

Quantum Microcanonical Entropy of a Pair of Observables

L. Truong

Department of Mathematics, University of California, Berkeley, California, USA

Received May 15, 1974

Abstract. The quantum analogue of the classical theory of the joint microcanonical entropy of a pair of observables is investigated for a system of a large number of identical non-interacting subsystems. It is shown that the quantum joint entropy coincides with the classical joint entropy of an appropriately chosen auxiliary classical system, and known results for classical systems are applied to prove the equivalence of the quantum microcanonical and quantum canonical ensembles.

1. Introduction and Statement of Results

The theory of the joint microcanonical entropy of two (or more) classical observables has been developed by D. Ruelle and O. Lanford III. This theory becomes particularly simple for systems consisting of a large number of non-interacting subsystems; indeed, it is shown in Section A.4 of [1] that what is involved in this case is a question about sums of independent identically distributed random variables. In this paper, the quantum analogues of these questions will be investigated. That is, the joint quantum microcanonical entropy of a pair of observables will be defined and its asymptotic behavior as the system in question becomes infinitely large will be investigated. The main result is that the quantum microcanonical entropy turns out to be identical with the classical microcanonical entropy of an appropriately chosen auxiliary “classical” system; the results of Section A.4 of [1] are then applied to give rather complete information about its behavior.

We will begin by describing in detail the sort of non-interacting classical systems considered here. Let Ω be a finite set, equipped with a measure M . (Ω is normally interpreted as the set of states of some atom or molecule, and the measure of a subset of Ω is taken to be the number of elements in that subset. For our purposes, however, it is essential to allow more general choices of M .) Let f_1, f_2 be two real-valued functions on Ω , and regard the pair $f = (f_1, f_2)$ as a mapping of Ω into \mathbb{R}^2 . (f_1 and f_2 can be interpreted as observables on the state space Ω , and of particular interest is the case where f_1 is the energy.) Let Ω^k be a product of k copies

of Ω , and let M^k be the product of k copies of M ; Ω^k is of course interpreted as the state space for a system of k identical atoms. Now define

$$\mathcal{V}(k, f, J) = M^k \{(\omega_1, \dots, \omega_k) : \frac{1}{k} [f(\omega_1) + \dots + f(\omega_k)] \in J\},$$

where J denotes an open convex subset of \mathbb{R}^2 .

It is desirable to know the behavior of $\mathcal{V}(k, f, J)$ for large k . The main facts are as follows (see A.4 of [1]):

1. $\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}(k, f, J)$ exists and is equal to $\sup_k \frac{1}{k} \log \mathcal{V}(k, f, J)$. This limit (which may be a real number or $-\infty$) is denoted by $S(f, J)$, and is called the microcanonical entropy.

2. If $x \in \mathbb{R}^2$, define

$$s(f, x) = \inf \{s(f, J) : J \text{ is an open convex set containing } x\}.$$

So defined, $s(f, x)$ is a concave upper semi-continuous function from \mathbb{R}^2 to $\mathbb{R} \cup \{-\infty\}$. If J is an open convex set in \mathbb{R}^2 ,

$$s(f, J) = \sup_{x \in J} s(f, x).$$

3. $\overline{\{x : s(f, x) > -\infty\}} = \overline{\text{conv}(f(\Omega))}$ ($= \text{conv}(f(\Omega))$, since Ω is finite). If the components f_1, f_2 of f are linearly independent modulo the constants (i.e., if f_1, f_2 , and 1 are linearly independent functions on Ω) then

$$\text{int}(\text{conv}(f(\Omega))) \neq \emptyset.$$

4. On $\text{int}(\text{conv}(f(\Omega)))$, $s(f, x)$ is analytic and strictly concave in x . In fact:

$$s(f, x) = x \cdot \gamma_x + \log \left[\int_{\Omega} \exp[-\gamma_x \cdot f(\omega)] M(d\omega) \right]$$

where γ_x is the unique solution of the equation

$$x = \frac{\int_{\Omega} f(\omega) e^{-\gamma \cdot f(\omega)} M(d\omega)}{\int_{\Omega} e^{-\gamma \cdot f(\omega)} M(d\omega)}.$$

(In the above, γ like x denotes an element of \mathbb{R}^2 and $\gamma \cdot x = \gamma_1 x_1 + \gamma_2 x_2$, etc.)

5. Suppose ε is an element in the interior of the convex hull of the range of f_1 , let $\bar{f}_2(\varepsilon)$ denote the unique value of x such that

$$S(f, (\varepsilon, x)) = \sup_y S(f, (\varepsilon, y)).$$

If β is chosen so that

$$\varepsilon = \frac{\int e^{-\beta f_1(\omega)} f_1(\omega) d\omega}{\int e^{-\beta f_1(\omega)} d\omega},$$

then

$$\bar{f}_2(\varepsilon) = \frac{\int e^{-\beta f_1(\omega)} f_2(\omega) d\omega}{\int e^{-\beta f_1(\omega)} d\omega}.$$

Let us next turn to the formulation of the quantum analogues of the systems described above. Let \mathcal{H} be an n -dimensional Hilbert space (whose unit vectors can be thought of as describing the pure quantum states of some atom or molecule). For two self-adjoint operators H, B , and for $k \in \mathbb{Z}^+$, let \mathcal{H}_k be the tensor product of k -copies of \mathcal{H} . For $1 \leq i \leq k$, let (H_i) and (B_i) be the operators

$$1 \otimes \cdots \otimes 1 \otimes \underset{\text{ith place}}{H} \otimes 1 \otimes \cdots \otimes 1$$

and

$$1 \otimes \cdots \otimes 1 \otimes \underset{\text{ith place}}{B} \otimes 1 \otimes \cdots \otimes 1$$

respectively, where 1 is the identity operator on \mathcal{H} . For $k \in \mathbb{Z}^+$, define

$$H_k = \frac{1}{k} \sum_{i=1}^k (H_i) \quad \text{and} \quad B_k = \frac{1}{k} \sum_{i=1}^k (B_i).$$

Clearly both H_k, B_k are self-adjoint operators on \mathcal{H}_k . Schematically, \mathcal{H} may be thought of as the analogue of Ω , the trace of an operator on \mathcal{H} as the analogue of the integral of a function on Ω with respect to M , and H, B as the analogues of f_1, f_2 . Next it must be decided what to choose as the analogue of the classical \mathcal{V} ; heuristically, this should be some sort of “quantum joint distribution” of H_k and B_k . However, since the operators H, B need not commute, the definition of this “quantum joint distribution” is not obvious, but the following choice is reasonably satisfactory from the physical point of view:

Definition 1.1. Let H, B be self-adjoint operators on \mathcal{H} , J be an open subset of \mathbb{R}^2 , and k be an element of \mathbb{Z}^+ . Define

$$\mathcal{V}_q(k, H, B, J) = \sum_{(\bar{\lambda}, \bar{\mu}) \in J} \text{tr}(P(k, \bar{\lambda}) Q(k, \bar{\mu}))$$

where $P(k, \bar{\lambda}) [Q(k, \bar{\mu})]$ denotes the eigenprojection of $H_k [B_k]$ belonging to the eigenvalue $\bar{\lambda} [\bar{\mu}]$. As in the classical case, we are interested in analyzing the asymptotic behavior of $\mathcal{V}_q(k, H, B, J)$ as $k \rightarrow \infty$. It will be shown that if H and B are given, $\Omega_{H,B}$ and M can be determined in terms of H and B , and two real-valued functions f_H, f_B can be defined on $\Omega_{H,B}$, so that for any $k \in \mathbb{Z}^+$ and any J ,

$$\mathcal{V}_q(k, H, B, J) = M^k \{(\omega_1, \dots, \omega_k) : \frac{1}{k} [f(\omega_1) + \dots + f(\omega_k)] \in J\},$$

or $\mathcal{V}_q(k, H, B, J) = \mathcal{V}(k, f, J)$ where $f = (f_H, f_B)$. It will be shown that if H, B , and I are linearly independent, then f_H, f_B are linearly independent

modulo constants on $\Omega_{H,B}$. Thus studying the asymptotic behavior of $\mathcal{V}_q(k, H, B, J)$ is reduced to studying the asymptotic behavior of $\mathcal{V}(k, f, J)$, which is already known from classical statistical mechanics. Therefore facts 1–4 stated above can be applied to $\mathcal{V}_q(k, H, B, J)$. In particular, $\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J)$ exists and will be denoted by $S_q(H, B, J)$ (the quantum microcanonical entropy). If $x \in \mathbb{R}^2$, define $S_q(H, B, x) = \inf \{S_q(H, B, J) : J \text{ is an open convex set containing } x\}$.

Now consider what physical consequences can be deduced from the above results. Let ε lie between the largest and smallest eigenvalues of H ; ε is interpreted as the average energy per subsystem. Holding ε fixed, and regarding $S_q(H, B, (\varepsilon, y))$ as a function of y alone, a function is obtained which is upper semi-continuous and strictly concave on the interval where it is finite. By upper semi-continuity and the fact that the function is different from $-\infty$ only on a finite interval, it must assume its maximum at least once; strict concavity implies that it assumes its maximum exactly once. Let $\bar{B}(\varepsilon)$ denote the unique value of y where the maximum is assumed, i.e. $\bar{B}(\varepsilon)$ is the unique solution of

$$S_q(H, B, (\varepsilon, \bar{B}(\varepsilon))) = \sup_r S_q(H, B, (\varepsilon, y)).$$

Roughly speaking, $\bar{B}(\varepsilon)$ should be interpreted as the overwhelmingly probable value of B_k , given that H_k is near ε , for large k . Since H_k and B_k are operators, this statement requires some explanation.

To give this explanation, the *quantum microcanonical ensemble* is introduced. If I is an interval in \mathbb{R} , let

$$P(k, I) = \sum_{\bar{\lambda} \in I} P(k, \bar{\lambda}),$$

$$Q(k, I) = \sum_{\bar{\mu} \in I} Q(k, \bar{\mu});$$

also, let $L(\mathcal{H}_k)$ be the algebra of all linear operators on \mathcal{H}_k .

Definition 1.2. Given ε , and given $\delta > 0$, define the *quantum microcanonical ensemble* to be the state $\varrho_{k,\varepsilon,\delta}^{\text{m.c.}}$ on $L(\mathcal{H}_k)$ given by

$$\varrho^{\text{m.c.}}(T) = \frac{\text{Tr}(T \cdot P(k, (\varepsilon - \delta, \varepsilon)))}{\text{Tr} P(k, (\varepsilon - \delta, \varepsilon))}.$$

(The parameter δ is an unfortunate technical necessity; it should be thought of as very small and will be allowed to go to zero whenever possible.)

The probabilistic interpretation of quantum mechanics indicates how to associate with $\varrho^{\text{m.c.}}$ and any self-adjoint $T \in L(\mathcal{H}_k)$ a probability measure on the spectrum of T : the probability of any subset of the spectrum is $\varrho^{\text{m.c.}}$ of the corresponding spectral projection. This con-

struction allows one to speak of the microcanonical probability that T belongs to some set, etc. Now apply this terminology to $T = B_k$. The result is that if I is any open interval whose closure does not contain $\bar{B}(\varepsilon)$, and if δ is small enough, then the microcanonical probability that B_k belongs to I converges to zero exponentially as $k \rightarrow \infty$. Note that this assertion characterizes $\bar{B}(\varepsilon)$ uniquely and justifies its interpretation as the most probable value of B_k given that $H_k \approx \varepsilon$.

To prove the above result, note that the microcanonical probability that B_k is in I is just

$$\frac{\text{Tr}(P(k, I_\varepsilon) Q(k, I))}{\text{Tr}(P(k, I_\varepsilon))} = \frac{\mathcal{V}_q(k, H, B, I_\varepsilon \times I)}{\mathcal{V}_q(k, H, B, I_\varepsilon \times \mathbb{R})},$$

where I_ε denotes $(\varepsilon - \delta, \varepsilon)$. For large k ,

$$\frac{\mathcal{V}_q(k, H, B, I_\varepsilon \times I)}{\mathcal{V}_q(k, H, B, I_\varepsilon \times \mathbb{R})} \approx \exp \{ + k [S_q(H, B, I_\varepsilon \times I) - S_q(H, B, I_\varepsilon \times \mathbb{R})] \}.$$

The fact that $\bar{B}(\varepsilon) \notin \bar{I}$ implies that, for sufficiently small δ ,

$$S_q(H, B, I_\varepsilon \times \mathbb{R}) \text{ is strictly larger than } S_q(H, B, I_\varepsilon \times I)$$

i.e., the probability behaves asymptotically for large k like $e^{-k \cdot \delta s}$, with $\delta s > 0$. An easy extension of this argument shows that $\bar{B}(\varepsilon)$ is also the large k limit of the microcanonical mean value of B_k , i.e. that

$$\bar{B}(\varepsilon) = \lim_{\delta \rightarrow 0} \lim_{k \rightarrow \infty} \varrho_{k, \varepsilon, \delta}^{\text{m.c.}}(B_k) \text{ (see Proposition 2.10).}$$

To find a simpler way of computing $\bar{B}(\varepsilon)$, introduce an auxiliary parameter β , interpreted as the inverse temperature. Then if ε, β are related by

$$\varepsilon = \frac{\text{Tr}(H e^{-\beta H})}{\text{Tr}(e^{-\beta H})},$$

we have

$$\bar{B}(\varepsilon) = \frac{\text{Tr}(B e^{-\beta H})}{\text{Tr}(e^{-\beta H})}.$$

This formula is referred to as the *canonical prescription* for computing $\bar{B}(\varepsilon)$.

Remark. The reasons for considering only two observables (as opposed to n) are that:

(a) In physical applications, it is usually enough to have two observables; one may be thought to be the energy of the system, and the other any observable.

(b) There are technical difficulties in considering more than two observables. For example, \mathcal{V}_q is defined to be the trace of a product of projections, but the trace of the product of three or more self-adjoint

operators may not yield a unique value, if the order in which the operators are multiplied is permuted. Furthermore, the trace of three or more self-adjoint operators might not even be a real number. However, $\text{tr}(HB)$ is independent of the order in which H and B are multiplied together, and, if H and B are positive, $\text{tr}(HB)$ is positive.

In Section 2, the results stated in this section will be proved.

I am grateful to Oscar E. Lanford III for his numerous helpful suggestions, in particular for his extensive assistance with revision of the wording of the introduction.

2. Proof of Results

In this section we shall prove the existence and analyze the properties of $\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J)$ where H, B are self-adjoint operators on an n -dimensional Hilbert space H , J is an open convex set of \mathbb{R}^2 , and $\mathcal{V}_q(k, H, B, J) = \sum_{(\bar{\lambda}, \bar{\mu}) \in J} \text{tr}(P(k, \bar{\lambda}) Q(k, \bar{\mu}))$. (See Definition 1.1.) Then the properties of $\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J)$ will be applied to obtain the results on the equivalence of ensembles for observables stated in Section 1.

Lemma 2.1. (a) $\text{Spec } H_k = \left\{ \bar{\lambda} : \bar{\lambda} = \frac{1}{k} \sum_{i=1}^k \lambda_i, \lambda_i \in \text{Spec } H, 1 \leq i \leq k \right\}$.

(b) If $\bar{\lambda} \in \text{Spec } H_k$, then the projection $P(k, \bar{\lambda})$ of H_k onto the eigensubspace of H_k with eigenvalue $\bar{\lambda}$ is given by

$$P(k, \bar{\lambda}) = \sum \left\{ P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_k} : \lambda_1, \dots, \lambda_k \in \text{Spec } H \text{ such that } \bar{\lambda} = \frac{1}{k} \sum_{i=1}^k \lambda_i \right\}.$$

Similar statements hold for B_k .

Proof. Statements (a) and (b) will be proved together. Let $\lambda_1, \dots, \lambda_k$ be eigenvalues of H (repetition is allowed), and let h_1, \dots, h_k be the respective eigenvectors. Then by applying the definition of H_k ,

$$\begin{aligned} H_k(h_1 \otimes \cdots \otimes h_k) &= \frac{1}{k} \left(\sum_{i=1}^k \lambda_i \right) h_1 \otimes \cdots \otimes h_k; \\ \left\{ \bar{\lambda} : \bar{\lambda} = \frac{1}{k} \left(\sum_{i=1}^k \lambda_i \right) \text{ where } \lambda_i \in \text{Spec } H \right\} &\subset \text{Spec } H_k. \end{aligned} \tag{1}$$

It is clear that if $(\lambda_1, \dots, \lambda_k) \in (\text{Spec } H)^k$ then $P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_k}$ is a projection of \mathcal{H}_k . If $(\lambda_1, \dots, \lambda_k)$ and $(\lambda'_1, \dots, \lambda'_k)$ are two distinct elements of $(\text{Spec } H)^k$, then $P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_k}$ and $P_{\lambda'_1} \otimes \cdots \otimes P_{\lambda'_k}$ are mutually orthogonal.

To see this suppose $\lambda_i \neq \lambda'_i$, then $P_{\lambda_i} \cdot P_{\lambda'_i} = 0$, so

$$\begin{aligned} & (P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_i} \otimes \cdots \otimes P_{\lambda_k}) \cdot (P_{\lambda'_1} \otimes \cdots \otimes P_{\lambda'_i} \otimes \cdots \otimes P_{\lambda'_k}) \\ &= P_{\lambda_1} P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_i} P_{\lambda'_i} \otimes \cdots \otimes P_{\lambda_k} P_{\lambda'_k} \\ &= 0. \end{aligned}$$

Now suppose $\bar{\lambda} \in \text{Spec } H_k$, then: (i) If $\bar{\lambda}$ is expressible as $\frac{1}{k} \sum_{i=1}^k \lambda_i$, where $\lambda_i \in \text{Spec } H$, let $\hat{P}(k, \bar{\lambda}) = \sum P_{\lambda_1} \otimes \cdots \otimes P_{\lambda_k}$, where the sum is taken over all $\lambda_1, \dots, \lambda_k$ such that $\bar{\lambda} = \frac{1}{k} \left(\sum_{i=1}^k \lambda_i \right)$; and (ii) If $\bar{\lambda}$ is not expressible as $\frac{1}{k} \sum_{i=1}^k \lambda_i$, then let $\hat{P}(k, \bar{\lambda}) = 0$ (the zero projection). A straight-forward computation will show that $I_k = \sum \hat{P}(k, \bar{\lambda}) \leq \sum P(k, \bar{\lambda}) = I_k$, which implies $\hat{P}(k, \bar{\lambda}) = P(k, \bar{\lambda})$ for all $\bar{\lambda}$, which implies both (a) and (b).

Let $\Omega = \{(\lambda, \mu) : (\lambda, \mu) \in \text{Spec } H \times \text{Spec } B, \text{tr}(P_\lambda Q_\mu) \neq 0\}$. Let $f_H, f_B : \Omega \rightarrow \mathbb{R}$ be defined by

$$f_H(\lambda, \mu) = \lambda, \quad f_B(\lambda, \mu) = \mu,$$

and $f = (f_H, f_B)$, i.e., f is the identity function on Ω . Let M be a measure on Ω defined by

$$M\{(\lambda, \mu)\} = \text{tr}(P_\lambda Q_\mu).$$

Note that: (i) $\text{tr}(P_\lambda Q_\mu) \geq 0$, since $P_\lambda Q_\mu$ is positive and (ii) $M(\Omega) = n$, independent of H, B , since

$$\begin{aligned} M(\Omega) &= \sum_{(\lambda, \mu) \in \Omega} \text{tr}(P_\lambda Q_\mu) \\ &= \sum_{(\lambda, \mu) \in \text{Spec } H \times \text{Spec } B} \text{tr}(P_\lambda Q_\mu) \\ &= \text{tr} \left(\sum_{(\lambda, \mu) \in \text{Spec } H \times \text{Spec } B} P_\lambda Q_\mu \right) \\ &= \text{tr } I \\ &= n. \end{aligned}$$

The object of introducing the above definitions is to show that $\mathcal{V}_q(k, H, B, J)$ can be expressed as

$$M^k \left\{ (\omega_1, \dots, \omega_k) \in \Omega^k : \frac{1}{k} \sum_{i=1}^k f(\omega_i) \in J \right\}.$$

Assuming that this expression is proved, the techniques of classical statistical mechanics¹ can be applied to analyze

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J).$$

¹ See Lanford, A.4 of [1].

Proposition 2.2. *Let $J \subset \mathbb{R}^2$ be open and convex, Ω^k be the Cartesian product of k copies of Ω , and M^k the product measure induced by M on Ω^k . Then*

$$\mathcal{V}_q(k, H, B, J) = M^k \left\{ (\omega_1, \dots, \omega_k) \in \Omega^k : \frac{1}{k} \sum_{i=1}^k f(\omega_i) \in J \right\}.$$

Proof. By definition

$$\mathcal{V}_q(k, H, B, J) = \sum_{(\bar{\lambda}, \bar{\mu}) \in J} \text{tr}(P(k, \bar{\lambda}) Q(k, \bar{\mu})).$$

This expression can be rewritten as

$$\Sigma \{ \text{tr}((P_{\lambda_1} \otimes \dots \otimes P_{\lambda_k})(Q_{\mu_1} \otimes \dots \otimes Q_{\mu_k})) : (\lambda_i, \mu_i) \in \text{Spec } H \times \text{Spec } B$$

where

$$\bar{\lambda} = \frac{1}{k} \sum_{i=1}^k \lambda_i, \quad \bar{\mu} = \frac{1}{k} \sum_{i=1}^k \mu_i \quad \text{and} \quad (\bar{\lambda}, \bar{\mu}) \in J \},$$

by Lemma 2.1.

Evidently only those k -tuples of pairs which are elements of Ω^k will contribute to the sum. For if $(\lambda_i, \mu_i) \notin \Omega$ then

$$\text{tr}(P_{\lambda_i} Q_{\mu_i}) = 0,$$

and

$$\text{tr}((P_{\lambda_1} \otimes \dots \otimes P_{\lambda_k})(Q_{\mu_1} \otimes \dots \otimes Q_{\mu_k})) = \prod_{i=1}^k \text{tr}(P_{\lambda_i} Q_{\mu_i}).$$

Let $\omega_i = (\lambda_i, \mu_i)$ and note that $\left(\frac{1}{k} \sum_{i=1}^k \lambda_i, \frac{1}{k} \sum_{i=1}^k \mu_i \right) = \frac{1}{k} \sum_{i=1}^k f(\omega_i)$. Thus $\mathcal{V}_q(k, H, B, J)$ can be rewritten as

$$\Sigma \left\{ \prod_{i=1}^k M(\omega_i) : (\omega_1, \dots, \omega_k) \in \Omega^k \quad \text{and} \quad \frac{1}{k} \sum_{i=1}^k f(\omega_i) \in J \right\}$$

which is precisely the expression

$$M^k \left\{ (\omega_1, \dots, \omega_k) \in \Omega^k : \frac{1}{k} \sum_{i=1}^k f(\omega_i) \in J \right\}.$$

Thus the proposition is proved.

Having the above proposition, all desired properties of $\mathcal{V}_q(k, H, B, J)$ can be deduced.

Proposition 2.3. *If H, B and J are fixed, then*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J) = \sup_k \frac{1}{k} \log \mathcal{V}_q(k, H, B, J).$$

Proof. The proposition follows immediately from Proposition 2.2 and fact 1 of Section 1.

Let $S_q(H, B, J) = \lim_{k \rightarrow \infty} \frac{1}{k} \log \mathcal{V}_q(k, H, B, J)$. If $x \in \mathbb{R}^2$, define

$$S_q(H, B, x) = \inf_J S_q(H, B, J), \tag{2}$$

where J runs over the convex open sets containing x .

Then

$$\begin{aligned} S_q(H, B, J) &= S(f, J), \\ S_q(H, B, x) &= S(f, x). \end{aligned} \tag{3}$$

These two equations in (3) follow immediately from the definitions of $S_q(H, B, J)$, $S_q(H, B, x)$, $S(f, J)$, $S(f, x)$ and Proposition 2.2.

The following proposition is an immediate consequence of the second equation of (3) and fact 2 of Section 1.

Proposition 2.4. $S_q(H, B, x)$ is a concave upper semi-continuous function from \mathbb{R}^2 to $\mathbb{R} \cup \{-\infty\}$.

Next it is of interest to determine for what region of \mathbb{R}^2 , $S_q(H, B, x) > -\infty$ and a condition on H and B so that the aforesaid set is not empty. Before the condition can be stated, the next two lemmas are needed.

Lemma 2.5. If P and Q are projections of \mathcal{H} such that $\text{tr}(PQ) = \text{tr} P = \text{tr} Q$, then $P = Q$.

Proof. $\text{tr}((P - Q)^2) = \text{tr} P - 2\text{tr}(PQ) + \text{tr} Q = 0$, by hypothesis. Since $(P - Q)^2$ is positive then $(P - Q)^2 = 0$ and so $P = Q$. Thus the lemma is proved.

Lemma 2.6. If H, B and I are linearly independent on \mathcal{H} then $\text{conv}(\Omega)$ has nonempty interior. (The converse is also true.)

Proof. It suffices to prove that Ω is not a subset of a straight line. For any $\lambda \in \text{Spec} H$ there exists a $\mu \in \text{Spec} B$ such that $(\lambda, \mu) \in \Omega$ since $0 < \text{tr} P_\lambda = \sum \{\text{tr}(P_\lambda Q_\mu) : \mu \in \text{Spec} B\}$. Thus there exists a μ such that $\text{tr}(P_\lambda Q_\mu) > 0$. Similarly the above assertion is valid with μ and λ interchanged. The proof of this lemma is by contradiction. If there exist constants a, b, c such that $a\lambda + b\mu = c$ for any pair $(\lambda, \mu) \in \Omega$, the only non-trivial case is where a and b are both nonzero. For if either a or $b = 0$ then respectively B or H is a dilation. Obviously in these two cases H, B , and I are linearly dependent. Thus the only case of interest is when Ω is a subset of a line $ax + by = c$ which is not parallel to the x or y axis. In this case, for each $\lambda \in \text{Spec} H$ there is a unique $\mu \in \text{Spec} B$ such that $(\lambda, \mu) \in \Omega$. The same assertion holds when λ and μ are interchanged, since the line is not parallel to either axis.

Thus, by the above assertion,

$$\Omega = \{(\lambda_1, \mu_1), \dots, (\lambda_l, \mu_l)\}$$

where

$$\{\lambda_1, \dots, \lambda_l\} = \text{Spec} H \quad \text{and} \quad \{\mu_1, \dots, \mu_l\} = \text{Spec} B.$$

From the definition of Ω ,

$$\begin{aligned} \operatorname{tr}(P_{\lambda_i} Q_{\mu_j}) &= 0 \quad \text{if } i \neq j \\ &> 0 \quad \text{if } i = j. \\ \operatorname{tr} P_{\lambda_i} &= \Sigma \{ \operatorname{tr}(P_{\lambda_i} Q_{\mu_j}) : \mu_j \in \operatorname{Spec} B \} \\ &= \operatorname{tr}(P_{\lambda_i} Q_{\mu_i}). \end{aligned}$$

Similarly

$$\operatorname{tr} Q_{\mu_i} = \operatorname{tr}(P_{\lambda_i} Q_{\mu_i}),$$

and from Lemma 2.5,

$$P_{\lambda_i} = Q_{\mu_i}, \quad 1 \leq i \leq l.$$

Thus

$$\begin{aligned} aH + bB &\dots \sum_{i=1}^l (a\lambda_i P_{\lambda_i} + b\mu_i Q_{\mu_i}) \\ &= \sum_{i=1}^l (a\lambda_i + b\mu_i) P_{\lambda_i}, \end{aligned}$$

but $a\lambda_i + b\mu_i = c$ by assumption. Therefore $aH + bB = cI$, which contradicts the hypothesis that H, B , and I are linearly independent.

Proposition 2.7. (a) $\{x: S_q(H, B, x) > -\infty\} = \overline{\operatorname{conv} f(\Omega)}$, and (b) if H, B, I are linearly independent then $\{x: S_q(H, B, x) > -\infty\}$ has non-empty interior.

Note that since $f(\Omega)$ is finite and $f(\Omega) \subset \mathbb{R}^2$, $\overline{(\operatorname{conv} f(\Omega))} = \operatorname{conv} f(\Omega)$.

Proof. (a) follows from the definition of $\Omega, f, (3)$, the first half of fact 3, Section 1 and the fact that $f(\Omega) = \Omega$; and (b) is an immediate consequence of (a), Lemma 2.6, and the second half of fact 3, Section 1.

The analyticity and strict concavity of $S_q(H, B, x)$ on $\operatorname{int} \operatorname{conv} \Omega$ follow directly from Proposition 2.7, fact 4 of Section 1 and the following observations:

If $\gamma \in \mathbb{R}^2$, define

$$\begin{aligned} Z(\gamma, H, B) &= \operatorname{tr}(e^{-\gamma_1 H} e^{-\gamma_2 B}) \\ &= \sum_{(\lambda, \mu) \in \operatorname{Spec} H \times \operatorname{Spec} B} e^{-(\gamma_1 \lambda + \gamma_2 \mu)} \operatorname{tr}(P_{\lambda} Q_{\mu}) \\ &= \sum_{(\lambda, \mu) \in \Omega} e^{-(\gamma_1 \lambda + \gamma_2 \mu)} \operatorname{tr}(P_{\lambda} Q_{\mu}). \end{aligned}$$

By definition, $(\lambda, \mu) = f(\lambda, \mu)$ and $\operatorname{tr}(P_{\lambda} Q_{\mu}) = M\{(\lambda, \mu)\}$. For simplicity, denote $\gamma_1 \lambda + \gamma_2 \mu$ by $\gamma \cdot f(\lambda, \mu)$. Thus one can write

$$Z(\gamma, H, B) = \int_{\Omega} e^{-\gamma \cdot f(\omega)} dM(\omega). \quad (4)$$

Let M_{γ} be the probability measure on Ω defined as follows: if $\omega \in \Omega$, then

$$M_{\gamma}(\{\omega\}) = \frac{e^{-\gamma \cdot f(\omega)} M(\{\omega\})}{Z(\gamma, H, B)}.$$

Theorem 2.8. *Suppose H, B, I are linearly independent. Then on $\text{int conv } \Omega$, $S_q(H, B, x)$ is analytic and strictly concave in x . Furthermore*

$$S_q(H, B, x) = \log Z(\gamma(x), H, B) + x \cdot \gamma(x),$$

where $x = \int_{\Omega} f(\omega) dM\gamma$.

Since $S_q(H, B, x)$ is strictly concave in x , the quantity $S_q(H, B, (\varepsilon, y))$, as a function of y with ε held fixed, takes on its maximum at a single point², which is denoted by $\bar{B}(\varepsilon)$.

Proposition 2.9. *If β, ε are chosen so that $\varepsilon = \frac{\text{tr}(He^{-\beta H})}{\text{tr} e^{-\beta H}}$, then*

$$\bar{B}(\varepsilon) = \frac{\text{tr}(Be^{-\beta H})}{\text{tr} e^{-\beta H}}.$$

Proof. By a straight-forward computation,

$$\frac{\text{tr}(He^{-\beta H})}{\text{tr} e^{-\beta H}} = \frac{\int_{\Omega} e^{-\beta f_H(\omega)} f_H(\omega) dM(\omega)}{\int_{\Omega} e^{-\beta f_H(\omega)} dM(\omega)}$$

and

$$\frac{\text{tr}(Be^{-\beta H})}{\text{tr} e^{-\beta H}} = \frac{\int_{\Omega} e^{-\beta f_H(\omega)} f_B(\omega) dM(\omega)}{\int_{\Omega} e^{-\beta f_H(\omega)} dM(\omega)}.$$

Then applying fact 5 of Section 1, the desired result is obtained.

Proposition 2.10. *For any ε in the interior of the convex hull of the spectrum of H ,*

$$\bar{B}(\varepsilon) = \lim_{\delta \downarrow 0} \lim_{k \rightarrow \infty} \frac{\text{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\text{tr}[P(k, (\varepsilon - \delta, \varepsilon))]}.$$

Proof. Let $\delta > 0$. Then there exists a unique pair ε_{δ}, B' with $\varepsilon - \delta \leq \varepsilon_{\delta} \leq \varepsilon$, such that

$$\sup \{S_q(H, B, \alpha, \beta) : \varepsilon - \delta < \alpha < \varepsilon\} = S_q(H, B, \varepsilon_{\delta}, B').$$

The existence of ε_{δ}, B' follows from the fact that $S_q(H, B, \alpha, \beta)$ is upper semi-continuous in α, β and equal to $-\infty$ for α, β outside $\text{conv}(\Omega)$; the uniqueness follows from the strict concavity of S_q . Clearly,

$$S_q(H, B, \varepsilon_{\delta}, B') = \sup_y S_q(H, B, \varepsilon_{\delta}, y),$$

so

$$B' = \bar{B}(\varepsilon_{\delta}).$$

Now, as $\delta = 0$, ε_{δ} converges to ε (since $\varepsilon - \delta \leq \varepsilon_{\delta} \leq \varepsilon$), and Proposition 2.9 implies that $\bar{B}(\varepsilon)$ is continuous in ε ; hence, it is only necessarily to show

² Provided $\varepsilon \in \text{int conv}(\text{Spec}(H))$.

that

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} = B'. \quad (5)$$

Note first that

$$\|B_k\| \leq \|B\|$$

for all k , and write the spectral resolution for B_k as

$$B_k = \sum_{\bar{\mu}} \bar{\mu} Q(k, \bar{\mu}).$$

Now let $a > 0$. Then

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \log \operatorname{tr}[Q(k, (-\infty, B' - a)) P(k, (\varepsilon - \delta, \varepsilon))] \\ &= \sup \{S_q(H, B, \alpha, \beta) : \varepsilon - \delta < \alpha < \varepsilon; \beta < B' - a\} \\ &< \sup \{S_q(H, B, \alpha, \beta) : \varepsilon - \delta < \alpha < \varepsilon\} \\ &= \lim_{k \rightarrow \infty} \frac{1}{k} \log \operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]. \end{aligned}$$

Thus,

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[Q(k, (-\infty, B' - a)) P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} = 0,$$

and since $Q(k, (-\infty, B' - a)) + Q(k, [B' - a, \infty)) = 1$,

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[Q(k, [B' - a, \infty)) P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} = 1.$$

Now

$$\begin{aligned} \frac{\operatorname{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} &\geq (B' - a) \frac{\operatorname{tr}[Q(k, [B' - a, \infty)) P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} \\ &\quad - \|B_k\| \frac{\operatorname{tr}[Q(k, (-\infty, B' - a)) P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} . \end{aligned}$$

So

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} \geq B' - a.$$

Since a is any strictly positive number

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} \geq B'.$$

A similar argument show that

$$\lim_{k \rightarrow \infty} \frac{\operatorname{tr}[B_k P(k, (\varepsilon - \delta, \varepsilon))]}{\operatorname{tr}[P(k, (\varepsilon - \delta, \varepsilon))]} \leq B';$$

these two inequalities together imply (5).

Remark. Suppose that H and B commute. Then they can be simultaneously diagonalized, and Ω is the set $\{(\lambda_1, \mu_1), \dots, (\lambda_j, \mu_j)\}$ of pairs of simultaneous eigenvalues. The mapping $H \mapsto f_H, B \mapsto f_B$ extends to an isomorphism from the algebra of matrices generated by H and B and the algebra of all complex-valued functions on Ω . In this special case, then, the quantum theory reduces completely to the classical theory.

References

1. Lanford III, O.E.: Entropy and equilibrium states in classical statistical mechanics. In: Lenard, A. (ed.): Statistical mechanics and mathematical problems, Battelle Seattle 1971 Rencontres, pp. 1–113. Berlin-Heidelberg-New York: Springer 1973

Communicated by G. Gallavotti

Lawrence Truong
University of California
Department of Mathematics
Berkeley, California 94720, USA

