \(\alpha(tH) \geq f(a_t)\) and \(\alpha(0) = f(0) = 0\) one concludes that both curves have a common tangent \(t \mapsto 2a_t = t(H\Omega, \Omega)\). This gives (13).

4) Without restriction assume that \((H\Omega, \Omega) = 0\). Then the previous result implies that \(\alpha(tH) \geq 0\) for all \(t\). Because \(\alpha(0) = 0\) and \(t \mapsto \alpha(tH)\) is convex one concludes that \(\alpha(tH)\) is continuous with a minimum at \(t = 0\).

Now calculate

\[
||[\exp_\phi(tH - \gamma t^2)]^{1/2}\Omega||^2 \\
= \int d(E\lambda\Omega, \Omega) \exp_\phi(t\lambda - \gamma t^2)
\]

10
\[
\int d(E\lambda\Omega, \Omega) \left[ 1 + t\lambda - \gamma t^2 \right] \\
- \int d(E\lambda\Omega, \Omega) \log \exp_\phi(t\lambda - \gamma t^2)
\]

\[
= 1 - \gamma t^2 - \int d(E\lambda\Omega, \Omega) \log \exp_\phi(t\lambda - \gamma t^2)
\]

\[
= 1 - \gamma t^2 - \int d(E\lambda\Omega, \Omega) \mathbb{1}_{\{\lambda - \gamma t^2 \leq -1\}} \log \exp_\phi(t\lambda - \gamma t^2)
\]

\[
- \int d(E\lambda\Omega, \Omega) \mathbb{1}_{\{\lambda - \gamma t^2 > -1\}} \log \exp_\phi(t\lambda - \gamma t^2).
\]

In the last of these terms use (7) and \( \log(1 + u) \geq u - u^2 \) when \( u > -1/2 \) to obtain

\[
||[\exp_\phi(tH - \gamma t^2)]^{1/2}\Omega||^2 \\
\leq 1 - \gamma t^2 - \int d(E\lambda\Omega, \Omega) \mathbb{1}_{\{\lambda - \gamma t^2 \leq -1\}} \log \exp_\phi(t\lambda - \gamma t^2)
\]

\[
- \int d(E\lambda\Omega, \Omega) \mathbb{1}_{\{\lambda - \gamma t^2 > -1\}} \left[ \frac{1}{2}(t\lambda - \gamma t^2) - \frac{1}{2}(t\lambda - \gamma t^2)^2 \right]
\]

\[
= 1 - \frac{1}{2} \gamma t^2 + \frac{t^2}{2} ||[H - \gamma t]\Omega||^2 - \int d(E\lambda\Omega, \Omega) f(t^2/2 - \lambda t/2)
\]

with

\[
f(u) = u + 2u^2 + \log \exp_\phi(-2u) \quad \text{if} \quad u \geq \frac{1}{2},
\]

\[
= 0 \quad \text{otherwise}.
\]

Because \( |f(u)| \leq 2u^2 \) for all \( u \) there follows

\[
||[\exp_\phi(tH - \gamma t^2)]^{1/2}\Omega||^2 \\
\leq 1 - \frac{1}{2} \gamma t^2 + \frac{t^2}{2} ||[H - \gamma t]\Omega||^2 + 2 \int d(E\lambda\Omega, \Omega) (t^2/2 - \lambda t/2)^2
\]

\[
= 1 - \frac{1}{2} \gamma t^2 + t^2 ||[H - \gamma t]\Omega||^2.
\]

The condition \( \gamma > 2 ||[H - \gamma t]\Omega||^2 \) has a solution when \( 8t^2 ||H\Omega||^2 < 1 \). If satisfied then the r.h.s. of the above expression is less than 1. This implies that \( \alpha(tH) \leq \gamma t^2 \) and hence that the function \( t \mapsto \alpha(tH) \) is differentiable at \( t = 0 \). \( \square \)

### 6 Definition of a family of state vectors

Let be given the normal faithful state \( \omega \), as before, with cyclic and separating vector \( \Omega \). The subset \( \mathcal{P}(\Omega) \) of \( \mathcal{H} \) is defined to be the set of elements \( \Psi \) of
$\mathcal{H}$ for which an operator $Y$ exists such that $\Psi = Y\Omega$ and which satisfies the conditions

- $\Omega$ is in the domain of $Y$ and $||Y\Omega|| = 1$;
- $Y$ is a closed operator affiliated with $\mathcal{A}$;
- $Y^*Y$ is strictly positive and $\mathcal{A}\Omega$ is a subset of the domain of $\log_\phi(Y^*Y)$;
- The range of $Y$ is dense in $\mathcal{H}$.

Note that $\Psi$ in the domain of $\log_\phi(Y^*Y)$ implies that $\Psi$ is in the domain of $|Y|$. This follows from Proposition 4.1 with $X = Y^*Y$, $H = \log_\phi(Y^*Y)$ and $\beta = 0$. As a consequence, $\mathcal{A}\Omega$ is a subset of the domain of $Y$.

Every $\Psi \in \mathcal{P}(\Omega)$ determines a normal state $\sigma_\Psi$ of $\mathcal{A}$ by the relation

$$\sigma_\Psi(A) = (A\Psi, \Psi), \quad A \in \mathcal{A}.$$ 

One cannot expect that the map $\Psi \mapsto \sigma_\Psi$ is one-to-one.

The following characterization of the set $\mathcal{P}(\Omega)$ holds.

**Theorem 6.1** Let $\Psi \in \mathcal{H}$. Are equivalent

i) $\Psi$ belongs to $\mathcal{P}(\Omega)$;

ii) There exists a unitary operator $U$ in $\mathcal{A}$ and a self-adjoint operator $H$ affiliated with $\mathcal{A}$, with domain including $\mathcal{A}\Omega$, such that $\Psi = Y\Omega$ with $Y = UX^{1/2}$ and $X = \exp_\phi(H - \alpha(H))$.

If wanted, $H$ can be chosen so that $(H\Omega, \Omega) = 0$.

**Proof**

i) $\Rightarrow$ ii) Given $\Psi \in \mathcal{P}(\Omega)$ write $\Psi = Y\Omega$ and let $Y = U|Y|$ be the polar decomposition of $Y$. Because $Y^*Y$ is strictly positive the isometry $U$ is not partial. Because the range of $Y$ is dense in $\mathcal{H}$ the isometry is unitary. Let

$$H = \log_\phi(Y^*Y) - (\log_\phi(Y^*Y)\Omega, \Omega).$$

It is a well-defined self-adjoint operator affiliated with $\mathcal{A}$. Clearly is $(H\Omega, \Omega) = 0$ and

$$\exp_\phi(H - \alpha(H)) = Y^*Y = |Y|^2,$$

with

$$\alpha(H) = -(\log_\phi(Y^*Y)\Omega, \Omega).$$

Because $\Psi$ belongs to $\mathcal{P}(\Omega)$ the domain of $\log_\phi(Y^*Y)$ includes $\mathcal{A}\Omega$. This implies that $\mathcal{A}\Omega$ is a subset of the domain of $H$. 

12
ii) ⇒ i) Four statements must be proved.

a) \( \Omega \) belongs to the domain of \( Y \) and \( \| Y \Omega \| = 1 \). The domain of the self-adjoint operator \( H \) includes \( A \). Hence Proposition 4.1 proves that \( X = \exp(\phi(H - \alpha(H))) \) satisfies \( \| X^{1/2} \Omega \| = 1 \). This implies \( \| Y \Omega \|^2 = \| UX^{1/2} \Omega \|^2 + \| X^{1/2} \Omega \|^2 = 1 \).

b) \( Y \) is a closed operator affiliated with \( A \). This follows because it is the product of an isometry \( U \) belonging to \( A \) and a self-adjoint operator \( X^{1/2} \) affiliated with \( A \).

c) \( Y^* Y \) is strictly positive and \( \mathcal{A} \Omega \) is a subset of the domain of \( \log \phi(Y^* Y) \). The function \( \exp \phi(u) \) is strictly positive defined on all of the real axis. Hence, the operator \( X = \exp \phi(H - \alpha(H)) \) is strictly positive. From \( Y^* Y = X \) then follows that \( Y^* Y \) is strictly positive. Because

\[
\log \phi(X) = H - \alpha(H)
\]

and \( \mathcal{A} \Omega \) is a subset of the domain of \( H \) it is also a subset of the domain of \( \log \phi(X) \).

d) The range of \( Y \) is dense in \( \mathcal{H} \). Because \( X \) is self-adjoint the closure of its range is the orthogonal complement of its kernel. The latter is trivial because \( X \) is strictly positive. Hence the range of \( X \) is dense in \( \mathcal{H} \). Because \( U \) is unitary this implies that also the range of \( Y = UX^{1/2} \) is dense.

One concludes that \( \Psi \) belongs to \( \mathcal{P}(\Omega) \).

This ends the proof of ii) ⇒ i).

\[ \Box \]

7 The manifold

The set \( \mathcal{P}(\Omega) \) is considered to be the non-commutative generalization of a deformed exponential family of probability distributions. In the Literature the operator \( X = Y^* Y \) appearing in Theorem 6.1 is called the non-commutative generalization of the Radon-Nikodym derivative of \( \omega_X \) w.r.t. \( \omega \), where \( \omega_X \) is defined by \( \omega_X(A) = (AX^{1/2} \Omega , X^{1/2} \Omega) \). The insertion of an unitary operator \( U \) in the definition of \( \mathcal{P}(\Omega) \) is important for what follows. In the commutative case it drops out because it satisfies \( UU^* = U^* U = I \).

For convenience, let \( \mathcal{A}_1 \) denote the group of unitary operators in \( \mathcal{A} \). Let \( \mathcal{A}_\Omega \) denote the set of self-adjoint operators \( H \) affiliated with \( \mathcal{A} \), with domain including \( \mathcal{A} \Omega \) and satisfying \( (H \Omega , \Omega) = 0 \). The previous theorem shows that any pair \( U, H \), with \( U \) in \( \mathcal{A}_1 \) and \( H \) in \( \mathcal{A}_\Omega \), determines an element \( \Psi_{U,H} \) of \( \mathcal{P}(\Omega) \). It is given by

\[
\Psi_{U,H} = U[\exp(\phi(H - \alpha(H)))]^{1/2} \Omega.
\]
Proposition 7.1 The map $\phi : U, H \mapsto \Psi_{U,H}$ is one-to-one.

Proof
Assume that
\[ U[\exp_{\phi}(H_1 - \alpha(H_1))]^{1/2}\Omega = V[\exp_{\phi}(H_2 - \alpha(H_2))]^{1/2}\Omega. \]
This can be written as
\[ [\exp_{\phi}(H_2 - \alpha(H_2))]^{1/2}\Omega = V^*U[\exp_{\phi}(H_1 - \alpha(H_1))]^{1/2}\Omega. \]
Because the state $\omega$ is faithful this implies that
\[ [\exp_{\phi}(H_2 - \alpha(H_2))]^{1/2} = V^*U[\exp_{\phi}(H_1 - \alpha(H_1))]^{1/2}. \]
From the uniqueness of the polar decomposition follows $V^*U = 1$ and
\[ [\exp_{\phi}(H_2 - \alpha(H_2))]^{1/2} = [\exp_{\phi}(H_1 - \alpha(H_1))]^{1/2}. \]
Take the phi-logarithm to obtain
\[ H_2 - \alpha(H_2) = H_1 - \alpha(H_1). \]
From $(H_i\Omega, \Omega) = 0$, $i = 1, 2$ follows that $\alpha(H_2) = \alpha(H_1)$ and therefore also $H_2 = H_1$.

Because the range of the map $\phi : U, H \mapsto \Psi_{U,H}$ is all of $\mathcal{P}(\Omega)$ the previous result shows that $\mathcal{A}_1 \times \mathcal{A}_\Omega, \phi$ is a chart for the manifold $\mathcal{P}(\Omega)$.

8 Affine spaces

It is shown in the previous section that the pairs $U, H$ with $U$ in $\mathcal{A}_1$ and $H$ in $\mathcal{A}_\Omega$ can be used as coordinates for the manifold $\mathcal{P}(\Omega)$. However, in the general case neither $\mathcal{A}_1$ nor $\mathcal{A}_\Omega$ is an affine space, a property which one expects for the coordinate spaces of exponential families (see Section 2.8.1 of [6]).

Any unitary operator $U$ in $\mathcal{A}_1$ can be written in exponential form $U = \exp(iA)$ with $A$ a self-adjoint operator affiliated with $\mathcal{A}$. This suggests that an affine parametrization of $\mathcal{A}_1$ is feasible. This possibility is not further explored here. Let us rather concentrate on $\mathcal{A}_\Omega$. The problem here is that
the sum of two self-adjoint operators $H$ and $K$ in $\mathcal{A}_\Omega$ need not to be self-adjoint. However, the convex combination $(1 - \lambda)H + \lambda K$ is essentially self-adjoint if $\lambda$ is close to 0. This is worked out below.

The \textit{gap} distance $\delta(H, K)$ between two self-adjoint operators $H$ and $K$ is the maximal distance between their graphs when intersected with the unit ball of $\mathcal{H} \times \mathcal{H}$. See Section IV.2.1 of [13].

**Proposition 8.1** Fix an operator $H$ in $\mathcal{A}_\Omega$. There exists $\epsilon > 0$ such that for any operator $K$ in $\mathcal{A}_\Omega$, $K \neq H$ and for any $\lambda$ in $[0, 1]$ satisfying

$$\lambda \leq \frac{\epsilon}{\delta(H, K)}$$

the closure of the operator $(1 - \lambda)H + \lambda K$ exists and belongs to $\mathcal{A}_\Omega$.

**Proof**

Theorem 4.1 of Chapter V of [13] asserts the existence of $\epsilon > 0$ such that any closed symmetric operator within gap distance $\epsilon$ of $H$ is self-adjoint. Let $L = (1 - \lambda)H + \lambda K$. It is densely defined because both the domains of $H$ and of $K$ contain $\mathcal{A}_\Omega$ as a subset. It is straightforward to verify that $L$ is a symmetric operator and that $\delta(H, L) = \lambda \delta(H, K)$. The closure of $L$ is symmetric again and its gap distance to $H$ is less than epsilon if (16) is satisfied. Hence it is self-adjoint. In addition it is affiliated with $\mathcal{A}$ and satisfies $\langle L\Omega, \Omega \rangle = 0$. Therefore it belongs to $\mathcal{A}_\Omega$.

The previous proposition supports the statement that $\mathcal{A}_\Omega$ looks locally like an affine space.

9 The tangent plane at the reference point

Given an unitary operator $U \in \mathcal{A}_1$ define a subset $\mathcal{P}_U(\Omega)$ of $\mathcal{P}(\Omega)$ by

$$\mathcal{P}_U(\Omega) = \{ \Psi_{U,H} : H \in \mathcal{A}_\Omega \},$$

and a map $\phi_U : \mathcal{P}_U(\Omega) \mapsto \mathcal{H}$ by

$$\phi_U(\Psi_{U,H}) = H.$$

The plane tangent to $\mathcal{P}_U(\Omega)$ at the origin $\Omega$ is now constructed.
Proposition 9.1 For all $U$ in $A_1$ and for all $H$ in $A_\Omega$ is
\[ \frac{d}{dt} \bigg|_{t=0} \Psi_{U,tH} = \frac{1}{4} U H \Omega. \]

Proof

The proof uses Proposition 5.1. The calculation is straightforward because the operators $tH$ mutually commute.

From this result follows that the tangent plane $T_\Omega P_U(\Omega)$ consists of all elements of $U A_\Omega \Omega$.

Proposition 9.2 $T_\Omega P_U(\Omega)$ is an affine subspace of $\mathcal{H}$.

Proof

From the definition follows immediately that if $H$ belongs to $A_\Omega$ then also $tH$ belongs to it for all real $t$. By the argument used in the proof of Proposition 8.1 there exists $\epsilon > 0$ such that, if the gap distance $\delta(\mathbb{I}, A)$ is less than $\epsilon$ and $A$ is symmetric and closed then $A$ is self-adjoint. Take any $H, K$ in $A_\Omega$ and $\lambda$ in $[0, 1]$. Then there exist $t > 0$ such that $\delta(\mathbb{I}, t(1-\lambda)H) + t\lambda K \leq \epsilon$. Then the closure $L$ of $t(1-\lambda)H + t\lambda K$ belongs to $A_\Omega$ and so does $L/t$. This shows that $U[(1-\lambda)H + \lambda K] \Omega$ belongs to $T_\Omega P_U(\Omega)$. One concludes that $T_\Omega P_U(\Omega)$ is convex. In combination with the first property, namely $\Psi \in T_\Omega P_U(\Omega)$ implies $t\Psi \in T_\Omega P_U(\Omega)$ for all real $t$, this implies that it is a plane, this is an affine subspace of $\mathcal{H}$.

10 Parallel transport

Consider now an arbitrary $\Psi$ in $\mathcal{P}(\Omega)$. From the definition follows that $\Psi$ is a cyclic and separating vector for $A$. Hence one can replace the reference point $\Omega$ by this new point $\Psi$ and reuse the above results but now for the family $\mathcal{P}(\Psi)$. The obvious question is then what is the relation between $\mathcal{P}(\Omega)$ and $\mathcal{P}(\Psi)$.

Introduce a subset $Q(\Omega)$ of $\mathcal{P}(\Omega)$ defined by
\[ Q(\Omega) = \{ \Psi = Y \Omega \in \mathcal{P}(\Omega) : \text{ both } Y \text{ and } Y^{-1} \text{ belong to } A \}. \]

Let
\[ Q_U(\Omega) = Q(\Omega) \cap \mathcal{P}_U(\Omega). \]
Proposition 10.1 For any $\Psi \in \mathcal{P}(\Omega)$ is $Q(\Psi) \subset \mathcal{P}(\Omega)$ and $Q_U(\Psi) \subset \mathcal{P}_V(\Omega)$ with $V = U|Z|W|Y||[Y^*Z^*ZW|^2Y||^{1/2}$.

Proof
Let $\Xi = Z\Psi$ be an arbitrary element of $Q(\Psi)$. Then $\Xi = ZY\Omega$ holds. For $\Xi$ to belong to $\mathcal{P}(\Omega)$ one has to show that

- $\Omega$ is in the domain of $ZY$ and $||ZY\Omega|| = 1$;
  This is obvious.
- $ZY$ is a closed operator affiliated with $\mathcal{A}$;
  $Y$ is affiliated with $\mathcal{A}$ and $Z$ belongs to $\mathcal{A}$. Hence $ZY$ is affiliated with $\mathcal{A}$.
  Assume $\Phi_n$ converge to zero and $ZY\Phi_n$ converge to $\Phi$. Because $Z^{-1}$ is continuous $Y\Phi_n$ converge to $Z^{-1}\Phi$. Because $Y$ is closed this implies that $Z^{-1}\Phi = 0$ and hence also $\Phi = 0$. One concludes that $ZY$ is a closed operator.
- $(ZY)^*ZY$ is strictly positive and $\mathcal{A}\Omega$ is a subset of the domain of $\log_\phi((ZY)^*ZY)$;
  $(ZY)^*ZY = Y^*Z^*ZY \geq 0$ is straightforward. Assume $Y^*Z^*ZY\Phi = 0$. This implies $ZY\Phi = 0$. Because $Z$ and $Y$ are invertible $\Phi = 0$ follows. Because $\log_\phi$ is operator-monotone $Y^*Z^*ZY \leq ||Z||^2Y^*Y$ implies that $\log_\phi((ZY)^*ZY) \leq \log_\phi(||Z||^2Y^*Y) = (||Z|| - 1)Y^*Y + \log_\phi(Y^*Y)$.
  Because $\mathcal{A}\Omega$ is a subset of the domain of both $Y^*Y$ and $\log_\phi(Y^*Y)$ it is also a subset of the domain of $\log_\phi((ZY)^*ZY)$.
- The range of $ZY$ is dense in $\mathcal{H}$.
  Assume $(ZY\Xi, \Phi) = 0$ for all $\Xi$ in the domain of $ZY$, which equals the domain of $Y$. This implies $(Y\Xi, Z^*\Phi) = 0$ for all $\Xi$ in the domain of $Y$. Because the range of $Y$ is dense in $\mathcal{H}$ there follows $Z^*\Phi = 0$ and hence $\Phi = 0$ because $Z$ and $Z^*$ are invertible. One concludes that the range of $ZY$ is dense in $\mathcal{H}$.
This shows that $Q(\Psi) \subset P(\Omega)$.

Assume now that $\Xi \in Q_U(\Psi)$. This means $Z = U|Z|$. Let $Y = W|Y|$. Then

$$ZY = U|Z|W|Y| = VX^{1/2} \quad \text{with} \quad X = |Y|W^*Z\ast ZW|Y|$$

with $V = U|Z|W|Y|^{-1/2}$. This implies $\Psi \in P_V(\Omega)$. □

The following characterization of the set $Q(\Omega)$ holds.

**Theorem 10.2** Let $\Psi \in \mathcal{H}$. Are equivalent

i) $\Psi$ belongs to $Q(\Omega)$;

ii) There exists a unitary operator $U$ in $\mathcal{A}$ and a self-adjoint operator $H$ in $\mathcal{A}$ such that $\Psi = Y\Omega$ with $Y = UX^{1/2}$ and $X = \exp_{\phi}(H - \alpha(H))$.

If wanted, $H$ can be chosen so that $(H\Omega, \Omega) = 0$.

**Proof**

i) $\Rightarrow$ ii) Given $\Psi \in Q(\Omega)$ write $\Psi = Y\Omega$ and let $Y = U|Y|$ be the polar decomposition of $Y$. By assumption both $Y$ and $Y^{-1}$ are in $\mathcal{A}$. From Theorem 6.1 follows that there exists a self-adjoint operator $H$ affiliated with $\mathcal{A}$, with domain including $\mathcal{A}\Omega$, such that $X = \exp_{\phi}(H - \alpha(H))$ with $X = Y^*Y$. This implies

$$H - \alpha(H) = \log_{\phi}(Y^*Y).$$

Because

$$\frac{1}{||Y^{-1}||^2} \leq Y^*Y \leq ||Y||^2$$

and $\log_{\phi}$ is increasing there follows

$$\log_{\phi}\left(\frac{1}{||Y^{-1}||^2}\right) \leq H - \alpha(H) \leq \log_{\phi}(||Y||^2). \quad (17)$$

This shows that $H$ is bounded and hence belongs to $\mathcal{A}$. 

18
From Theorem 6.1 follows that $\Psi$ belongs to $\mathcal{P}(\Omega)$. From 17 follows that $\log \phi(Y^*Y)$ is bounded. Since the function $\log \phi$ is bijective from $(0, +\infty)$ onto $(-\infty, +\infty)$ there exists constants $0 < \epsilon \leq \delta$ such that $\epsilon \leq Y^*Y \leq \delta$. This implies that both $Y$ and its inverse $Y^{-1}$ are bounded operators. One concludes that $\Psi$ belongs to $\mathcal{Q}(\Omega)$.

□

Given $\Psi$ in $\mathcal{P}(\Omega)$, vectors tangent to $\mathcal{Q}(\Psi)$ can be transported into vectors tangent to $\mathcal{P}(\Omega)$ in a simple manner. This is shown in the following proposition.

**Proposition 10.3** Let $\Psi \in \mathcal{P}(\Omega)$ and $\Phi = \Psi_{U,H} \in \mathcal{Q}(\Psi)$. Then $\frac{1}{4}UK\Omega$ with $K = H - (H\Omega, \Omega)$ is tangent to $\mathcal{Q}(\Omega)$.

**Proof**

By the previous theorem is $H$ a bounded operator belonging to $\mathcal{A}_\Psi$. Then $K$ belongs to $\mathcal{A}_\Omega$. The requirement that the domain of $K$ contains $\mathcal{A}\Omega$ is trivially satisfied. From Proposition 9.11 follows

$$\frac{d}{dt} \bigg|_{t=0} \Psi_{U,tK} = \frac{1}{4}UK\Omega.$$ 

This means that $\frac{1}{4}UK\Omega$ is tangent to $\mathcal{Q}(\Omega)$.

□

11 
**Escort state vectors**

The definition of escort state vectors relies on the following result.

**Proposition 11.1** Take $\Psi = Y\Omega$ in $\mathcal{P}(\Omega)$ and let $X = Y^*Y$. Then

1) $0 < \phi(X) < 1$; in particular, $\phi(X)$ is a bounded strictly positive operator with norm $\|\phi(X)\| \leq 1$.

2) $0 < (\phi(X)\Omega, \Omega) \leq 1/2$.

**Proof**

1) This follows because $X > 0$ and $0 < \phi(u) < 1$ for all $u > 0$. 

19
2) Use the spectral decomposition of $X$

$$X = \int_0^{+\infty} \lambda \, dE_\lambda,$$

together with the concavity of the function $\phi$ and the normalization $||X^{1/2}\Omega||^2 = 1$ to write

$$\langle \phi(X)\Omega, \Omega \rangle = \int_0^{+\infty} \phi(\lambda) \, d(E_\lambda\Omega, \Omega) \leq \phi \left( \int_0^{+\infty} \lambda \, d(E_\lambda\Omega, \Omega) \right) = \phi \left( ||X^{1/2}\Omega||^2 \right) = \phi(1) = \frac{1}{2}.$$  

Finally $\langle \phi(X)\Omega, \Omega \rangle = 0$ implies $X = 0$ because $\phi(0) = 0$, $\phi$ is strictly increasing and $X > 0$. However, $X = 0$ contradicts $||X^{1/2}\Omega|| = 1$. One concludes that $0 < \langle \phi(X)\Omega, \Omega \rangle \leq 1/2$.

□

Given $\Psi = Y\Omega$ in $\mathcal{P}(\Omega)$ and the polar decomposition $Y = UX^{1/2}$ with $X = Y^*Y$, the escort state vector $\tilde{\Psi}$ is defined by

$$\tilde{\Psi} = 1^{1/2} \phi \left( \frac{1}{\phi(X)\Omega, \Omega} \right)^{1/2} U\phi(X)^{1/2}\Omega.$$  

Note that $\phi(X)$ is a bounded positive operator and $\langle \phi(X)\Omega, \Omega \rangle > 0$. Hence $\tilde{\Psi}$ is well-defined. By construction is $||\tilde{\Psi}|| = 1$. The escort state $\tilde{\sigma}_\Psi$ is defined by $\tilde{\sigma}_\Psi = \sigma_{\tilde{\Psi}}$.

**Proposition 11.2** The map $\Psi \to \tilde{\Psi}$ is injective.

**Proof**

Assume $\tilde{\Psi}_1 = \tilde{\Psi}_2$ with $\Psi_i = Y_i\Omega$, $X_i = Y_i^*Y_i$, $Y_i = U_iX_i^{1/2}$, for $i = 1, 2$. Then

$$\frac{1}{(\phi(X_1)\Omega, \Omega)} U_1\phi(X_1)^{1/2}\Omega = \frac{1}{(\phi(X_2)\Omega, \Omega)} U_2\phi(X_2)^{1/2}\Omega.$$  

With the same arguments as in the proof of Proposition 7.1 this implies $U_1 = U_2$ and

$$\phi(X_1)^{1/2} = \sqrt{\kappa}\phi(X_2)^{1/2}$$  

20
with
\[ \kappa = \frac{(\phi(X_1)\Omega, \Omega)}{(\phi(X_2)\Omega, \Omega)}. \]

Take the square to obtain
\[ \phi(X_1) = \kappa \phi(X_2) \]
and hence
\[ \frac{X_1}{1 + X_1} = \frac{\kappa X_2}{1 + X_2}. \]

Without restriction assume that \( \kappa \leq 1 \); If \( \kappa > 1 \) then swap \( X_1 \) and \( X_2 \).
Let us show that \( \kappa = 1 \) by a reasoning ex absurdo. Assume that the strict inequality \( \kappa < 1 \) holds. Then \( X_1 \) is a bounded operator given by
\[ X_1 = \frac{\kappa X_2}{1 + (1 - \kappa)X_2}. \]

Let
\[ X_2 = \int_0^{+\infty} \lambda dE_\lambda \]
denote the spectral representation of \( Y \). Then one has
\[ X_1 = \int_0^{+\infty} \frac{\kappa \lambda}{1 + (1 - \kappa)\lambda} dE_\lambda \]
and hence
\[ 1 = ||X_1^{1/2}\Omega||^2 = \int_0^{+\infty} \frac{\kappa \lambda}{1 + (1 - \kappa)\lambda} d(E_\lambda \Omega, \Omega). \]

Subtract
\[ 1 = ||X_2^{1/2}\Omega||^2 = \int_0^{+\infty} \lambda d(E_\lambda \Omega, \Omega) \]
to obtain
\[ 0 = \int_0^{+\infty} \left[ \frac{\kappa \lambda}{1 + (1 - \kappa)\lambda} - \lambda \right] d(E_\lambda \Omega, \Omega) = -(1 - \kappa) \int_0^{+\infty} \frac{\lambda(1 + \lambda)}{1 + (1 - \kappa)\lambda} d(E_\lambda \Omega, \Omega). \]
Because
\[
\frac{\lambda(1+\lambda)}{1+(1-\kappa)\lambda} \geq 0
\]
with equality if and only if \(\lambda = 0\) this implies that \(X_2\Omega = 0\). The latter contradicts \(||X_2^{1/2}\Omega||^2 = 1\). One concludes that \(\kappa < 1\) is not possible.

From \(\kappa = 1\) follows that
\[
\frac{X_1}{1+X_1} = \frac{X_2}{1+X_2}.
\]
This implies \(X_1 = X_2\). One concludes that \(\Psi_1 = \Psi_2\).

\(\square\)

Introduce the notation
\[
\tilde{\mathcal{P}}(\Omega) \equiv \{ \tilde{\Psi} : \Psi \in \mathcal{P}(\Omega) \}.
\]
By the previous proposition the map \(\Psi \rightarrow \tilde{\Psi}\) is a bijection of \(\mathcal{P}(\Omega)\) onto \(\tilde{\mathcal{P}}(\Omega)\).

**Proposition 11.3** \(\tilde{\mathcal{P}}(\Omega) \subset \mathcal{P}(\Omega)\).

**Proof**
Take \(\tilde{\Psi}\) in \(\tilde{\mathcal{P}}(\Omega)\). By definition there exists \(\Psi = Y\Omega\) in \(\mathcal{P}(\Omega)\) such that
\[
\tilde{\Psi} = \frac{1}{(\phi(X)\Omega,\Omega)^{1/2}} U\phi(X)^{1/2}\Omega,
\]
with polar decomposition \(Y = UX^{1/2}\) and with \(X = Y^*Y\). Let \(\tilde{X} = \phi(X)/(\phi(X)\Omega,\Omega)\). Positivity of \(\tilde{X}\) follows because \(\phi\) is strictly positive on \((0, +\infty)\), and \(X > 0\). We have to show that \(\tilde{Y}\) exists so that \(\tilde{\Psi} = \tilde{Y}\Omega\) and the following conditions are fulfilled.

- \(\Omega\) is in the domain of \(\tilde{Y}\) and \(||\tilde{Y}\Omega|| = 1\);
- \(\tilde{Y}\) is a closed operator affiliated with \(\mathcal{A}\);
- \(\tilde{Y}^*\tilde{Y}\) is strictly positive and \(\mathcal{A}\Omega\) is a subset of the domain of \(\log_\phi(\tilde{Y}^*\tilde{Y})\);
- The range of \(\tilde{Y}\) is dense in \(\mathcal{H}\).
Let $\tilde{Y} = U\phi(X)^{1/2}/(\phi(X)\Omega, \Omega)^{1/2}$. It is a bounded operator. Hence $\Omega$ is in its domain. The normalization $||\tilde{Y}\Omega|| = 1$ is realized by construction. The operator $\tilde{Y}$ belongs to $\mathcal{A}$ because it is a bounded function of an operator affiliated with $\mathcal{A}$. One has

$$\tilde{Y}^* \tilde{Y} = \frac{\phi(X)}{(\phi(X)\Omega, \Omega)}.$$ 

This is a strictly positive operator because the function $\phi$ is strictly positive on $(0, +\infty)$ and the operator $X$ is strictly positive.

Normalisation follows by construction. It remains to be shown that $\mathcal{A}\Omega$ is a subset of the domain of $\log_{\phi}(X)$ and that the range of $\tilde{Y}$ is dense in $\mathcal{H}$. One has

$$\log_{\phi}(\tilde{X}) = \tilde{X} - \mathbb{I} + \log(\tilde{X})$$

$$= \tilde{X} - \mathbb{I} + \log(\phi(X)) - \log(\phi(X)\Omega, \Omega).$$ 

Because $\tilde{X}$ is a bounded operator it suffices to show that any $\Psi$ in $\mathcal{A}\Omega$ belongs to the domain of $\log(\phi(X))$. Let

$$X = \int \lambda dE_{\lambda}$$

be the spectral decomposition of $X$. Then Proposition 3.1 guarantees the existence of a constant $C$ such that

$$||\log(\phi(X))\Psi||^2 = \int [\log(\phi(\lambda))]^2 d(E_{\lambda}\Psi, \Psi)$$

$$\leq C + \int [\log_{\phi}(\lambda)]^2 d(E_{\lambda}\Psi, \Psi)$$

$$= C + ||\log_{\phi}(X)\Psi||^2$$

$$< +\infty.$$

The latter follows because $\Psi$ belongs to the domain of $\log_{\phi}(X)$.

Finally, the range of $\tilde{Y}$ is dense in $\mathcal{H}$ because the null space of $\tilde{X}$ is trivial.

\section*{12 Tracial states}

In the parametrized commutative case escort probabilities occur naturally when considering the gradient of the probability distribution. See Theorem 2.1 of [11]. Proposition 3 of [1] shows in the parameter-free commutative
context that the function $K(u)$, the analogue of $\alpha(H)$ here, is Fréchet differentiable and the derivatives involve the escort distribution. A generalization of these results to the non-commutative case is not straightforward. The closest to a commutative situation is realized when the reference state $\omega$ is tracial, i.e. when $\omega(AB) = \omega(BA)$ for all $A, B \in A$. This property is often referred to as ‘cyclic permutation under the trace’.

The following result is derived using the series expansion approach found in Section 10.8 of [12].

**Proposition 12.1** Assume that the Hilbert space $H$ is finite-dimensional and assume that $\omega$ is a tracial state of $A$, with corresponding cyclic and separating vector $\Omega$. Fix $H$ in $A_\Omega$ and assume it satisfies $||H - \alpha(H)|| < \pi$. Let $X = \exp_{\phi}(H - \alpha(H))$. Then one has

1) For any $A = A^*$ in $A$ satisfying $\omega(A\phi(X)) = 0$ the first derivative of $\alpha(H + tA)$ at $t = 0$ exists and vanishes.

2) For any $A = A^*$ in $A$ the first derivative of $\alpha(H + tA)$ at $t = 0$ exists and equals $\omega(A\phi(X))/\omega(\phi(X))$.

3) The Fréchet derivative of the function $\alpha(H)$ is given by

$$D\alpha(H)(A) = \frac{\omega(A\phi(X))}{\omega(\phi(X))} = \tilde{\sigma}_{A}$$

with $\Psi = X^{1/2}\Omega$.

**Proof**

1) Let

$$\exp_{\phi}(u) = \sum_{n=0}^{\infty} a_n u^n$$

be the series expansion of the deformed exponential function. The first few coefficients are $a_0 = 1$, $a_1 = 1/2$, $a_2 = 1/16$. The radius of convergence equals $\pi$.

Without restriction assume that $H \neq \alpha(H)$. One has for

$$|t| < \frac{||A||}{||H - \alpha(H)||}$$
that
\[
\omega(\exp \phi (H - \alpha(H) + tA)) = \sum_{n=0} a_n \omega([H - \alpha(H) + tA]^n).
\]

The derivative is given by
\[
\frac{d}{dt} \omega(\exp \phi (H - \alpha(H) + tA)) \bigg|_{t=0} = \sum_{n=1} a_n \sum_{p=0}^{n-1} \omega([H - \alpha(H)]^p A[H - \alpha(H)]^{n-1-p}).
\]

Because \(\omega\) is tracial this equals
\[
= \sum_{n=1} a_n \sum_{p=0}^{n-1} \omega(A[H - \alpha(H)]^{n-1})
= \sum_{n=1} n a_n \omega(A[H - \alpha(H)]^{n-1})
= \omega(A\phi(X)).
\]

The latter vanishes by assumption.

Similarly, the second derivative is given by
\[
\frac{d^2}{dt^2} \omega(\exp \phi (H - \alpha(H) + tA)) \bigg|_{t=0} = \omega(A\phi(X)A) \geq 0.
\]

Hence one has
\[
\omega(\exp \phi (H - \alpha(H) + tA)) = 1 + \frac{1}{2} t^2 \omega(A\phi(X)A) + O(t^3).
\]

Without restriction assume \(A \neq 0\). From the definition of \(\alpha(H + tA)\) then follows that \(\alpha(H + tA) \geq \alpha(H)\) for \(t\) in a neighborhood of \(t = 0\), with equality if and only if \(t = 0\). This suffices to conclude that the first derivative of \(\alpha(H + tA)\) vanishes at \(t = 0\).

2) Let \(\lambda = \omega(A\phi(X))/\omega(\phi(X))\). From 1) follows that
\[
0 = \frac{d}{dt} \alpha(H + t[A - \lambda]) \\
= \frac{d}{dt} \alpha(H + tA) + \frac{d}{dt} \alpha(H - t\lambda))
\]

25
\[
\left. \frac{d}{dt} \right|_{t=0} \alpha(H + tA) + \left. \frac{d}{dt} \right|_{t=0} \alpha(H) - t\lambda
\]

This yields 2).

3) From 2) follows that

\[
\frac{d}{dt} \alpha(H + tA) = \omega(A\phi(X_t))
\]

with \(X_t = \exp(H + tA - \alpha(H + tA))\). Integration gives

\[
|\alpha(H + tA) - \alpha(H) - t\omega(A\phi(X))| = \left| \int_0^t ds \left[ \omega(A\phi(X_s)) - \omega(A\phi(X)) \right] \right|
\]

\[
\leq \int_0^t ds \left| \omega(A\phi(X_s)) - \omega(A\phi(X)) \right|
\]

\[
\leq 2t||A||.
\]

This proves the Fréchet differentiability with derivative given by \(D\alpha(H)(A) = \omega(A\phi(X))\).

□

13 Example

Let \(\sigma_1, \sigma_2, \sigma_3\) denote the Pauli matrices. The GNS representation \(\pi, \mathcal{H}, \Omega\) of the tracial state \(\omega\) in 2 dimensions involves a 4-dimensional Hilbert space \(\mathcal{H}\) spanned by orthonormal vectors \(\Omega_0 = \Omega\) and \(\Omega_\alpha = \hat{\sigma}_\alpha \Omega\), \(\alpha = 1, 2, 3\). The notation \(\hat{A}\) is used instead of \(\pi(A)\). For convenience, let \(\sigma_0\) denote the 2-dimensional unit matrix.

Choose \(H = \hat{\sigma}_z\) and \(A = \hat{\sigma}_y\). Let

\[
U = \begin{pmatrix} \cos \phi & i \sin \nu \\ i \sin \nu & \cos \nu \end{pmatrix}
\]

with \(\nu = \frac{1}{2} \arctan t\). Then

\[
U^*(\sigma_z + t\sigma_y)U = \frac{1}{\cos(2\nu)}\sigma_z.
\]

This is a diagonal matrix. Hence

\[
\exp_{\phi}(\sigma_z + t\sigma_y - \beta) = U \begin{pmatrix} a^+ & 0 \\ 0 & a^- \end{pmatrix} U^*
\]

26
\[= \frac{1}{2}(a_+ + a_-) + \frac{1}{2}(a_+ - a_-) \left[ \cos(2\nu)\sigma_z + \sin(2\nu)\sigma_y \right] \]

with

\[a_\pm = \exp(\pm \frac{1}{\cos(2\nu)} - \beta) = \exp(\pm \sqrt{1 + t^2} - \beta).\]

This implies

\[\exp(\phi(H + tA - \beta) = \frac{1}{2}(a_+ + a_-)\hat{\sigma}_0 + \frac{1}{2}(a_+ - a_-) \left[ \cos(2\nu)\hat{\sigma}_z + \sin(2\nu)\hat{\sigma}_y \right].\]

\[\beta = \alpha(H + tA) \text{ is now determined by the requirement that} \]

\[1 = \langle X_t\Omega, \Omega \rangle = \frac{1}{2}(a_+ + a_-) \]

with \(X_t = \exp(\phi(H + tA - \alpha(H + tA))\). Take the derivative of

\[2 = a_+ + a_-\]

\[= \exp(\sqrt{1 + t^2} - \alpha(H + tA)) + \exp(-\sqrt{1 + t^2} - \alpha(H + tA)).\]

This gives

\[D\alpha(H + tA)(A) = \frac{d}{dt} \alpha(H + tA) = \frac{t}{\sqrt{1 + t^2}} \left( \begin{array}{c} \phi(a_+) - \phi(a_-) \\ \phi(a_+) + \phi(a_-) \\ a_+ - a_- \\ \end{array} \right) \]

\[= \frac{1}{2} \sqrt{1 + t^2} \left( a_+ - a_- \right). \tag{18} \]

In particular, \(D\alpha(H)(A) = 0\), as is expected on the basis of symmetry arguments.

On the other hand is

\[\phi(X_t) = U \left( \begin{array}{c} \phi(a_+) \\ 0 \\ \phi(a_-) \end{array} \right) U^* \]

\[= \frac{1}{2} (\phi(a_+) + \phi(a_-)) + \frac{1}{2} (\phi(a_+) - \phi(a_-)) \left[ \cos(2\nu)\sigma_z + \sin(2\nu)\sigma_y \right].\]

This yields

\[\langle A\phi(X_t)\Omega, \Omega \rangle = \frac{1}{2} (\phi(a_+) - \phi(a_-)) \sin(2\nu)\]

\[= \frac{1}{2} \frac{t}{\sqrt{1 + t^2}} (\phi(a_+) - \phi(a_-)).\]
and
\[(\phi(X_t)\Omega, \Omega) = \frac{1}{2}(\phi(a_+) + \phi(a_-)).\]

Hence,
\[
\frac{(A\phi(X_t)\Omega, \Omega)}{(\phi(X_t)\Omega, \Omega)} = \frac{t}{\sqrt{1+t^2}} \frac{\phi(a_+)}{\phi(a_+)} - \frac{\phi(a_-)}{\phi(a_-)},
\]

which is the same result as (18), in agreement with Proposition 12.1.

### 14 Discussion

Part of the work of Montrucchio and Pistone [1] can be transferred to a non-commutative setting in a rather straightforward manner. The probability distributions are replaced by state vectors in a Hilbert space. These are called probability amplitudes because only their squared modulus has the interpretation of a probability. As a consequence, elements of the deformed exponential family \(P(\omega)\) are characterized not by a single positive function but by a pair of operators \(U, X\) where \(U\) is an unitary operator and \(X\) is self-adjoint and strictly positive. See Section 6. The unitary operator is irrelevant in the case of an abelian algebra of operators. However, the freedom of choosing \(U\) is essential in Proposition 10.1.

The main obstacle in generalizing all of [1] to a non-commutative context is that certain monotone functions and convex functions, appearing in the proofs of [1], are not operator-monotone, respectively operator-convex. See the Appendix below. In addition, technical difficulties arise because the sum of two self-adjoint operators is in general not self-adjoint due to problems with the domain of definition. An attempt to circumvent the problem is found in Section 8.

### Appendix

The following negative results give an indication of the kind of problems that one encounters with functions of operators. In this Appendix the function \(\phi\) is defined by
\[\phi(u) = \frac{u}{\lambda + u},\]

where \(\lambda\) is a fixed positive constant. The deformed logarithm equals
\[\log_\phi(u) = u - 1 + \lambda \log u.\]
**Proposition**  The function $f(u) = u - \exp_\phi(u)$ is *not* operator-monotone.

**Proof**

Introduce the shorthands $x = \exp_\phi(u)$ and $y = \exp_\phi(v)$. One has

$$f'(u) = 1 - \frac{\phi(\exp_\phi(u))}{x} = 1 - \frac{\lambda}{\lambda + x} = \frac{\lambda}{\lambda + x}.$$ 

A necessary condition (Theorem 11.17 of [7]) for $f(u)$ to be operator-monotone is that the following determinant is positive

$$D = \begin{vmatrix} f'(u) & 1 - \frac{x-y}{u-v} \\ 1 - \frac{x-y}{u-v} & f'(v) \end{vmatrix} = \frac{\lambda}{\lambda + x} \frac{\lambda}{\lambda + y} - \left[ 1 - \frac{x-y}{u-v} \right]^2.$$

Consider the case $u > v$. This implies $x > y$. Introduce $\epsilon > 0$ defined by $x = (1 + \epsilon)y$. Then

$$u = \log_\phi(x) = x - 1 + \lambda \log x = y - 1 + \lambda \log y + \epsilon y + \lambda \log(1 + \epsilon) = v + \epsilon y + \lambda \log(1 + \epsilon).$$

One obtains

$$D = \frac{\lambda}{\lambda + y + \epsilon y} \frac{\lambda}{\lambda + y} - \left[ \frac{\lambda \log(1 + \epsilon)}{\epsilon y + \lambda \log(1 + \epsilon)} \right]^2.$$ 

The condition that $D > 0$ becomes

$$(\lambda + y)(\lambda + y + \epsilon y) \left[ \frac{\log(1 + \epsilon)}{\epsilon y + \lambda \log(1 + \epsilon)} \right]^2 < 1 \quad \text{for all } y > 0, \epsilon > 0.$$

Let $\delta > 1$ be given by

$$\delta = \frac{\epsilon}{\log(1 + \epsilon)}.$$

Then this condition becomes

$$(\lambda + y)(\lambda + y + \epsilon y) < [\lambda + \delta y]^2,$$
or, equivalently,
\[ [\delta^2 - (1 + \epsilon)]y > \lambda [\epsilon - 2(\delta - 1)]. \]  
(19)

This equation puts a condition on the choice of \( y \). Take for instance \( \epsilon = \delta = e - 1 \). Then the condition reads
\[ y > \lambda \frac{3 - e}{e^2 - 3e + 1}. \]

The r.h.s. of this condition is positive. Hence there exist choices of \( y > 0 \) which do not satisfy the condition.

\[ \square \]

**Corollary** There exist hermitian matrices \( A \) and \( B \) which violate the operator version of (6), i.e. for which
\[ \exp \phi(A) - \exp \phi(B) \leq A - B \] if \( A \geq B \)
does not hold.

**Corollary** The function \( g(u) = \log \exp \phi(u) \) is increasing and concave, but not operator-monotone.

**References**

[1] Montrucchio, L., Pistone, G.: Deformed exponential bundle: The linear growth case. In: Geometric Science of Information, GSI 2017 LNCS proceedings, F. Nielsen and F. Barbaresco eds., (Springer, 2017), p. 239–246.

[2] Amari, S.: Differential-geometric methods in statistics. Lecture Notes in Statistics 28 (Springer, 1985).

[3] Amari, S., Nagaoka, H.: Methods of information geometry, Translations of mathematical monographs 191 (Am. Math. Soc., 2000; Oxford University Press, 2000); Originally in Japanese (Iwanami Shoten, Tokyo, 1993)

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

30
[5] Zhang, J., Hästö, P.: Statistical manifold as an affine space: A functional equation approach. J. Math. Psych. 50, 60–65 (2006).

[6] Ay, N., Jost, J., Ván Lê, H., Schwachhöfer, L.: Information Geometry (Springer, 2017).

[7] Petz D.: Quantum Information Theory and Quantum Statistics (Springer, 2008)

[8] Newton, N. J.: An infinite-dimensional statistical manifold modeled on Hilbert space. J. Funct. Anal. 263, 1661–1681 (2012).

[9] Tsallis, C.: Possible Generalization of Boltzmann-Gibbs Statistics. J. Stat. Phys. 52, 479–487 (1988).

[10] Naudts, J.: Deformed exponentials and logarithms in generalized thermostatistics. Physica A316, 323-334 (2002).

[11] Naudts, J.: Estimators, escort probabilities, and phi-exponential families in statistical physics. J. Ineq. Pure Appl. Math. 5, 102 (2004).

[12] Naudts, J.: Generalised Thermostatistics (Springer, 2011).

[13] Kato, T.: Perturbation theory for linear operators (Springer, 1966)