Vol. 46 (2000) REPORTS ON MATHEMATICAL PHYSICS No. 3

THE QUANTUM INFORMATION MANIFOLD FOR ε-BOUNDED FORMS*

M. R. GRASSELLI† and R. F. STREATER

Department of Mathematics, King's College, Strand, London, WC2R 2LS, UK
(e-mail: ray.streater@kcl.ac.uk)

(Received October 20, 1999)

Let $H_0 \geq I$ be a self-adjoint operator and let V be a form-small perturbation such that $\|V\|_\epsilon := \left\| R_0^{1/2+\varepsilon} V R_0^{1/2-\varepsilon} \right\| < \infty$, where $\varepsilon \in (0, 1/2)$ and $R_0 = H_0^{-1}$. Suppose that there exists a positive $\beta < 1$ such that $Z_0 := \operatorname{Tr} e^{-\beta H_0} < \infty$. Let $Z := \operatorname{Tr} e^{-(H_0+V)}$. Then we show that the free energy $\Psi = \log Z$ is an analytic function of V in the sense of Fréchet, and that the family of density operators defined in this way is an analytic manifold.

Introduction

The use of differential geometric methods in parametric estimation theory is by now a fairly sound subject, whose foundations, applications and techniques can be found in several books [1, 7, 10]. The nonparametric version of this *information geometry* had its mathematical basis laid down in recent years [4, 16]. It is a genuine branch of infinite-dimensional analysis and geometry. The theory of quantum information manifolds aims to be its noncommutative counterpart [6, 11–13].

In this paper we generalise the results obtained by one of us [18, 19] to a larger class of potentials. In Section 1 we introduce ε-bounded perturbations of a given Hamiltonian and review their relation with form-bounded and operator-bounded perturbations. In Section 2 we construct a Banach manifold of quantum mechanical states with $(+1)$-affine structure and $(+1)$-connection, using the ε-bounded perturbations. Finally, in Section 3 we prove analyticity of the free energy Ψ_X in sufficiently small neighbourhoods in this manifold, from which it follows that the (-1)-coordinates are analytic.

1. ε-bounded perturbations

We recall the concepts of operator-bounded and form-bounded perturbations [8]. Given operators H and X defined on dense domains $\mathcal{D}(H)$ and $\mathcal{D}(X)$ in a Hilbert space $\mathcal{H}$, we say that X is *H-bounded* if

*Invited lecture of the XXXI Symposium on Mathematical Physics, Toruń, May 18–21, 1999, delivered by R. F. Streater.

†Supported by a grant from CAPES-Brazil.

(i) $\mathcal{D}(H) \subset \mathcal{D}(X)$ and
(ii) there exist positive constants a and b such that

$$\|X\psi\| \le a\|H\psi\| + b\|\psi\|, \qquad \text{for all} \quad \psi \in \mathcal{D}(H).$$

Analogously, given a positive self-adjoint operator H with associated form q_H and form domain $Q(H)$, we say that a symmetric quadratic form X (or the symmetric sesquiform obtained from it by polarization) is *q_H-bounded* if

(i) $Q(H) \subset Q(X)$ and
(ii) there exist positive constants a and b such that

$$|X(\psi, \psi)| \le a q_H(\psi, \psi) + b(\psi, \psi), \qquad \text{for all} \quad \psi \in Q(H).$$

In both cases, the infimum of such a is called the relative bound of X (with respect to H or with respect to q_H, accordingly).

Suppose that X is a quadratic form with domain $Q(X)$ and A, B are operators on $\mathcal{H}$ such that A^* and B are densely defined. Suppose further that $A^* : \mathcal{D}(A^*) \to Q(X)$ and $B : \mathcal{D}(B) \to Q(X)$. Then the expression AXB means the form defined by

$$\phi, \psi \mapsto X(A^*\phi, B\psi), \qquad \phi \in \mathcal{D}(A^*), \quad \psi \in \mathcal{D}(B).$$

With this definition in mind, let us specialise to the case where $H_0 \ge I$ is 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. Then it is easy to show that a symmetric operator $X : \mathcal{D}(H_0) \to \mathcal{H}$ is H_0-bounded if and only if $\|XR_0\| < \infty$. The following lemma is also known [18, lemma 2].

LEMMA 1. *A symmetric quadratic form X defined on Q_0 is q_0-bounded if and only if $R_0^{1/2}XR_0^{1/2}$ is a bounded symmetric form defined everywhere. Moreover, if $\|R_0^{1/2}XR_0^{1/2}\| < \infty$ then the relative bound a of X with respect to q_0 satisfies $a \le \|R_0^{1/2}XR_0^{1/2}\|$.*

The set $\mathcal{T}_\omega(0)$ of all H_0-bounded symmetric operators X is a Banach space with norm $\|X\|_\omega(0) := \|XR_0\|$, since the map $A \mapsto AH_0$ from $\mathcal{B}(H)$ onto $\mathcal{T}_\omega(0)$ is an isometry.

The set $\mathcal{T}_0(0)$ of all q_0-bounded symmetric forms X is also a Banach space with norm $\|X\|_0(0) := \|R_0^{1/2}XR_0^{1/2}\|$, since the map $A \mapsto H_0^{1/2}AH_0^{1/2}$ from the set of all bounded self-adjoint operators on $\mathcal{H}$ onto $\mathcal{T}_0(0)$ is again an isometry.

Now, 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 $\|X\|_\varepsilon(0) := \|R_0^{1/2+\varepsilon}XR_0^{1/2-\varepsilon}\|$ is finite. Then the map $A \mapsto H_0^{1/2-\varepsilon}AH_0^{1/2+\varepsilon}$ 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 $\|\cdot\|_\varepsilon(0)$. We note that $\mathcal{D}(H_0^{1/2}) \subset \mathcal{D}(H_0^{1/2-\delta})$, for all $0 \le \delta \le 1/2$.

We can now prove the following lemma.

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

Proof: We have to prove that $\|R_0^y X R_0^{1-y}\|$ is increasing for $y \in [1/2, 1]$ and decreasing for $y \in [0, 1/2]$. Let $\frac{1}{2} \leq \delta \leq 1$ and suppose that $\|R_0^\delta X R_0^{1-\delta}\| < \infty$. Interpolation theory for Banach spaces [17] and the fact that $\|R_0^\delta X R_0^{1-\delta}\| = \|R_0^{1-\delta} X R_0^\delta\|$ then give

$$\|R_0^x X R_0^{1-x}\| \leq \|R_0^\delta X R_0^{1-\delta}\|, \qquad \text{for all} \quad x \in [1-\delta, \delta],$$

and particularly for $\frac{1}{2} \leq y \leq \delta \leq 1$, we have

$$\|R_0^y X R_0^{1-y}\| \leq \|R_0^\delta X R_0^{1-\delta}\|.$$

On the other hand, for $0 \leq 1-\delta \leq y \leq \frac{1}{2}$,

$$\|R_0^y X R_0^{1-y}\| \leq \|R_0^\delta X R_0^{1-\delta}\| = \|R_0^{1-\delta} X R_0^\delta\|.$$ □

2. Construction of the manifold

2.1. The first chart

Let $\mathcal{C}_p, 0 < p < 1$, denote the set of 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

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

We take the underlying set of the quantum information manifold to be

$$\mathcal{M} = \mathcal{C}_{<1} \cap \Sigma,$$

where $\Sigma \subseteq \mathcal{C}_1$ denotes the set of density operators. We do so because the next step of our project is to look at the Orlicz space geometry associated with the quantum information manifold [4] and the quantum analogue of classical Orlicz space $L \log L$ seems to be

$$\mathcal{C}_1 \log \mathcal{C}_1 := \{\rho \in \mathcal{C}_1 : S(\rho) = -\sum \lambda_i \log \lambda_i < \infty\},$$

where $\{\lambda_i\}$ are the singular numbers of ρ. With this notation, the set of normal states of finite entropy is $\mathcal{C}_1 \log \mathcal{C}_1 \cap \Sigma$ and we have $\mathcal{C}_{<1} \subset \mathcal{C}_1 \log \mathcal{C}_1$. At this level, $\mathcal{M}$ has a natural affine structure defined as follows: let $\rho_1 \in \mathcal{C}_{p_1} \cap \Sigma$ and $\rho_2 \in \mathcal{C}_{p_2} \cap \Sigma$; take $p = \max\{p_1, p_2\}$, then $\rho_1, \rho_2 \in \mathcal{C}_p \cap \Sigma$, since $p \leq q$ implies $\mathcal{C}_p \subseteq \mathcal{C}_q$ [15]; define "$\lambda\rho_1 + (1-\lambda)\rho_2$, $0 \leq \lambda \leq 1$" as the usual sum of operators in $\mathcal{C}_p$. This is called the (-1)-affine structure.

We want to cover $\mathcal{M}$ by a Banach manifold. In [18] this is achieved defining hoods of $\rho \in \mathcal{M}$ using form-bounded perturbations. The manifold obtained there is shown to have a Lipschitz structure. In [19] the same is done with the more restrictive class of operator-bounded perturbations. The result then is that the manifold has an analytic structure. We now proceed using ε-bounded perturbations, with a similar result.

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

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

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

$$\rho_X = Z_X^{-1} e^{-(H_0+X)} = e^{-(H_0+X+\Psi_X)}. \tag{2}$$

Then $\rho_X \in \mathcal{C}_{\beta_X} \cap \Sigma$, where $\beta_X = \frac{\beta_0}{1-a_0} < 1$ [18, lemma 4]. The state ρ_X does not change if we add to H_X a multiple of the identity in such a way that $H_X + cI \geq I$, so we can always assume that, for the perturbed state, we have $H_X \geq I$. We take as a hood $\mathcal{M}_0$ of ρ_0 the set of all such states, that is, $\mathcal{M}_0 = \{\rho_X : \|X\|_\varepsilon(0) < 1-\beta_0\}$.

Because $\rho_X = \rho_{X+\alpha I}$, we introduce in $\mathcal{T}_\varepsilon(0)$ the equivalence relation $X \sim Y$ iff $X - Y = \alpha I$ for some $\alpha \in \mathbb{R}$. We then identify ρ_X in $\mathcal{M}_0$ with the line $\{Y \in \mathcal{T}_\varepsilon(0) : Y = X + \alpha I, \alpha \in \mathbb{R}\}$ in $\mathcal{T}_\varepsilon(0)/_\sim$. This is a bijection from $\mathcal{M}_0$ onto the subset of $\mathcal{T}_\varepsilon(0)/_\sim$ defined by $\{\{X + \alpha I\}_{\alpha\in\mathbb{R}} : \|X\|_\varepsilon(0) < 1-\beta_0\}$ and $\mathcal{M}_0$ becomes topologised by transfer of structure. Now that $\mathcal{M}_0$ is a (Hausdorff) topological space, we want to parametrise it by an open set in a Banach space. By analogy with the finite dimensional case [14, 5, 11], we want to use the Banach subspace of centred variables in $\mathcal{T}_\varepsilon(0)$; in our terms, perturbations with zero mean (the 'scores'). For this, define the regularised mean of $X \in \mathcal{T}_\varepsilon(0)$ in the state ρ_0 as

$$\rho_0 \cdot X := \operatorname{Tr}(\rho_0^\lambda X \rho_0^{1-\lambda}), \qquad \text{for} \quad 0 < \lambda < 1. \tag{3}$$

Since $\rho_0 \in \mathcal{C}_{\beta_0} \cap \Sigma$ and X is q_0-bounded, lemma 5 of [18] ensures that $\rho_0 \cdot X$ is finite and independent of λ. It was shown there that $\rho_0 \cdot X$ is a continuous map from $\mathcal{T}_0(0)$ to $\mathbb{R}$, because its bound contained a factor $\|X\|_0(0)$. Exactly the same proof shows that $\rho_0 \cdot X$ is a continuous map from $\mathcal{T}_\varepsilon(0)$ to $\mathbb{R}$. Thus the set $\widehat{\mathcal{T}_\varepsilon}(0) := \{X \in \mathcal{T}_\varepsilon(0) : \rho_0 \cdot X = 0\}$ is a closed subspace of $\mathcal{T}_\varepsilon(0)$ and so is a Banach space with the norm $\|\cdot\|_\varepsilon$ restricted to it.

To each $\rho_X \in \mathcal{M}_0$, consider the unique intersection of the equivalence class of X in $\mathcal{T}_\varepsilon(0)/_\sim$ with the set $\widehat{\mathcal{T}_\varepsilon}(0)$, that is, the point in the line $\{X + \alpha I\}_{\alpha\in\mathbb{R}}$ with $\alpha = -\rho_0 \cdot X$. Write $\widehat{X} = X - \rho_0 \cdot X$ for this point. The map $\rho_X \mapsto \widehat{X}$ is a homeomorphism between $\mathcal{M}_0$ and the open subset of $\widehat{\mathcal{T}_\varepsilon}(0)$ defined by $\{\widehat{X} : \widehat{X} = X - \rho_0 \cdot X,\ \|X\|_\varepsilon < 1-\beta_0\}$. The map $\rho_X \mapsto \widehat{X}$ is then a chart for the Banach manifold $\mathcal{M}_0$ modelled by $\widehat{\mathcal{T}_\varepsilon}(0)$. As usual, we identify the tangent space at ρ_0 with $\widehat{\mathcal{T}_\varepsilon}(0)$, the tangent curve $\{\rho(\lambda) = Z_{\lambda X}^{-1} e^{-(H_0+\lambda X)}, \lambda \in [-\delta, \delta]\}$ being identified with $\widehat{X} = X - \rho_0 \cdot X$.

2.2. Enlarging the manifold

We extend our manifold by adding new patches compatible with $\mathcal{M}_0$. The idea is to construct a chart around each perturbed state ρ_X as we did around ρ_0. Let $\rho_X \in \mathcal{M}_0$ with Hamiltonian $H_X \geq I$ and consider the Banach space $\mathcal{T}_\varepsilon(X)$ of all symmetric forms Y on Q_0 such that the norm $\|Y\|_\varepsilon(X) := \left\|R_X^{1/2+\varepsilon} Y R_X^{1/2-\varepsilon}\right\|$ is finite, where $R_X = H_X^{-1}$ denotes the resolvent of H_X at the origin. In $\mathcal{T}_\varepsilon(X)$, take Y such that $\|Y\|_\varepsilon(X) < 1 - \beta_X$. From Lemma 2 we know that Y is q_X-bounded with bound a_X less than $1-\beta_X$. Let H_{X+Y} be the unique semi-bounded self-adjoint operator, given by the *KLMN* theorem, with form $q_{X+Y} = q_X + Y = q_0 + X + Y$ and form domain $Q_{X+Y} = Q_X = Q_0$. Then the operator

$$\rho_{X+Y} = Z_{X+Y}^{-1} e^{-H_{X+Y}} = Z_{X+Y}^{-1} e^{-(H_0+X+Y)} \tag{4}$$

is in $\mathcal{C}_{\beta_Y} \cap \Sigma$, where $\beta_Y = \frac{\beta_X}{1-a_X}$.

We take as a neighbourhood of ρ_X the set $\mathcal{M}_X$ of all such states. Again $\rho_{X+Y} = \rho_{X+Y+\alpha I}$, so we furnish $\mathcal{T}_\varepsilon(X)$ with the equivalence relation $Z \sim Y$ iff $Z - Y = \alpha I$ and we see that $\mathcal{T}_\varepsilon(X)$ is mapped bijectively onto the set of lines

$$\{\{Z = Y + \alpha I\}_{\alpha\in\mathbb{R}}, \|Y\|_\varepsilon(X) < 1 - \beta_X\}$$

in $\mathcal{T}_\varepsilon(X)/_\sim$. In this way we topologise $\mathcal{M}_X$, by transfer of structure, with the quotient topology of $\mathcal{T}_\varepsilon(X)/_\sim$.

Again we can define the mean of Y in the state ρ_X by

$$\rho_X \cdot Y := \mathrm{Tr}(\rho_X^\lambda Y \rho_X^{1-\lambda}), \qquad \text{for} \quad 0 < \lambda < 1. \tag{5}$$

and notice that it is finite and independent of λ. This is a continuous function of Y with respect to the norm $\|\cdot\|_\varepsilon(X)$, hence $\widehat{\mathcal{T}}_\varepsilon(X) = \{Y \in \mathcal{T}_\varepsilon(X) : \rho_X \cdot Y = 0\}$ is closed and so is a Banach space with the norm $\|\cdot\|_\varepsilon(X)$ restricted to it. Finally, let $\widehat{Y}$ be the unique intersection of the line $\{Z = Y + \alpha I\}_{\alpha\in\mathbb{R}}$ with the hyperplane $\widehat{\mathcal{T}}_\varepsilon(X)$, given by $\alpha = -\rho_X \cdot Y$. Then $\rho_{X+Y} \mapsto \widehat{Y}$ is a homeomorphism between $\mathcal{M}_X$ and the open subset of $\widehat{\mathcal{T}}_\varepsilon(X)$ defined by $\left\{\widehat{Y} \in \widehat{\mathcal{T}}_\varepsilon(X) : \widehat{Y} = Y - \rho_X \cdot Y, \|Y\|_\varepsilon(X) < 1 - \beta_X\right\}$. We obtain that $\rho_{X+Y} \mapsto \widehat{Y}$ is a chart for the manifold $\mathcal{M}_X$ modelled by $\widehat{\mathcal{T}}_\varepsilon(X)$. The tangent space at ρ_X is identified with $\widehat{\mathcal{T}}_\varepsilon(X)$ itself.

We now look to the union of $\mathcal{M}_0$ and $\mathcal{M}_X$. We need to show that our two previous charts are compatible in the overlapping region $\mathcal{M}_0 \cap \mathcal{M}_X$. But first we prove the following series of technical lemmas.

LEMMA 3. *Let X be a symmetric form defined on Q_0 such that $\left\|R_0^{1/2} X R_0^{1/2}\right\| < 1$. Then $\mathcal{D}(H_0^{1/2-\varepsilon}) = \mathcal{D}(H_X^{1/2-\varepsilon})$, for any $\varepsilon \in (0, 1/2)$.*

Proof: We know that $\mathcal{D}(H_0^{1/2}) = \mathcal{D}(H_X^{1/2})$, since X is q_0-small. Moreover, H_X and H_0 are comparable as forms, that is, there exists $c > 0$ such that

$$c^{-1} q_0(\psi) \leq q_X(\psi) \leq c q_0(\psi), \qquad \text{for all} \quad \psi \in Q_0.$$

Using the fact that $x \mapsto x^{\alpha}$ $(0 < \alpha < 1)$ is an operator monotone function [3, Lemma 4.20], we conclude that

$$c^{-(1-2\varepsilon)} H_0^{1-2\varepsilon} \leq H_X^{1-2\varepsilon} \leq c^{1-2\varepsilon} H_0^{1-2\varepsilon},$$

which implies that $\mathcal{D}(H_0^{1/2-\varepsilon}) = \mathcal{D}(H_X^{1/2-\varepsilon})$. □

The conclusion remains true if we now replace H_X by $H_X + I$, if necessary in order to have $H_X \geq I$. This is assumed in the next corollary.

COROLLARY 1. *The operator* $H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon}$ *is bounded and has bounded inverse* $H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon}$.

Proof: $R_0^{1/2-\varepsilon}$ is bounded and maps $\mathcal{H}$ into $\mathcal{D}(H_0^{1/2-\varepsilon}) = \mathcal{D}(H_X^{1/2-\varepsilon})$. Then $H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon}$ is bounded, since $H_X^{1/2-\varepsilon}$ is closed. By exactly the same argument, we obtain that $H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon}$ is bounded. Finally, $(H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon})(H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon}) = (H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon})(H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon}) = I$. □

LEMMA 4. *For* $\varepsilon \in (0, 1/2)$, *let* X *be a symmetric form defined on* Q_0 *such that* $\|R_0^{1/2+\varepsilon} X R_0^{1/2-\varepsilon}\| < 1$. *Then* $R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon}$ *is bounded and has bounded inverse* $R_X^{1/2+\varepsilon} H_0^{1/2+\varepsilon}$. *Moreover,* $\mathcal{D}(H_0^{1/2+\varepsilon}) = \mathcal{D}(H_X^{1/2+\varepsilon})$.

Proof: From Lemma 2, we know that $\|R_0^{1/2} X R_0^{1/2}\| < 1$, so Lemma 3 and its corollary apply. We have that

$$1 > \|R_0^{1/2+\varepsilon} X R_0^{1/2-\varepsilon}\| = \|R_0^{1/2+\varepsilon}(H_X - H_0) R_0^{1/2-\varepsilon}\| = \|R_0^{1/2+\varepsilon} H_X R_0^{1/2-\varepsilon} - I\|,$$

thus $\|R_0^{1/2+\varepsilon} H_X R_0^{1/2-\varepsilon}\| < \infty$. We write this as

$$\|R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon} H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon}\| < \infty .$$

Since $H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon}$ is bounded and invertible, so is $R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon}$. Finally, the fact that $\|R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon}\| < \infty$ and $\|R_X^{1/2+\varepsilon} H_0^{1/2+\varepsilon}\| < \infty$ implies that $H_X^{1/2+\varepsilon}$ and $H_0^{1/2+\varepsilon}$ are comparable, hence $\mathcal{D}(H_0^{1/2+\varepsilon}) = \mathcal{D}(H_X^{1/2+\varepsilon})$. □

The next theorem ensures the compatibility between the two charts in the overlapping region $\mathcal{M}_0 \cap \mathcal{M}_X$.

THEOREM 1. $\|\cdot\|_\varepsilon(X)$ *and* $\|\cdot\|_\varepsilon(0)$ *are equivalent norms.*

Proof: We need to show that there exist positive constants m and M such that $m\|Y\|_\varepsilon(0) \leq \|Y\|_\varepsilon(X) \leq M\|Y\|_\varepsilon(0)$. We just write

$$\begin{aligned} \|Y\|_\varepsilon(X) &= \|R_X^{1/2+\varepsilon} H_0^{1/2+\varepsilon} R_0^{1/2+\varepsilon} Y R_0^{1/2-\varepsilon} H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon}\| \\ &\leq \|R_X^{1/2+\varepsilon} H_0^{1/2+\varepsilon}\| \|H_0^{1/2-\varepsilon} R_X^{1/2-\varepsilon}\| \|Y\|_\varepsilon(0) \\ &= M\|Y\|_\varepsilon(0) \end{aligned}$$

and, for the inequality in the other direction, we write

$$\begin{aligned}\|Y\|_\varepsilon(0) &= \left\| R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon} R_X^{1/2+\varepsilon} Y R_X^{1/2-\varepsilon} H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon} \right\| \\ &\leq \left\| R_0^{1/2+\varepsilon} H_X^{1/2+\varepsilon} \right\| \left\| H_X^{1/2-\varepsilon} R_0^{1/2-\varepsilon} \right\| \|Y\|_\varepsilon(X) \\ &= m^{-1} \|Y\|_\varepsilon(X). \qquad \square\end{aligned}$$

We see that $\mathcal{T}_\varepsilon(0)$ and $\mathcal{T}_\varepsilon(X)$ are, in fact, the same Banach space furnished with two equivalent norms, and observe that the quotient spaces $\mathcal{T}_\varepsilon(0)/_\sim$ and $\mathcal{T}_\varepsilon(X)/_\sim$ are exactly the same set. The general theory of Banach manifolds does the rest [9].

We continue in this way, adding a new patch around another point $\rho_{X'}$ in $\mathcal{M}_0$ or around some other point in $\mathcal{M}_X$ but outside $\mathcal{M}_0$. Whichever point we start from, we get a third piece $\mathcal{M}_X$ with chart into an open subset of the Banach space $\left\{Y \in \mathcal{T}_\varepsilon(X') : \rho_{X'} \cdot Y = 0\right\}$, with norm $\|Y\|_\varepsilon(X') := \left\| R_{X'}^{1/2+\varepsilon} Y R_{X'}^{1/2-\varepsilon} \right\|$ equivalent to the previously defined norms. We can go on inductively, and all the norms of any overlapping regions will be equivalent.

DEFINITION 1. The *information manifold* $\mathcal{M}(H_0)$ defined by H_0 consists of all states obtainable in a finite number of steps, by extending $\mathcal{M}_0$ as explained above.

These states are well defined in the following sense. If, for two different sets of perturbations $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, we have $X_1 + \cdots + X_n = Y_1 + \cdots + Y_m$ as forms on $\mathcal{D}(H_0^{1/2-\varepsilon})$, then we arrive at the same state either taking the route $X_1, \ldots, X_n$ or taking the route $Y_1, \ldots, Y_m$, since the self-adjoint operator associated with the form $q_0 + X_1 + \cdots + X_n = q_0 + Y_1 + \cdots + Y_m$ is unique.

2.3. Affine geometry in $\mathcal{M}(H_0)$

The set $A = \left\{\widehat{X} \in \widehat{\mathcal{T}}_\varepsilon(0) : \widehat{X} = X - \rho_0 \cdot X, \|X\|_\varepsilon(0) < 1 - \beta_0\right\}$ is a convex subset of the Banach space $\widehat{\mathcal{T}}_\varepsilon(0)$ and so has an affine structure coming from its linear structure. We provide $\mathcal{M}_0$ with an affine structure induced from A using the patch $\widehat{X} \mapsto \rho_X$ and call this the canonical or (+1)-affine structure. The (+1)-convex mixture of ρ_X and ρ_Y in $\mathcal{M}_0$ is then $\rho_{\lambda X+(1-\lambda)Y}$, $(0 \leq \lambda \leq 1)$, which differs from the previously defined (−1)-convex mixture $\lambda\rho_X + (1-\lambda)\rho_Y$.

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 (+1)-parallel transport U_L of $(Z - \rho_X \cdot Z) \in \widehat{\mathcal{T}}_\varepsilon(X)$ along any continuous path L connecting ρ_X and ρ_Y in the manifold to be the point $(Z - \rho_Y \cdot Z) \in \widehat{\mathcal{T}}_\varepsilon(Y)$. Clearly $U_L(0) = 0$ for every L, so the (+1)-affine connection given by U_L is torsion free. Moreover, U_L is independent of L by construction, thus the (+1)-affine connection is flat. We see that the (+1)-parallel transport just moves the representative point in the line $\{Z + \alpha I\}_{\alpha\in\mathbb{R}}$ from one hyperplane to another.

Now consider a second piece of the manifold, say $\mathcal{M}_X$. We have the (+1)-affine structure on it again by transfer of structure from $\widehat{\mathcal{T}}_\varepsilon(X)$. Since both $\widehat{\mathcal{T}}_\varepsilon(0)$ and $\widehat{\mathcal{T}}_\varepsilon(X)$ inherit their affine structures from the linear structure of the same set (either $\mathcal{T}_\varepsilon(0)$ or $\mathcal{T}_\varepsilon(X)$), we see that the (+1)-affine structures of $\mathcal{M}_0$ and $\mathcal{M}_X$ are the same on their overlap. We define the parallel transport in $\mathcal{M}_X$ again by moving representative points around. To parallel transport a point between any two tangent spaces in the union of the two pieces, we proceed by stages. For instance, if U denotes the parallel transport from ρ_0 to ρ_X, it is straightforward to check that U takes a convex mixture in $\widehat{\mathcal{T}_\varepsilon}(0)$ to a convex mixture in $\widehat{\mathcal{T}_\varepsilon}(X)$. So, if $\rho_Y \in \mathcal{M}_0$ and $\rho_{Y'} \in \mathcal{M}_X$ are points outside the overlap, we parallel transport from ρ_Y to $\rho_{Y'}$ following the route $\rho_Y \to \rho_0 \to \rho_X \to \rho_{Y'}$. Continuing in this way, we furnish the whole $\mathcal{M}(H_0)$ with a $(+1)$-affine structure and a flat, torsion free, $(+1)$-affine connection.

Although each hood in $\mathcal{M}(H_0)$ is clearly $(+1)$-convex, we have not been able to prove that $\mathcal{M}(H_0)$ is itself $(+1)$-convex.

3. Analyticity of the free energy

The free energy of the state $\rho_X = Z_X^{-1}e^{-H_X} \in \mathcal{C}_{\beta_X} \subset \mathcal{M}, \beta_X < 1$, is the function $\Psi : \mathcal{M} \to \mathbb{R}$ given by

$$\Psi(\rho_X) := \log Z_X. \tag{6}$$

In this section we show that $\Psi_X \equiv \Psi(\rho_X)$ is infinitely Fréchet differentiable and that it has a convergent Taylor series for sufficiently small hoods of ρ_X in $\mathcal{M}$.

We say that Y is an ε-bounded direction if $Y \in \mathcal{T}_\varepsilon(X)$. The n-th variation of the partition function Z_X in the ε-bounded directions $V_1, \dots, V_n$ is given by $(n!)^{-1}$ times the Kubo n-point function [2]

$$\mathrm{Tr} \int_0^1 d\alpha_1 \int_0^1 d\alpha_2 \cdots \int_0^1 d\alpha_{n-1} [\rho_X^{\alpha_1} V_1 \rho_X^{\alpha_2} V_2 \cdots \rho_X^{\alpha_n} V_n], \tag{7}$$

where $\alpha_n = 1 - \alpha_1 - \cdots - \alpha_{n-1}$. Our first task is to show that this is finite. Since for an operator of trace class A we have $|\mathrm{Tr} A| \le \|A\|_1$, we only need to check that the multiple integral is of trace class.

We begin by estimating the trace of $[\rho_X^{\alpha_1} V_1 \rho_X^{\alpha_2} V_2 \cdots \rho_X^{\alpha_n} V_n]$ as written as

$$\begin{aligned}[\rho_X^{\alpha_1\beta_X}][H_X^{1-\delta_n+\delta_1}\rho_X^{(1-\beta_X)\alpha_1}][R_X^{\delta_1} V_1 R_X^{1-\delta_1}][\rho_X^{\alpha_2\beta_X}][H_X^{1-\delta_1+\delta_2}\rho_X^{(1-\beta_X)\alpha_2}]\\ [R_X^{\delta_2} V_2 R_X^{1-\delta_2}]\cdots[\rho_X^{\alpha_n\beta_X}][H_X^{1-\delta_{n-1}+\delta_n}\rho_X^{(1-\beta_X)\alpha_n}][R_X^{\delta_n} V_n R_X^{1-\delta_n}],\end{aligned}$$

with $\delta_j \in \left[1/2 - \varepsilon, 1/2 + \varepsilon\right]$ to be specified soon. In this product, we have n factors of the form $[\rho_X^{\alpha_j\beta_X}]$, n factors of the form $[R_X^{\delta_j} V_j R_X^{1-\delta_j}]$, and n factors of the form $[H_X^{1-\delta_{j-1}+\delta_j}\rho_X^{(1-\beta_X)\alpha_j}]$, with δ_0 standing for δ_n.

For the factors $[\rho_X^{\alpha_j\beta_X}]$, putting $p_j = 1/\alpha_j$, Hölder's inequality leads to the trace norm bound

$$\left\|[\rho_X^{\alpha_1\beta_X}]\cdots[\rho_X^{\alpha_n\beta_X}]\right\|_1 \le \left\|\rho_X^{\beta_X}\right\|_1^{\alpha_1}\cdots\left\|\rho_X^{\beta_X}\right\|_1^{\alpha_n} = \left\|\rho_X^{\beta_X}\right\|_1 < \infty. \tag{8}$$

By virtue of Lemma 2, we know that the factors $[R_X^{\delta_j} V_j R_X^{1-\delta_j}]$ are bounded in operator norm by

$$\|R_X^{\delta_j} V_j R_X^{1-\delta_j}\| \le \|R_X^{1/2+\varepsilon} V_j R_X^{1/2-\varepsilon}\| = \|V_j\|_\varepsilon(X) < \infty. \tag{9}$$

In both these cases, the bounds are independent of α. The hardest case turns out to be the factors $[H_X^{1-\delta_{j-1}+\delta_j} \rho_X^{(1-\beta_X)\alpha_j}]$, where the estimate, as we will see, does depend on α and we have to worry about integrability. For them, the spectral theorem gives the operator norm bound

$$\begin{aligned}\|H_X^{1-\delta_{j-1}+\delta_j} \rho_X^{(1-\beta_X)\alpha_j}\| &= Z_X^{-\alpha_j(1-\beta_X)} \sup_{x\ge 1} \left\{x^{1-\delta_{j-1}+\delta_j} e^{-(1-\beta_X)\alpha_j x}\right\} \\ &\le Z_X^{-\alpha_j(1-\beta_X)} \left(\frac{1-\delta_{j-1}+\delta_j}{(1-\beta_X)\alpha_j}\right)^{1-\delta_{j-1}+\delta_j} e^{-(1-\delta_{j-1}+\delta_j)}. \end{aligned}\tag{10}$$

Apart from $\alpha_j^{-(1-\delta_{j-1}+\delta_j)}$, the other terms in (10) will be bounded independently of α. To deal with the integral of $\alpha_j^{-(1-\delta_{j-1}+\delta_j)} d\alpha_j$, we divide the region of integration in n (overlapping) regions $S_j := \{\alpha : \alpha_j \ge 1/n\}$ (since $\sum \alpha_j = 1$). For the region S_n, for instance, the integrability at $\alpha_j = 0$ is guaranteed if we choose δ_j such that $\delta_j < \delta_{j-1}$. So we take $\delta_n = \delta_0 > \delta_1 > \cdots > \delta_{n-1}$. We must have $\delta_j \in \left[\frac{1}{2} - \varepsilon, 1/2 + \varepsilon\right]$, then we choose $\delta_n = \frac{1}{2} + \varepsilon$, $\delta_1 = \frac{1}{2} + \varepsilon - \frac{2\varepsilon}{n}$, $\delta_2 = \frac{1}{2} + \varepsilon - \frac{4\varepsilon}{n}$, $\ldots$, $\delta_{n-1} = \frac{1}{2} - \varepsilon + \frac{2\varepsilon}{n}$. Then each of the $(n-1)$ integrals, for $j = 1, \ldots, n-1$, is

$$\int_0^1 \alpha_j^{-(1-\delta_{j-1}+\delta_j)} d\alpha_j = (\delta_{j-1} - \delta_j)^{-1} = \frac{n}{2\varepsilon}$$

resulting in a contribution of $\left(\frac{n}{2\varepsilon}\right)^{n-1}$. The last integrand in S_n is $\alpha_n^{-(1-\delta_{n-1}+\delta_n)} \le n^2$. The same bound holds for the other regions S_j, $j = 1, \ldots, n-1$, giving a total bound

$$\prod_{j=1}^n \int_0^1 \alpha_j^{-(1-\delta_{j-1}+\delta_j)} d\alpha_j \le n \left[\frac{n^2 n^{n-1}}{(2\varepsilon)^{n-1}}\right] = \frac{n^2 n^n}{(2\varepsilon)^{n-1}}. \tag{11}$$

Now that we have fixed δ_j, the promised bound for the other terms in (10) is

$$\prod_{j=1}^n Z_X^{-\alpha_j(1-\beta_X)} \left(\frac{1-\delta_{j-1}+\delta_j}{1-\beta_X}\right)^{1-\delta_{j-1}+\delta_j} \le 4 Z_X^{-(1-\beta_X)} (1-\beta_X)^{-n} e^{-n} \tag{12}$$

since $(1 - \delta_{j-1} + \delta_j) < 1$ except for one term, which is less than 2.

Collecting the estimates (8), (9), (11) and (12), we get the following bound for the n-point function

$$4 \|\rho_X^{\beta_X}\|_1 Z_X^{-(1-\beta_X)} (2\varepsilon) n^2 n^n e^{-n} \prod_j \left[\frac{\|V_j\|_\varepsilon(X)}{2\varepsilon(1-\beta_X)}\right]. \tag{13}$$

Thus the n-th variation of Z_X exists for any ε-bounded directions and is an n-linear bounded map. Hence [21, Prop. 4.20], Z has an n-th Gateaux derivative at X. Since this holds for any n, we see that Z is infinitely often Gateaux differentiable at X. Moreover, when using Duhamel's formula [18, Theorem 9] to deduce the expression (7) for the n-th variation (as in [19, Theorem 3]), we actually find that the limit procedure is uniform in V, thence [20, Theorem 3.3] the Gateaux derivatives of Z at X are, in fact, Fréchet derivatives.

Therefore, Z is infinitely Fréchet differentiable with convergent Taylor expansion for $Z(X+V)$ if $\|V\|_\varepsilon(X) < (1-\beta_X)2\varepsilon$. Since Z_X is positive, the same is true for its logarithm, the free energy Ψ_X. Notice that the condition $\|V\|_\varepsilon(X) < (1-\beta_X)2\varepsilon$ is stronger than to require that ρ_{V+X} lie in an ε-hood of ρ_X.

Finally, let us say that a map $\Phi : \mathcal{U} \to \mathbb{R}$, on a hood $\mathcal{U}$ in $\mathcal{M}$, is $(+1)$-analytic in $\mathcal{U}$ if it is infinitely often Fréchet differentiable and $\Phi(X+V) \equiv \Phi(\rho_{X+V})$ has a convergent Taylor expansion for ρ_{X+V} in this hood. In particular, the (-1)-coordinates $\eta_X = \rho \cdot X$ (mixture coordinates) are analytic, since they are derivatives of the free energy Ψ_X. This specification of the sheaf of germs of analytic functions defines a real analytic structure on the manifold.

Acknowledgments

We thank E. B. Davies for help with Lemma 3. RFS thanks Rector Jamiołkowski of Toruń University for hospitality, May 1999, where this work was started.

REFERENCES

[1] S.-I. Amari: *Differential Geometric Methods in Statistics, Lecture Notes in Statistics* **28**, Springer, New York 1985.
[2] H. Araki: Relative Hamiltonians for Faithful Normal States of a von Neumann Algebra, *Publ. R. I. M. S. (Kyoto)*, **9** (1968), 165–209.
[3] E. B. Davies: *One-parameter Semigroups*, Academic Press, New York 1980.
[4] P. Gibilesco and G. Pistone: Connections on nonparametric statistical manifolds by Orlicz space geometry, *Inf. Dim. Analysis, Quant. Prob. and Rel. Top.*, **1** (1998), 325–347.
[5] H. Hasagawa: *Rep. Math. Phys.* **33** (1993), 87–93.
[6] H. Hasagawa: *Noncommutative Extension of the Information Geometry*, in *Quantum Communication and Measurement*, eds. V. P. Balavkin, O. Hirota and R. L. Hudson, Plenum Press, New York 1995.
[7] R. A. Kass and P. W. Vos: *Geometric Foundations of Asymptotic Inference*, Wiley, New York 1997.
[8] T. Kato: *Perturbation Theory for Linear Operators*, Springer, Berlin 1966.
[9] S. Lang: *Differential and Riemannian Manifolds*, Springer, Berlin 1995.
[10] M. K. Murray and J. W. Rice: *Differential Geometry and Statistics, Monographs on Statistics and Applied Probability* **48**, Chapman & Hall, London 1993.
[11] H. Nagaoka: *Differential Geometric Aspects of Quantum State Estimation and Relative Entropy*, in *Quantum Communication and Measurement*, eds. V. P. Balavkin, O. Hirota and R. L. Hudson, Plenum Press, 1995.
[12] D. Petz: *J. Math. Phys.* **35** (1994), 780–795.
[13] D. Petz and C. Sudar: *J. Math. Phys.* **37** (1996), 2662–2673.
[14] D. Petz and G. Toth: *Lett. Math. Phys.* **27** (1993), 205–216.
[15] A. Pietsch: *Nuclear Locally Convex Spaces*, Springer, Berlin 1972.
[16] G. Pistone and C. Sempi: *Ann. Stat.* **33** (1995), 1543–1561.
[17] M. Reed and B. Simon: *Methods of Modern Mathematical Physics*, vol. 2, Academic Press, 1975.

[18] R. F. Streater: *The Information Manifold for Relatively Bounded Potentials*, *Proc. Steklov Inst. Math.* **228** (2000), 205–223.
[19] R. F. Streater: *The Analytic Quantum Information Manifold*, to appear in *Stochastic Processes, Physics and Geometry*, eds. F. Gesztesy, S. Paycha and H. Holden, Canad. Math. Soc.
[20] M. M. Vainberg: *Variational Methods for the Study of Nonlinear Operators*, Holden-Day, 1964.
[21] E. Zeidler: *Nonlinear Functional Analysis and its Applications*, vol. 1, Springer, Berlin 1985.