The Quantum Information Manifold for $\varepsilon$-Bounded Forms

M. R. Grasselli*    R. F. Streater
Dept. of Mathematics,
King’s College,
Strand,
London, WC2R 2LS
29/09/1999

Abstract

Let $H_0 \geq I$ be a self-adjoint operator and let $V$ be a form-small perturbation such that $\| R_0^{\frac{1}{2} + \varepsilon} VR_0^{\frac{1}{2} - \varepsilon} \| < \infty$, where $\varepsilon \in (0, 1/2)$ and $R_0 = H_0^{-1}$.

Suppose that there exists a positive $\beta < 1$ such that $Z := \text{Tr} e^{-\beta H_0} < \infty$. Then we show that the free energy $\Psi = \log Z$ is an analytic function in the sense of Fréchet, and that the family of density operators defined in this way is an analytic manifold.

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 non-parametric 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, 12].

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

*Supported by a grant from CAPES-Brazil.
1 $\varepsilon$-Bounded Perturbations

We recall the concepts of operator-bounded and form-bounded perturbations [18].

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

i. $\mathcal{D}(H) \subset \mathcal{D}(X)$ and

ii. there exist positive constants $a$ and $b$ such that

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

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

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

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

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

With this definition in mind, let us specialise to the case where $H_0 \geq I$ 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 $\left\|R_0^{1/2}XR_0^{1/2}\right\| < \infty$ then the relative bound $a$ of $X$ with respect to $q_0$ satisfies $a \leq \left\|R_0^{1/2}XR_0^{1/2}\right\|$.

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 $\mathcal{B}(\mathcal{H})$ 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 $\|X\|_0(0) := \left\|R_0^{1/2}XR_0^{1/2}\right\|$, 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 $T_\varepsilon(0)$ be the set of all symmetric forms $X$ with $D(H_0^{1/2-\varepsilon}) \subset Q(X)$ and such that $\|X\|_\varepsilon(0) := \left\| R_0^{1/2 + \varepsilon} X R_0^{1/2 - \varepsilon} \right\|$ is finite. Then the map $A \mapsto H_0^{1/2 - \varepsilon} A H_0^{1/2 + \varepsilon}$ is an isometry from the set of all bounded self-adjoint operators on $\mathcal{H}$ onto $T_\varepsilon(0)$. Hence $T_\varepsilon(0)$ is a Banach space with the $\varepsilon$-norm $\| \cdot \|_\varepsilon(0)$.

We note that $D(H_0^{1/2}) \subset D(H_0^{1/2 - \delta})$, for all $0 \leq \delta \leq 1/2$.

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 $\left\| R_0^{1/2 \cdot x} X R_0^{1/2 - y} \right\|$ is increasing for $y \in [1/2, 1]$ and decreasing for $y \in [0, 1/2]$. Let $1/2 \leq \delta \leq 1$ and suppose that $\left\| R_0^{\delta} X R_0^{1 - \delta} \right\| < \infty$. Interpolation theory for Banach spaces [17] and the fact that $\left\| R_0^{\delta} X R_0^{1 - \delta} \right\| = \left\| R_0^{1 - \delta} X R_0^{\delta} \right\|$ then give

$$\left\| R_0^{y} X R_0^{1 - y} \right\| \leq \left\| R_0^{\delta} X R_0^{1 - \delta} \right\|,$$

for all $x \in [1 - \delta, \delta]$, and particularly for $1/2 \leq y \leq \delta \leq 1$, we have

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

By the other hand, for $0 \leq 1 - \delta \leq y \leq 1/2$,

$$\left\| R_0^{y} X R_0^{1 - y} \right\| \leq \left\| R_0^{\delta} X R_0^{1 - \delta} \right\| = \left\| R_0^{1 - \delta} X R_0^{\delta} \right\|. \quad \square$$

## 2 Construction of the Manifold

### 2.1 The First Chart

Let $C_p, 0 < p < 1$, denote the set of compact operators $A : \mathcal{H} \mapsto \mathcal{H}$ such that $|A|^p \in C_1$, where $C_1$ is the set of trace-class operators on $\mathcal{H}$. Define

$$C_{<1} := \bigcup_{0<p<1} C_p.$$

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

$$\mathcal{M} = C_{<1} \cap \Sigma$$

where $\Sigma \subseteq C_1$ 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 [1] and the quantum analogue of classical Orlicz space $L \log L$ seems to be

$$C_1 \log C_1 := \{ \rho \in C_1 : S(\rho) = -\sum \lambda_i \log \lambda_i < \infty \},$$

3
where \( \{\lambda_i\} \) are the singular numbers of \( \rho \). With this notation, the set of normal states of finite entropy is \( C_{<1}\log C_1 \cap \Sigma \) and we have \( C_{<1} \subset C_{1}\log C_1 \). At this level, \( \mathcal{M} \) has a natural affine structure defined as follows: let \( \rho_1 \in C_{p_1} \cap \Sigma \) and \( \rho_2 \in C_{p_2} \cap \Sigma \); take \( p = \max\{p_1, p_2\} \), then \( \rho_1, \rho_2 \in C_p \cap \Sigma \), since \( p \leq q \) implies \( C_p \subseteq C_q \). Define "\( \lambda \rho_1 + (1 - \lambda)\rho_2, 0 \leq \lambda \leq 1 \)" as the usual sum of operators in \( C_p \). 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 \( \varepsilon \)-bounded perturbations, with a similar result.

To each \( \rho_0 \in C_{\beta_0} \cap \Sigma \), \( \beta_0 < 1 \), let \( H_0 = -\log \rho_0 + cI \geq I \) be a self-adjoint operator with domain \( \mathcal{D}(H_0) \) such that
\[
\rho_0 = Z_0^{-1}e^{-H_0} = e^{-(H_0+\Psi_0)}. \tag{3}
\]

In \( T_\varepsilon(0) \), 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
\[
\rho_X = Z_X^{-1}e^{-(H_0+X)} = Z_X^{-1}e^{-(H_0+X+\Psi_X)}. \tag{4}
\]

Then \( \rho_X \in C_{\beta_X} \cap \Sigma \), where \( \beta_X = \frac{\beta_0}{1-a_0} < 1 \) \( [13] \), lemma 4 \). The state \( \rho_X \) 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 \( \rho_0 \) the set of all such states, that is, \( \mathcal{M}_0 = \{\rho_X : \|X\|_\varepsilon(0) < 1 - \beta_0\} \).

Because \( \rho_X = \rho_X + aI \), we introduce in \( T_\varepsilon(0) \) the equivalence relation \( X \sim Y \) if \( X - Y = aI \) for some \( a \in \mathbb{R} \). We then identify \( \rho_X \) in \( \mathcal{M}_0 \) with the line \( \{Y \in T_\varepsilon(0) : Y = X + aI, a \in \mathbb{R}\} \) in \( T_\varepsilon(0)/\sim \). This is a bijection from \( \mathcal{M}_0 \) onto the subset of \( T_\varepsilon(0)/\sim \) defined by \( \{X + aI \in \mathbb{R} : \|X\|_\varepsilon(0) < 1 - \beta_0\} \) 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 \( [14, 6, 17] \), we want to use the Banach subspace of centred variables in \( T_\varepsilon(0) \); in our terms, perturbations with zero mean (the ‘scores’). For this, define the regularised mean of \( X \in T_\varepsilon(0) \) in the state \( \rho_0 \) as
\[
\rho_0 \cdot X := Tr(\rho_0 \cdot X \rho_0^{1-\lambda}), \quad \text{for } 0 < \lambda < 1. \tag{5}
\]

Since \( \rho_0 \in C_{\beta_0} \cap \Sigma \) and \( X \) is \( q_0 \)-bounded, lemma 5 of \( [18] \) ensures that \( \rho_0 \cdot X \) is finite and independent of \( \lambda \). It was a shown there that \( \rho_0 \cdot X \) is a continuous map from \( T_\varepsilon(0) \) 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 \( T_\varepsilon(0) \) to \( \mathbb{R} \). Thus the set
\[ \hat{T}_\varepsilon(0) := \{ X \in T_\varepsilon(0) : \rho_0 \cdot X = 0 \} \] is a closed subspace of \( T_\varepsilon(0) \) 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 \( T_\varepsilon(0)/\sim \) with the set \( \hat{T}_\varepsilon(0) \), that is, the point in the line \( (X + \alpha I)_{\alpha \in \mathbb{R}} \) with \( \alpha = -\rho_0 \cdot X \). Write \( \hat{X} = X - \rho_0 \cdot X \) for this point. The map \( \rho_X \mapsto \hat{X} \) is a homeomorphism between \( \mathcal{M}_0 \) and \( M \) restricted to it. Finally, let \( \hat{\alpha}_I \) is to construct a chart around each perturbed state and notice that it is finite and independent of \( \varepsilon \).

We extend our manifold by adding new patches compatible with \( \mathcal{M}_0 \) modeled by \( \hat{T}_\varepsilon(0) \). As usual, we identify the tangent space at \( \rho_0 \) with \( \hat{T}_\varepsilon(0) \), the tangent curve \( \{ (\lambda) = Z_{\lambda X}^1 e^{-(H_0 + \lambda X)}, \lambda \in [-\delta, \delta] \} \) being identified with \( \hat{X} = X - \rho_0 \cdot X \).

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 \( \rho_X \) as we did around \( \rho_0 \). Let \( \rho_X \in \mathcal{M}_0 \) with Hamiltonian \( H_X \geq I \) and consider the Banach space \( T_\varepsilon(X) \) of all symmetric forms \( Y \) such that the norm \( \| Y \|_\varepsilon(X) := \left\| R_X^{1+\varepsilon} Y R_X^{1/\varepsilon} \right\| \) is finite, where \( R_X = H_X^{-1} \) denotes the resolvent of \( H_X \) at the origin. In \( T_\varepsilon(X) \), take \( Y \) such that \( \| Y \|_\varepsilon(X) < 1 - \beta X \). From lemma \[ \ref{lemma:boundedness} \] 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

\[ \rho_{X+Y} = Z_{X+Y}^{-1} e^{-H_{X+Y}} = Z_{X+Y}^{-1} e^{-(H_0 + X + Y)} \quad (6) \]

is in \( C_{\beta_Y} \cap \Sigma \), where \( \beta_Y = \frac{\beta}{1 - a_X} \).

We take as a neighbourhood of \( \rho_X \) the set \( \mathcal{M}_X \) of all such states. Again \( \rho_{X+Y} = \rho_X + \alpha I \), so we furnish \( T_\varepsilon(X) \) with the equivalence relation \( Z \sim Y \) iff \( Z - Y = \alpha I \) and we see that \( T_\varepsilon(X) \) is mapped bijectively onto the set of lines \( \{ Z = Y + \alpha I \}_{\alpha \in \mathbb{R}} : \| Y \|_\varepsilon(X) < 1 - \beta X \} \) in \( T_\varepsilon(X)/\sim \). In this way we topologise \( \mathcal{M}_X \), by transfer of structure, with the quotient topology of \( T_\varepsilon(X)/\sim \).

Again we can define the mean of \( Y \) in the state \( \rho_X \) by

\[ \rho_X \cdot Y := Tr(\rho_X Y \rho_X^{1-\lambda}) \quad \text{for } 0 < \lambda < 1. \quad (7) \]

and notice that it is finite and independent of \( \lambda \). This is a continuous function of \( Y \) with respect to the norm \( \| \cdot \|_\varepsilon(X) \), hence \( \tilde{T}_\varepsilon(X) = \{ Y \in T_\varepsilon(X) : \rho_X \cdot Y = 0 \} \) is closed and so is a Banach space with the norm \( \| \cdot \|_\varepsilon(X) \) restricted to it. Finally, let \( \tilde{Y} \) be the unique intersection of the line \( \{ Z = Y + \alpha I \}_{\alpha \in \mathbb{R}} \) with the hyperplane \( \tilde{T}_\varepsilon(X) \), given by \( \alpha = -\rho_X \cdot Y \). Then \( \rho_{X+Y} \mapsto \tilde{Y} \) is a homeomorphism between \( \mathcal{M}_X \) and the open subset of \( \tilde{T}_\varepsilon(X) \) defined by \( \{ \tilde{Y} \in \tilde{T}_\varepsilon(X) : \tilde{Y} = Y - \rho_X \cdot Y, \| Y \|_\varepsilon(X) < 1 - \beta X \} \}. We obtain that \( \rho_{X+Y} \mapsto \tilde{Y} \) is a global chart for the manifold \( \mathcal{M}_X \) modeled by \( \tilde{T}_\varepsilon(X) \). The tangent space at \( \rho_X \) is identified with \( \tilde{T}_\varepsilon(X) \) 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 8** Let $X$ be a symmetric form defined on $Q_0$ such that $\|R_0^{1/2}XR_0^{1/2}\| < 1$. Then $\mathcal{D}(H_0^{\frac{1}{2}-\varepsilon}) = \mathcal{D}(H_X^{\frac{1}{2}-\varepsilon})$, 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

$$c^{-1}q_0(\psi) \leq q_X(\psi) \leq cq_0(\psi), \quad \text{for all } \psi \in Q_0.$$

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

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

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

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

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

**Lemma 10** For $\varepsilon \in (0, 1/2)$, let $X$ be a form defined on $\mathcal{D}(H_0^{\frac{1}{2}-\varepsilon})$ such that $\|R_0^{\frac{1}{2}+\varepsilon}XR_0^{\frac{1}{2}-\varepsilon}\| < 1$. Then $R_0^{\frac{1}{2}+\varepsilon}H_X^{\frac{1}{2}+\varepsilon}$ is bounded and has bounded inverse $R_X^{\frac{1}{2}+\varepsilon}H_0^{\frac{1}{2}+\varepsilon}$.

Moreover, $\mathcal{D}(H_0^{\frac{1}{2}+\varepsilon}) = \mathcal{D}(H_X^{\frac{1}{2}+\varepsilon})$.

Proof: From lemma 8, we know that $\|R_0^{1/2}XR_0^{1/2}\| < 1$, so lemma 8 and its corollary apply. We have that

$$1 > \left\| R_0^{\frac{1}{2}+\varepsilon}XR_0^{\frac{1}{2}-\varepsilon} \right\| = \left\| R_0^{\frac{1}{2}+\varepsilon}(H_X - H_0)R_0^{\frac{1}{2}-\varepsilon} \right\| = \left\| R_0^{\frac{1}{2}+\varepsilon}H_XR_0^{\frac{1}{2}-\varepsilon} - I \right\|,$$
thus \( \| R_0^{1+\varepsilon} H X R_0^{1-\varepsilon} \| < \infty \). We write this as

\[
\left\| R_0^{1+\varepsilon} H X R_0^{1-\varepsilon} \right\| < \infty.
\]

Since \( H_0^{1\pm\varepsilon} R_0^{1\pm\varepsilon} \) is bounded and invertible, so is \( R_0^{1+\varepsilon} H X^+ \). Finally, the fact that \( \left\| R_0^{1+\varepsilon} H X^+ \right\| < \infty \) and \( \left\| H_0^{1+\varepsilon} H_0^{1-\varepsilon} \right\| < \infty \) implies that \( H_0^{1+\varepsilon} \) and \( H_0^{1+\varepsilon} \) are comparable, hence \( D(H_0^{1+\varepsilon}) = D(H_0^{1-\varepsilon}). \)

The next theorem ensures the compatibility between the two charts in the overlapping region \( M_0 \cap M_X \).

**Theorem 11** \( \| \cdot \|_{e}(X) \) and \( \| \cdot \|_{e}(0) \) are equivalent norms.

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

\[
\| Y \|_{e}(X) = \left\| R_0^{1+\varepsilon} H X_0^{1+\varepsilon} R_0^{1-\varepsilon} Y R_0^{1-\varepsilon} H_0^{1-\varepsilon} R_0^{1+\varepsilon} \right\|
\leq \left\| R_0^{1+\varepsilon} H X_0^{1+\varepsilon} \right\| \left\| H_0^{1-\varepsilon} R_0^{1-\varepsilon} \right\| \| Y \|_{e}(0)
= M\| Y \|_{e}(0)
\]

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

\[
\| Y \|_{e}(0) = \left\| R_0^{1+\varepsilon} H X_0^{1+\varepsilon} R_0^{1-\varepsilon} Y R_0^{1-\varepsilon} H_0^{1-\varepsilon} R_0^{1+\varepsilon} \right\|
\leq \left\| R_0^{1+\varepsilon} H X_0^{1+\varepsilon} \right\| \left\| H_0^{1-\varepsilon} R_0^{1-\varepsilon} \right\| \| Y \|_{e}(X)
= m^{-1}\| Y \|_{e}(X). \]

We see that \( T_e(0) \) and \( T_e(X) \) are, in fact, the same Banach space furnished with two equivalent norms, and observe that the quotient spaces \( T_e(0)/\sim \) and \( T_e(X)/\sim \) are exactly the same set. The general theory of Banach manifolds does the rest.

We continue in this way, adding a new patch around another point \( \rho_X \) in \( M_0 \) or around some other point in \( M_X \) but outside \( M_0 \). Whichever point we start from, we get a third piece \( M_X \) with chart into an open subset of the Banach space \( \{ Y \in T_e(X') : \rho_X \cdot Y = 0 \} \), with norm \( \| Y \|_{e}(X') := \left\| R_X^{1+\varepsilon} Y R_X^{1-\varepsilon} \right\| \) equivalent to the previously defined norms. We can go on inductively, and all the norms of any overlapping regions will be equivalent.

**Definition 12** The information manifold \( M(H_0) \) defined by \( H_0 \) consists of all states obtainable in a finite numbers of steps, by extending \( M_0 \) 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 \( \mathcal{D}(H_0^{1-\varepsilon}) \), 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 \( A = \{ \tilde{X} \in \mathcal{T}_\varepsilon(0) : \tilde{X} = X - \rho_0 \cdot X, \|X\|_\varepsilon(0) < 1 - \beta_0 \} \) is a convex subset of the Banach space \( \mathcal{T}_\varepsilon(0) \) 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 \( \tilde{X} \to \rho_X \) and call this the canonical or \((1)-affine\) structure. The \((1)-convex\) mixture of \( \rho_X \) and \( \rho_Y \) in \( \mathcal{M}_0 \) is then \( \rho_{X+(1-\lambda)Y}, (0 \leq \lambda \leq 1) \), which differs from the previously defined \((-1)-convex\) mixture \( \lambda \rho_X + (1-\lambda)\rho_Y \).

Given two points \( \rho_X \) and \( \rho_Y \) in \( \mathcal{M}_0 \) and their tangent spaces \( \mathcal{T}_\varepsilon(X) \) and \( \mathcal{T}_\varepsilon(Y) \), we define the \((1)-parallel\) transport \( U_L \) of \( (Z - \rho_X \cdot Z) \in \mathcal{T}_\varepsilon(X) \) along any continuous path \( L \) connecting \( \rho_X \) and \( \rho_Y \) in the manifold to be the point \( (Z - \rho_Y \cdot Z) \in \mathcal{T}_\varepsilon(Y) \). 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 \( \{Z + \alpha I\} \) 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 \( \mathcal{T}_\varepsilon(X) \). Since both \( \mathcal{T}_\varepsilon(0) \) and \( \mathcal{T}_\varepsilon(X) \) 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 \( \rho_0 \) to \( \rho_X \), it is straightforward to check that \( U \) takes a convex mixture in \( \mathcal{T}_\varepsilon(0) \) to a convex mixture in \( \mathcal{T}_\varepsilon(X) \). So, if \( \rho_Y \in \mathcal{M}_0 \) and \( \rho_{Y'} \in \mathcal{M}_X \) are points outside the overlap, we parallel transport from \( \rho_Y \) 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 hoo 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 \( \Psi : \mathcal{M} \to \mathbb{C} \) given by

\[
\Psi(\rho_X) := \log Z_X.
\]
We say that $Y$ is an $\varepsilon$-bounded direction if $Y \in T_\varepsilon(X)$. We now show that $\Psi(\rho_X)$ is infinitely Fréchet differentiable when the Fréchet derivatives are taken in an $\varepsilon$-bounded direction $\lambda V$ and that, in this case, it has a convergent Taylor series for sufficiently small $\lambda$.

The $n$-th Fréchet derivative of $\Psi_X \equiv \Psi(\rho_X)$ in the $\varepsilon$-bounded directions $V_1, \ldots, V_n$ is given by $(n!)^{-1}$ times the Kubo $n$-point function $[2]$

$$
\text{Tr} \int_0^1 d\alpha_1 \int_0^1 d\alpha_2 \cdots \int_0^1 d\alpha_n \frac{1}{\alpha_n} \rho_X^\alpha V_1 \rho_X^\alpha V_2 \cdots \rho_X^\alpha V_n],
$$

(14)

where $\alpha_n = 1 - \alpha_1 - \cdots - \alpha_{n-1}$.

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

$$
\begin{align*}
&[\rho_X^{\alpha_1} V_1 \rho_X^{\alpha_2} V_2 \cdots \rho_X^{\alpha_n} V_n] \\
&= [\rho_X^{\alpha_1} V_1] [\rho_X^{\alpha_2} V_2] \cdots [\rho_X^{\alpha_n} V_n].
\end{align*}
$$

(15)

$$
\begin{align*}
&\left\| \rho_X^{\alpha_1} \cdots \rho_X^{\alpha_n} \right\|_1 \\
&\leq \left\| \rho_X^{\alpha_1} \right\|_1^{\alpha_1} \cdots \left\| \rho_X^{\alpha_n} \right\|_1^{\alpha_n} \\
&= \left\| \rho_X^{\alpha} \right\|_1 < \infty.
\end{align*}
$$

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

$$
\left\| R_X^\delta V_j R_X^{-\delta} \right\| \leq \left\| R_X^{\frac{\delta}{2} + \varepsilon} V_j R_X^{\frac{\delta}{2} - \varepsilon} \right\| = \| V_j \|_\varepsilon(X) < \infty.
$$

(16)

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

$$
\begin{align*}
&\left\| H_X^{1-\delta_j-x-j} \rho_X^{(1-\beta X)\alpha_j} \right\| \\
&\leq Z_X^{-\alpha_j(1-\beta X)} \sup_{x \geq 1} \left\{ x^{1-\delta_j-x-j} e^{-(1-\beta X)\alpha_j x} \right\} \\
&\leq Z_X^{-\alpha_j(1-\beta X)} \left( \frac{1 - \delta_j-x-j}{1 - (1-\beta X)\alpha_j} \right)^{1-\delta_j-x-j} e^{-(1-\delta_j-x-j)}. \quad (17)
\end{align*}
$$

Apart from $\alpha_j^{-(1-\delta_j-x-j)}$, the other terms in (17) will be bounded independently of $\alpha$. To deal with the integral of $\alpha_j^{-(1-\delta_j-x-j)} d\alpha_j$, we divide the region of integration in $n$ (overlapping) regions $S_j := \{ \alpha : \alpha_j \geq 1/n \}$ (since $\sum \alpha_j = 1$). For
the region $S_n$, for instance, the integrability at $\alpha_j = 0$ is guaranteed if we choose $\delta_j$ such that $\delta_j < \delta_{j-1}$. So we take $\delta_n = \delta_0 > \delta_1 > \cdots > \delta_{n-1}$. We must have $\delta_j \in \left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\right]$, 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 $(\frac{n}{2\varepsilon})^{n-1}$. The last integrand in $S_n$ is $\alpha_n^{-(1-\delta_{n-1}+\delta_n)} \leq n^2$. The same bound holds for the other regions $S_j, j = 1, \ldots, n-1$, giving a total bound
\[ \prod_{j=1}^n \int_0^1 \alpha_j^{-(1-\delta_{j-1}+\delta_j)} d\alpha_j \leq n \left[ \frac{n^2}{(2\varepsilon)^{n-1}} \right] = \frac{n^2}{(2\varepsilon)^{n-1}}. \tag{18} \]

Now that we have fixed $\delta_j$, the promised bound for the other terms in (17) is
\[ \prod_{j=1}^n Z_X^{-\alpha_j(1-\beta_X)} \left( \frac{1 - \delta_{j-1} + \delta_j}{1 - \beta_X} \right)^{1-\delta_{j-1}+\delta_j} \]
\[ \leq 4 Z_X^{-(1-\beta_X)} (1 - \beta_X)^{-n} e^{-n} \tag{19} \]
since $(1 - \delta_{j-1} + \delta_j) < 1$ except for one term, when it is less than 2.

Collecting the estimates (15), (16), (18) and (19), we get the following bound for the $n$-point function
\[ 4 \left\| \rho_X^{(1-\beta_X)} \right\|_1 Z_X^{-(1-\beta_X)} (2\varepsilon)n^2 e^{-n} \prod_{j} \frac{\|V_j\|_e(X)}{2\varepsilon(1 - \beta_X)}. \tag{20} \]

Thus $\Psi_X$ is infinitely Fréchet differentiable in $\varepsilon$-bounded direction and the Taylor series converges if $\|V_j\|_e(X) < (1 - \beta_X)2\varepsilon$. Notice that this is stronger than the condition that $\rho_{V_+X}$ lies in a $\varepsilon$-hood of $\rho_X$.

Finally, let us say that a map $\Phi : U \to C$, on a hood $U$ in $M$, is $(+1)$-analytic in $U$ if it is infinitely often Fréchet differentiable in all $\varepsilon$-bounded directions $AV$ and $\Phi(\rho_X)$ has a convergent Taylor expansion for sufficiently small of $\lambda$. In particular, the $(+1)$-coordinates $\eta_X = \rho \cdot X$ (mixture coordinates) are analytic, since they are derivatives of the free energy $\Psi_X$. 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 8. RFS thanks Rektor Jamiokowski of Torun 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-Verlag, New York, 1985.
[2] H. Araki, *Relative Hamiltonians for Faithful Normal States of a von Neumann Algebra*, Publ. R. I. M. S. (Kyoto), 9, 165-209, 1968.

[3] E. B. Davies, *One-parameter semigroups*, Academic Press, 1980.

[4] P. Gibilescu and G. Pistone, *Connections on nonparametric statistical manifolds by Orlicz space geometry*, Inf. Dim. Analysis, Quant. Prob. and Rel. Top., 1, 325-347, 1998.

[5] H. Hasagawa, *α-Divergence of the Noncommutative Information Geometry*, Rep. Math. Phys., 33, 87-93, 1993.

[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, 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-Verlag, 1966.

[9] S. Lang, *Differential and Riemannian Manifolds*, Springer-Verlag, 1995.

[10] M. K. Murray and J. W. Rice, *Differential Geometry and Statistics, Monographs on Statistics and Applied Probability*, 48, Chapman & Hall, 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, *Geometry of Canonical Correlation on the State Space of a Quantum System*, J. Math. Phys., 35, 780-795, 1994.

[13] D. Petz and C. Sudar, *Geometries of Quantum States*, J. Math. Phys., 37, 2662-2673, 1996.

[14] D. Petz and G. Toth, *The Bogoliubov inner product in quantum statistics*, Lett. Math. Phys., 27, 205-216, 1993.

[15] A. Pietsch, *Nuclear Locally Convex Spaces*, Springer-Verlag, 1972.

[16] G. Pistone and C. Sempi, *An infinite dimensional geometric structure on the space of all probability measures equivalent to a given one*, Ann. Stat., 33, 1543-1561, 1995.

[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*, to appear in the Bogoliubov Memorial Volume, ed. A. A. Slavnov, Steklov Institute.

[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.