Noncommutative Orlicz Spaces in Quamtum Information Geometry

M. R. Grasselli

SharcNet Chair in Financial Mathematics
Department of Mathematics and Statistics
McMaster University

Supported by NSERC and MITACS

Guelph, November 30, 2004

1. Classical Parametric Information Geometry

- Study of differential geometric properties of families of classical probability densities.

- Given a probability space (Ω, Σ, μ), a family of probability densities $\mathcal{M} = \{p(x, \theta)\}$, for sample points $x \in \Omega$ and parameters $\theta = (\theta^1, \ldots, \theta^n) \in \mathbb{R}^n$ can be viewed as a Riemannian manifold equipped with the Fisher metric

$$g_{ij} = \int \frac{\partial \log p(x, \theta)}{\partial \theta^i} \frac{\partial \log p(x, \theta)}{\partial \theta^j} p(x, \theta) dx \quad (1)$$

Apart from the Levi-Civita connection associated with g, the statistical manifold $\mathcal{M}$ can be equipped with the exponential connection

$$\left(\nabla^{(1)}_{\frac{\partial}{\partial\theta^i}} \frac{\partial}{\partial\theta^j}\right)(p) = \frac{\partial^2 \log p}{\partial\theta^i \partial\theta^j} - E_p\left(\frac{\partial^2 \log p}{\partial\theta^i \partial\theta^j}\right),$$

and the mixture connection

$$\left(\nabla^{(-1)}_{\frac{\partial}{\partial\theta^i}} \frac{\partial}{\partial\theta^j}\right)(p) = \frac{\partial^2 \log p}{\partial\theta^i \partial\theta^j} + \frac{\partial \log p}{\partial\theta^i}\frac{\partial \log p}{\partial\theta^j},$$

which are dual to the metric g in the sense that $\langle\cdot,\cdot\rangle_p$ if

$$v\left(g(s_1, s_2)\right) = g(\nabla^{(1)}_v s_1, s_2) + g(s_1, \nabla^{(-1)}_v s_2) \tag{2}$$

for all $v \in T_p\mathcal{M}$ and all smooth vector fields s_1 and s_2.

One can also define a family of α–connections induced by the embeddings

$$\begin{aligned} \ell_\alpha \ : \ & \mathcal{M} \to \mathcal{A} \\ & p \mapsto \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}}, \end{aligned}$$

where $\mathcal{A}$ is the algebra of random variables on Ω and prove that they satisfy

$$\nabla^\alpha = \frac{1+\alpha}{2}\nabla^{(1)} + \frac{1-\alpha}{2}\nabla^{(-1)}. \tag{3}$$

The Fisher metric is the unique Riemannian metric reduced by all stochastic maps on the tangent bundle to $\mathcal{M}$. The duality of the α–connections with respect to it lead to rich minimization/projection theorems related to their associated α–divergences.

2. Classical Nonparametric Information Geometry

2.1. Classical Orlicz Spaces

Consider Young functions of the form

$$\Phi(x) = \int_0^{|x|} \phi(t)dt, \quad x \geq 0, \tag{4}$$

where $\phi : [0,\infty) \mapsto [0,\infty)$ is nondecreasing, continuous and such that $\phi(0) = 0$ and $\lim_{x\to\infty} \phi(x) = +\infty$. This include the monomials $|x|^r/r$, for $1 < r < \infty$, as well as the following examples:

$$\Phi_1(x) = \cosh x - 1, \tag{5}$$

$$\Phi_2(x) = e^{|x|} - |x| - 1, \tag{6}$$

$$\Phi_3(x) = (1 + |x|)\log(1 + |x|) - |x| \tag{7}$$

The complementary of a Young function Φ of the form (4) is given by

$$\Psi(y) = \int_0^{|y|} \psi(t)dt, \quad y \geq 0, \tag{8}$$

where ψ is the inverse of ϕ. One can verify that (Φ_2, Φ_3) and $(|x|^r/r, |x|^s/s)$, with $r^{-1} + s^{-1} = 1$, are examples of complementary pairs.

Now let (Ω, Σ, P) be a probability space. The Orlicz space associated with a Young function Φ defined as

$$L^{\Phi}(P) = \left\{ f : \Omega \mapsto \overline{\mathbf{R}}, \text{measurable} : \int_{\Omega} \Phi(\alpha f)dP < \infty, \text{ for some } \alpha > 0 \right\}.$$

If we identify functions which differ only on sets of measure zero, then L^{Φ} is a Banach space when furnished with the Luxembourg norm

$$N_{\Phi}(f) = \inf\left\{k > 0 : \int_{\Omega} \Phi(\frac{f}{k})dP \leq 1\right\}, \tag{9}$$

or with the equivalent Orlicz norm

$$\|f\|_{\Phi} = \sup\left\{\int_{\Omega} |fg|d\mu : g \in L^{\Psi}(\mu), \int_{\Omega} \Psi(g)dP \leq 1\right\}, \tag{10}$$

where Ψ is the complementary Young function to Ψ.

If Φ and Ψ are complementary Young functions, $f \in L^{\Phi}(P)$, $g \in L^{\Psi}(P)$, then we have the generalized Hölder inequality:

$$\int_{\Omega} |fg| dP \leq 2 N_{\Phi}(f) N_{\Psi}(g). \tag{11}$$

It follows that $L^{\Phi} \subset \left(L^{\Psi}\right)^{*}$ for any pair of complementary Young functions.

If $\Psi_2 \prec \Psi_1$ then there exist a constant k such that $\|\cdot\|_{\Psi_2} \leq k \|\cdot\|_{\Psi_1}$ and therefore $L^{\Psi_1}(P) \subset L^{\Psi_2}(P)$.

If two Young functions are equivalent, the Banach spaces associated with them coincide as sets and have equivalent norms.

2.2. The Pistone-Sempi Manifold

Consider the set

$$\mathcal{M} \equiv \mathcal{M}(\Omega, \Sigma, \mu) = \{f : \Omega \mapsto \mathbf{R}, f > 0 \text{ a.e. and } \int_\Omega f d\mu = 1\}.$$

For each point $p \in \mathcal{M}$, let $L^{\Phi_1}(p)$ be the exponential Orlicz space over the probability space $(\Sigma, \Omega, pd\mu)$ and consider its closed subspace of p-centred random variables

$$B_p = \{u \in L^{\Phi_1}(p) : \int_\Omega upd\mu = 0\} \tag{12}$$

as the coordinate Banach space.

In probabilistic terms, the set $L^{\Phi_1}(p)$ correspond to random variables whose moment generating function with respect to the probability $pd\mu$ is finite on a neighborhood of the origin.

They define one dimensional exponential models $p(t)$ associated with a point $p \in \mathcal{M}$ and a random variable u:

$$p(t) = \frac{e^{tu}}{Z_p(tu)} p, \qquad t \in (-\varepsilon, \varepsilon). \tag{13}$$

Define the inverse of a local chart around $p \in \mathcal{M}$ as

$$\begin{aligned} e_p : \mathcal{V}_p &\rightarrow \mathcal{M} \\ u &\mapsto \frac{e^u}{Z_p(u)} p. \end{aligned} \tag{14}$$

Denote by $\mathcal{U}_p$ the image of $\mathcal{V}_p$ under e_p. Let e_p^{-1} be the inverse of e_p on $\mathcal{U}_p$. Then a local chart around p is given by

$$\begin{aligned} e_p^{-1} : \mathcal{U}_p &\rightarrow B_p \\ q &\mapsto \log\left(\frac{q}{p}\right) - \int_\Omega \log\left(\frac{q}{p}\right) p d\mu. \end{aligned} \tag{15}$$

For any $p_1, p_2 \in \mathcal{M}$, the transition functions are given by

$$\begin{aligned} e_{p_2}^{-1} e_{p_1} : e_{p_1}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) &\rightarrow e_{p_2}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \\ u &\mapsto u + \log\left(\frac{p_1}{p_2}\right) - \int_\Omega \left(u + \log\frac{p_1}{p_2}\right) p_2 d\mu. \end{aligned}$$

Proposition 1 *For any $p_1, p_2 \in \mathcal{M}$, the set $e_{p_1}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ is open in the topology of B_{p_1}.*

We then have that the collection $\{(\mathcal{U}_p, e_p^{-1}), p \in \mathcal{M}\}$ satisfies the three axioms for being a C^∞–atlas for $\mathcal{M}$. Moreover, since all the spaces B_p are isomorphic as topological vector spaces, we can say that $\mathcal{M}$ is a C^∞–manifold modeled on $B_p \equiv T_p\mathcal{M}$.

Given a point $p \in \mathcal{M}$, the connected component of $\mathcal{M}$ containing p coincides with the maximal exponential model obtained from p: $\mathcal{E}(p) = \left\{\frac{e^u}{Z_p(u)} p, u \in B_p\right\}$.

2.3. The Fisher Information and Dual Connections

Let $\langle \cdot, \cdot \rangle_p$ be a continuous positive definite symmetric bilinear form assigned continuously to each $B_p \simeq T_p\mathcal{M}$. A pair of connection (∇, ∇^*) are said to be dual with respect to $\langle \cdot, \cdot \rangle_p$ if

$$\langle \tau u, \tau^* v \rangle_q = \langle u, v \rangle_p \tag{16}$$

for all $u, v \in T_p\mathcal{M}$, where τ and τ^* denote the parallel transports associated with ∇ and ∇^*, respectively.

Equivalently, (∇, ∇^*) are dual with respect to $\langle \cdot, \cdot \rangle_p$ if

$$v\left(\langle s_1, s_2 \rangle_p\right) = \langle \nabla_v s_1, s_2 \rangle_p + \langle s_1, \nabla_v^* s_2 \rangle_p \tag{17}$$

for all $v \in T_p\mathcal{M}$ and all smooth vector fields s_1 and s_2.

The infinite dimensional generalisation of the Fisher information is given by

$$\langle u, v\rangle_p = \int_\Omega (uv) p d\mu, \quad \forall u, v \in B_p. \tag{18}$$

This is clearly bilinear, symmetric and positive definite. Moreover, continuity follows from that fact that, since $L^{\Phi_1}(p) \simeq L^{\Phi_2}(p) \subset L^{\Phi_3}(p)$, the generalised Hölder inequality gives

$$|\langle u, v\rangle_p| \leq K\|u\|_{\Phi_1,p}\|v\|_{\Phi_1,p}, \quad \forall u, v \in B_p. \tag{19}$$

If p and q are two points on the same connected component of $\mathcal{M}$, then the exponential parallel transport is given by

$$\begin{aligned} \tau_{pq}^{(1)} : T_p\mathcal{M} &\rightarrow T_q\mathcal{M} \\ u &\mapsto u - \int_\Omega uq d\mu. \end{aligned} \qquad (20)$$

To obtain duality with respect to the Fisher information, we define the mixture parallel transport on $T\mathcal{M}$ as

$$\begin{aligned} \tau_{pq}^{(-1)} : T_p\mathcal{M} &\rightarrow T_q\mathcal{M} \\ u &\mapsto \frac{p}{q}u, \end{aligned} \qquad (21)$$

for p and q in the same connected component of $\mathcal{M}$.

Theorem 2 *The connections $\nabla^{(1)}$ and $\nabla^{(-1)}$ are dual with respect to the Fisher information.*

2.4. α–connections

We begin with Amari's α-embeddings

$$\begin{aligned} \ell_\alpha : \mathcal{M} &\rightarrow L^r(\mu) \\ p &\mapsto \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}}, \quad \alpha \in (-1,1), \end{aligned} \tag{22}$$

where $r = \frac{2}{1-\alpha}$. Observe that $\ell_\alpha(p) \in S^r(\mu)$, the sphere of radius r in $L^r(\mu)$.

Using the chain rule, the push-forward of the map ℓ_α can be implemented as

$$\begin{aligned} (\ell_\alpha)_{*(p)} : T_p\mathcal{M} = B_p &\rightarrow T_{rp^{1/r}} S^r(\mu) \\ u &\mapsto p^{\frac{1-\alpha}{2}} u. \end{aligned} \tag{23}$$

We are now ready to define the α-connections. In what follows, $\widetilde{\nabla}$ is used to denote the trivial connection on $L^r(\mu)$.

Definition 3 *For $\alpha \in (-1,1)$, let $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ be a smooth curve such that $p = \gamma(0)$ and $v = \dot{\gamma}(0)$ and let $s \in S(T\mathcal{M})$ be a differentiable vector field. The α-connection on $T\mathcal{M}$ is given by*

$$\left(\nabla_v^\alpha s\right)(p) = (\ell_\alpha)_{*(p)}^{-1}\left[\Pi_{rp^{1/r}}\widetilde{\nabla}_{(\ell_\alpha)_{*(p)}v}(\ell_\alpha)_{*(\gamma(t))}s\right]. \tag{24}$$

Theorem 4 *The exponential, mixture and α-covariant derivatives on $T\mathcal{M}$ satisfy*

$$\nabla^\alpha = \frac{1+\alpha}{2}\nabla^{(1)} + \frac{1-\alpha}{2}\nabla^{(-1)}. \tag{25}$$

Corollary 5 *The connections ∇^α and $\nabla^{-\alpha}$ are dual with respect to the Fisher information $\langle \cdot, \cdot \rangle_p$.*

3. Finite Dimensional Quantum Systems

- $\mathcal{H}^N$: finite dimensional complex Hilbert space;
- $\mathcal{B}(\mathcal{H}^N)$: algebra of operators on $\mathcal{H}^N$;
- $\mathcal{A}$: N^2-dimensional real vector subspace of self-adjoint operators;
- $\mathcal{M}$: n-dimensional submanifold of all invertible density operators on $\mathcal{H}^N$, with $n = N^2 - 1$.

3.1 Quantum α-connections

For $\alpha \in (-1, 1)$, define the α-embedding of $\mathcal{M}$ into $\mathcal{A}$ as

$$\begin{aligned} \ell_\alpha \ : \ & \mathcal{M} \to \mathcal{A} \\ & \rho \mapsto \frac{2}{1-\alpha} \rho^{\frac{1-\alpha}{2}}. \end{aligned}$$

At each point $\rho \in \mathcal{M}$, consider the subspace of $\mathcal{A}$ defined by

$$\mathcal{A}_\rho^{(\alpha)} = \left\{ A \in \mathcal{A} : \mathrm{Tr} \left(\rho^{\frac{1+\alpha}{2}} A \right) = 0 \right\},$$

and define the isomorphism

$$\begin{aligned} (\ell_\alpha)_{*(\rho)} \ : \ & T_\rho \mathcal{M} \to \mathcal{A}_\rho^{(\alpha)} \\ & v \mapsto (\ell_\alpha \circ \gamma)'(0). \end{aligned} \tag{26}$$

If $(\theta^1, \ldots, \theta^n)$ is a coordinate system for $\mathcal{M}$, then the α-representation of a basis tangent vector is

$$\frac{\partial}{\partial \theta^i} \mapsto \frac{\partial \ell_\alpha(\rho)}{\partial \theta^i}.$$

Lemma 6 (Hasegawa, 1996) *Let $= \rho(\theta)$ be a smooth manifold of invertible density matrices. Then there exists a anti-selfadjoint operator Δ_i such that*

$$\frac{\partial \rho}{\partial \theta^i} = \frac{\partial^c \rho}{\partial \theta^i} + [\rho, \Delta_i], \tag{27}$$

where $\frac{\partial^c \rho}{\partial \theta^i} \in \mathcal{C}(\rho)$ and $[\rho, \Delta_i] \in \mathcal{C}(\rho)^\perp$. Moreover, for any function F which is differentiable on a neighbourhood of the spectrum of ρ we have

$$\frac{\partial F(\rho)}{\partial \theta^i} = \frac{\partial^c F(\rho)}{\partial \theta^i} + [F(\rho), \Delta_i]. \tag{28}$$

Let $r = \frac{2}{1-\alpha}$. If we equip $\mathcal{A}$ with the the r-norm

$$\|A\|_r := (\mathrm{Tr}|A|^r)^{1/r}\,,$$

then the α-embedding can be vied as a mapping from $\mathcal{M}$ to the positive part of the sphere of radius r.

Definition 7 *For $\alpha \in (-1,1)$, let $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{M}$ be a smooth curve such that $\rho = \gamma(0)$ and $v = \dot{\gamma}(0)$ and let $s \in S(T\mathcal{M})$ be a differentiable vector field. The α-connection on $T\mathcal{M}$ is given by*

$$\left(\nabla_v^{(\alpha)} s\right)(\rho) = (\ell_\alpha)_{*(\rho)}^{-1} \left[\Pi_{r\rho^{1/r}} \widetilde{\nabla}_{(\ell_\alpha)_{*(\rho)} v} (\ell_\alpha)_{*(\gamma(t))} s\right], \qquad (29)$$

where $\Pi_{r\rho^{1/r}}$ is the canonical projection from the tangent space $T_{r\rho^{1/r}}\mathcal{A} = \mathcal{A}$ onto the tangent space $T_{r\rho^{1/r}} S^r = \mathcal{A}^{(\alpha)}$.

3.2. Monotone Metrics

Now let us consider the extended manifold of faithful weights $\widehat{\mathcal{M}}$ (the positive definite matrices) and use the -1-representation (the limiting case $\alpha = -1$ of the α-representations) in order to define a Riemannian metric g on $\widehat{\mathcal{M}}$ by means of the inner product $\langle \cdot, \cdot \rangle_\rho$ in $\mathcal{A} \subset B(\mathcal{H}^N)$. We say that $\hat{g}$ is monotone if and only if

$$\left\langle S(A^{(-1)}), S(A^{(-1)}) \right\rangle_{S(\rho)} \leq \left\langle A^{(-1)}, A^{(-1)} \right\rangle_\rho, \tag{30}$$

for every $\rho \in \mathcal{M}$, $A \in T_\rho \mathcal{M}$, and every completely positive, trace preserving map $S : \mathcal{A} \to \mathcal{A}$.

For any metric $\widehat{g}$ on $T\widehat{\mathcal{M}}$, define the positive (super) operator K_σ on $\mathcal{A}$ by

$$\widehat{g}_\rho(\widehat{A},\widehat{B}) = \left\langle \widehat{A}^{(-1)}, K_\rho\left(\widehat{B}^{(-1)}\right)\right\rangle_{HS} = \mathrm{Tr}\left(\widehat{A}^{(-1)} K_\rho\left(\widehat{B}^{(-1)}\right)\right). \quad (31)$$

Define also the (super) operators, $L_\rho X := \rho X$ and $R_\rho X := X\rho$, for $X \in \mathcal{A}$, which are also positive.

Theorem 8 (Petz 96) *A Riemannian metric g on $\mathcal{A}$ is monotone if and only if*

$$K_\sigma = \left(R_\sigma^{1/2} f(L_\sigma R_\sigma^{-1}) R_\sigma^{1/2}\right)^{-1},$$

where K_σ is defined in (31) and $f : R^+ \to R^+$ is an operator monotone function satisfying $f(t) = tf(t^{-1})$.

In particular, the BKM (Bogoluibov–Kubo–Mori) metric

$$g_\rho^B(A,B) = \int_0^\infty \mathrm{Tr}\left(\frac{1}{t+\rho}A^{(-1)}\frac{1}{t+\rho}B^{(-1)}\right)dt \qquad (32)$$

and the WYD (Wigner-Yanase-Dyson) metric

$$g_\rho^{(\alpha)}(A,B) := \mathrm{Tr}\left(A^{(\alpha)}B^{(-\alpha)}\right), \qquad A,B \in T_\rho\mathcal{M}, \qquad (33)$$

for $\alpha \in (-1,1)$ are special cases of monotone metrics corresponding respectively to the operator monotone functions

$$f^B(t) = \frac{t-1}{\log t}$$

and

$$f_p(x) = \frac{p(1-p)(x-1)^2}{(x^p-1)(x^{1-p}-1)}$$

for $p = \frac{1+\alpha}{2}$.

Theorem 9 *If the connections* $\nabla^{(1)}$ *and* $\nabla^{(-1)}$ *are dual with respect to a monotone Riemannian metric* g *on* $\mathcal{M}$*, then* g *is a* scalar *multiple of the BKM metric.*

Theorem 10 *If the connections* $\nabla^{(\alpha)}$ *and* $\nabla^{(-\alpha)}$ *are dual with respect to a monotone Riemannian metric* $\widehat{g}$ *on* $\widehat{\mathcal{M}}$*, then* $\widehat{g}$ *is a* scalar *multiple of the* WYD *metric.*

Corollary 11

$$\nabla^{\alpha} \neq \frac{1+\alpha}{2}\nabla^{(1)} + \frac{1-\alpha}{2}\nabla^{(-1)}. \tag{34}$$

4. Infinite dimensional quantum systems

- $\mathcal{H}$: infinite dimensional complex Hilbert space;

- $\mathcal{B}(\mathcal{H})$: algebra of operators on $\mathcal{H}$;

- $\mathcal{C}_p, 0 < p < 1$: compact operators $A : \mathcal{H} \mapsto \mathcal{H}$ such that $|A|^p \in \mathcal{C}_1$, where $\mathcal{C}_1$ is the set of trace-class operators on $\mathcal{H}$. Define

$$\mathcal{C}_{<1} := \bigcup_{0<p<1} \mathcal{C}_p.$$

- $\mathcal{M} = \mathcal{C}_{<1} \cap \Sigma$ where $\Sigma \subset \mathcal{C}_1$ denotes the set of normal faithful states on H.

4.1. ε-Bounded Perturbations

Let $H_0 \geq I$ be a self-adjoint operator with domain $\mathcal{D}(H_0)$, quadratic form q_0 and form domain $Q_0 = \mathcal{D}(H_0^{1/2})$, and let $R_0 = H_0^{-1}$ be its resolvent at the origin.

For $\varepsilon \in (0, 1/2)$, let $\mathcal{T}_\varepsilon(0)$ be the set of all symmetric forms X defined on Q_0 and such that $\|X\|_\varepsilon(0) := \left\| R_0^{\frac{1}{2}+\varepsilon} X R_0^{\frac{1}{2}-\varepsilon} \right\|$ is finite. Then the map $A \mapsto H_0^{\frac{1}{2}-\varepsilon} A H_0^{\frac{1}{2}+\varepsilon}$ is an isometry from the set of all bounded self-adjoint operators on $\mathcal{H}$ onto $\mathcal{T}_\varepsilon(0)$. Hence $\mathcal{T}_\varepsilon(0)$ is a Banach space with the ε-norm $\|\cdot\|_\varepsilon(0)$.

Lemma 12 *For fixed symmetric X, $\|X\|_\varepsilon$ is a monotonically increasing function of $\varepsilon \in [0, 1/2]$.*

4.2. Construction of the Manifold

To each $\rho_0 \in \mathcal{C}_{\beta_0} \cap \Sigma$, $\beta_0 < 1$, let $H_0 = -\log \rho_0 + cI \geq I$ be a self-adjoint operator with domain $\mathcal{D}(H_0)$ such that

$$\rho_0 = Z_0^{-1} e^{-H_0} = e^{-(H_0 + \Psi_0)}. \tag{35}$$

In $\mathcal{T}_\varepsilon(0)$, take X such that $\|X\|_\varepsilon(0) < 1 - \beta_0$. Since $\|X\|_0(0) \leq \|X\|_\varepsilon(0) < 1 - \beta_0$, X is also q_0-bounded with bound a_0 less than $1 - \beta_0$. The *KLMN* theorem then tells us that there exists a unique semi-bounded self-adjoint operator H_X with form $q_X = q_0 + X$ and form domain $Q_X = Q_0$. Following an unavoidable abuse of notation, we write $H_X = H_0 + X$ and consider the operator

$$\rho_X = Z_X^{-1} e^{-(H_0 + X)} = e^{-(H_0 + X + \Psi_X)}. \tag{36}$$

Then $\rho_X \in \mathcal{C}_{\beta_X} \cap \Sigma$, where $\beta_X = \frac{\beta_0}{1-a_0} < 1$ [Streater 2000]. We take as a neighbourhood $\mathcal{M}_0$ of ρ_0 the set of all such states, that is, $\mathcal{M}_0 = \{\rho_X : \|X\|_\varepsilon(0) < 1 - \beta_0\}$.

We want to use the Banach subspace of centred variables in $\mathcal{T}_\varepsilon(0)$ as generalized coordinates. For this, define the regularised mean of $X \in \mathcal{T}_\varepsilon(0)$ in the state ρ_0 as

$$\rho_0 \cdot X := \mathrm{Tr}(\rho_0^\lambda X \rho_0^{1-\lambda}), \qquad \text{for } 0 < \lambda < 1, \tag{37}$$

is finite, independent of λ and defines a continuous map from $\mathcal{T}_0(0)$ to $\mathbb{R}$. The map

$$\rho_X \mapsto \widehat{X} = X - \rho_0 \cdot X$$

is then a global chart for the Banach manifold $\mathcal{M}_0$ modeled by $\widehat{\mathcal{T}}_\varepsilon(0) = \{X \in \mathcal{T}_\varepsilon(0) : \rho_0 \cdot X = 0\}$. We extend our manifold by adding new patches compatible with $\mathcal{M}_0$.

4.3. Affine Structure and Analyticity of the Free Energy

Given two points ρ_X and ρ_Y in $\mathcal{M}_0$ and their tangent spaces $\widehat{\mathcal{T}}_\varepsilon(X)$ and $\widehat{\mathcal{T}}_\varepsilon(Y)$, we define the torsion free, flat, (+1)-parallel transport $\tau^{(1)}$ of $(Z - \rho_X \cdot Z) \in \widehat{\mathcal{T}}_\varepsilon(X)$ along any continuous path γ connecting ρ_X and ρ_Y in the manifold to be the point $(Z - \rho_Y \cdot Z) \in \widehat{\mathcal{T}}_\varepsilon(Y)$.

Theorem 13 *The free energy of the state* $\rho_X = Z_X^{-1} e^{-H_X} \in \mathcal{C}_{\beta_X} \subset \mathcal{M}, \beta_X < 1$, *defined by*

$$\Psi(\rho_X) := \log Z_X, \tag{38}$$

is infinitely Fréchet differentiable and has a convergent Taylor series for sufficiently small neighbourhoods of ρ_X *in* $\mathcal{M}$.

5. Noncommutative Orlicz Spaces

- $\mathcal{A}$: semifinite von Neumann algebra of operators on $\mathcal{H}$ with a faithful semifinite normal trace τ.

- $L^0(\mathcal{A}, \tau)$: closed densely defined operators $x = u|x|$ affiliated with $\mathcal{A}$ with the property that, for each $\varepsilon > 0$, there exits a projection $p \in \mathcal{A}$ such that $p\mathcal{H} \subset \mathcal{D}(x)$ and $\tau(1-p) \leq \varepsilon$ (trace measurable operators).

- $\tilde{x}(t) := \inf\{s > 0 : \tau(e_{(s,\infty)}) \leq t\}$, where $e_{(\cdot)}$ are the spectral projections of $|x|$ (rearrangement function).

Lemma 14 (Fack/Kosaki) *Let* $0 \leq a \in L^0(\mathcal{A}, \tau)$. *Then*

$$\tau(\phi(a)) = \int_0^\infty \phi(\tilde{a}(t))dt$$

for any continuous increasing function $\phi : [0, \infty) \to [0, \infty)$.

Definition 15 (Kunze) *The* *Orlicz space* *associated with* $(\mathcal{A}, \tau, \phi)$ *is*

$$L^\phi(\mathcal{A}, \tau) = \{x \in L^0(\mathcal{A}, \tau) : \tau(\phi(\lambda|x|)) \leq 1, \text{ for some } \lambda > 0\}.$$

Definition 16 (Zegarlinski) *Given a state* $\omega(a) = \tau(\rho a)$, *for an invertible density operator* ρ, *the* *Orlicz space* *associated with* $(\mathcal{A}, \omega, \phi)$ *is*

$$L^\phi(\mathcal{A}, \tau) = \{x \in L^0(\mathcal{A}, \tau) : O(\lambda x) \leq 1, \text{ for some } \lambda > 0\},$$

where, for a given $s \in [0, 1]$,

$$O(x) = \tau(\phi(|(\phi^{-1}(\rho))^s x (\phi^{-1}(\rho))^{1-s}|)).$$