JOURNAL OF MATHEMATICAL PHYSICS **47**, 073303 (2006)

Quantum macrostates, equivalence of ensembles, and an *H*-theorem

Wojciech De Roeck[a] and Christian Maes[b]
Instituut voor Theoretische Fysica, K.U.Leuven, Leuven, 3001 Belgium

Karel Netočný[c]
Institute of Physics AS CR, Prague, 182 21 Czech Republic

(Received 12 January 2006; accepted 2 June 2006; published online 31 July 2006)

Before the thermodynamic limit, macroscopic averages need not commute for a quantum system. As a consequence, aspects of macroscopic fluctuations or of constrained equilibrium require a careful analysis, when dealing with several observables. We propose an implementation of ideas that go back to John von Neumann's writing about the macroscopic measurement. We apply our scheme to the relation between macroscopic autonomy and an *H*-theorem, and to the problem of equivalence of ensembles. In particular, we show how the latter is related to the asymptotic equipartition theorem. The main point of departure is an expression of a law of large numbers for a sequence of states that start to concentrate, as the size of the system gets larger, on the macroscopic values for the different macroscopic observables. Deviations from that law are governed by the entropy. *© 2006 American Institute of Physics.* [DOI: 10.1063/1.2217810]

I. INTRODUCTION

"It is a fundamental fact with macroscopic measurements that everything which is measurable at all, is also simultaneously measurable, i.e. that all questions which can be answered separately can also be answered simultaneously." That statement by von Neumann enters his introduction to the macroscopic measurement.[16] He then continues to discuss in more detail how that view could possibly be reconciled with the non-simultaneous-measurability of quantum mechanical quantities. The main qualitative suggestion by von Neumann is to consider, for a set of noncommuting operators $A,B,\ldots$ a corresponding set of *mutually commuting* operators $A',B',\ldots$ which are each, in a sense, good approximations, $A'\approx A, B'\approx B,\ldots$. The whole question is: *in exactly what sense?* Especially in statistical mechanics, one is interested in fluctuations of macroscopic quantities or in the restriction of certain ensembles by further macroscopic constraints which only make sense for finite systems. In these cases, general constructions of a common subspace of observables become very relevant. Interestingly, at the end of his discussion on the macroscopic measurement,[16] von Neumann turns to the quantum *H*-theorem and to the relation between entropy and macroscopic measurement. He refers to the then recent work of Pauli,[13,15] who by using "disorder assumptions" or what we could call today, a classical Markov approximation, obtained a general argument for the *H*-theorem.

In the present paper, we are dealing exactly with the problems above and as discussed in Chapter V.4 of Ref. 16. While it is indeed true that averages of the form $A=(a_1+\cdots+a_N)/N, B=(b_1+\cdots+b_N)/N$, for which all commutators $[a_i,b_j]=0$ for $i\neq j$, have their commutator $[A,B]=O(1/N)$ going to zero (in the appropriate norm, corresponding to $[a_i,b_i]=O(1)$) as $N\uparrow+\infty$, it is not true in general that

[a] Also at: U.Antwerpen; electronic mail: wojciech.deroeck@fys.kuleuven.be
[b] Electronic mail: christian.maes@fys.kuleuven.be
[c] Electronic mail: netocny@fzu.cz

$$\lim_{N\to+\infty}\frac{1}{N}\log \mathrm{Tr}[e^{NA}e^{NB}] \stackrel{?}{=} \lim_{N\to+\infty}\frac{1}{N}\log \mathrm{Tr}[e^{NA+NB}].$$

These generating functions are obviously important in fluctuation theory, such as in the problem of large deviations for quantum systems.[12] It is still very much an open question to discuss the joint large deviations of quantum observables, or even to extend the Laplace-Varadhan formula to applications in quantum spin systems. The situation is better for questions about normal fluctuations and the central limit theorem, for which the so-called fluctuation algebra provides a nice framework, see, e.g., Ref. 8. There the pioneering work of André Verbeure will continue to inspire coming generations who are challenged by the features of noncommutativity in quantum mechanics.

These issues are also important for the question of convergence to equilibrium. For example, one would like to specify or to condition on various macroscopic values when starting off the system. Under these constrained equilibria not only the initial energy but also, e.g., the initial magnetization or particle density, etc., are known, and simultaneously installed. As with the large deviation question above, we enter here again in the question of equivalence of ensembles but we are touching also a variety of problems that deal with nonequilibrium aspects. The very definition of configurational entropy as related to the size of the macroscopic subspace has to be rethought when the macroscopic variables get their representation as noncommuting operators. One could again argue that all these problems vanish in the macroscopic limit, but the question (indeed) arises before the limit, for very large but finite N where one can still speak about finite dimensional subspaces or use arguments like the Liouville-von Neumann theorem.

In the following, there are three sections. In Sec. II we write about quantum macrostates and about how to define the macroscopic entropy associated to values of several noncommuting observables. As in the classical case, there is the Gibbs equilibrium entropy. The statistical interpretation, going back to Boltzmann for classical physics, is however not immediately clear in a quantum context. We will define various quantum H-functions. Second, in Sec. III, we turn to the equivalence of ensembles. The main result there is to give a counting interpretation to the thermodynamic equilibrium entropy. In that light we discuss quantum aspects of large deviation theory. Finally, in Sec. IV, we study the relation between macroscopic autonomy and the second law, as done before in Ref. 5 for classical dynamical systems. We prove that if the macroscopic observables give rise to a first-order autonomous equation, then the H-function, defined on the macroscopic values, is monotone. That is further illustrated using a quantum version of the Kac ring model.

II. QUANTUM MACROSTATES AND ENTROPY

Having in mind a macroscopically large closed quantum dynamical system, we consider a sequence $\mathcal{H}=(\mathcal{H}^N)_{N\uparrow+\infty}$ of finite-dimensional Hilbert spaces with the index N labeling different finitely extended approximations, and playing the role of the volume or the particle number, for instance. On each space $\mathcal{H}^N$ we have the standard trace Tr^N. Macrostates are usually identified with subspaces of the Hilbert spaces or, equivalently, with the projections on these subspaces. For any collection $(X_k^N)_{k=1}^n$ of mutually commuting self-adjoint operators there is a projection-valued measure (Q^N) on $\mathbb{R}^n$ such that for any function $F\in C(\mathbb{R}^n)$,

$$F(X_1^N,\ldots,X_n^N)=\int_{\mathbb{R}^n} Q^N(\mathrm{d}z)F(z).$$

A macrostate corresponding to the respective values $x=(x_1,x_2,\ldots,x_n)$ is then represented by the projection

$$Q^{N,\delta}(x) = \int_{\times_k (x_k-\delta, x_k+\delta)} Q^N(\mathrm{d}z)$$

for small enough $\delta > 0$. Furthermore, the Boltzmann H-function, in the classical case counting the cardinality of macrostates, is there defined as

$$H^{N,\delta}(x) = \frac{1}{N} \log \mathrm{Tr}^N[Q^{N,\delta}(x)]$$

with possible further limits $N \uparrow +\infty$, $\delta \downarrow 0$. However, a less trivial problem that we want to address here, emerges if the observables (X_k^N) chosen to describe the system on a macroscopic scale do not mutually commute.

Consider a family of sequences of self-adjoint observables $(X_k^N)_{N\uparrow+\infty, k\in K}$ where K is some index set, and let each sequence be uniformly bounded, $\sup_N \|X_k^N\| < +\infty$, $k \in K$. We call these observables macroscopic, having in mind mainly averages of local observables but that will not always be used explicitly in what follows; it will however serve to make the assumptions plausible.

In what follows, we define concentrating states as sequences of states for which the observables X_k^N assume sharp values. Those concentrating states will be labeled by possible "outcomes" of the observables X_k^N; for these values we write $x=(x_k)_{k\in K}$ where each $x_k \in \mathbb{R}$.

A. Microcanonical setup

1. Concentrating sequences

A sequence $(P^N)_{N\uparrow+\infty}$ of projections is called concentrating at x whenever

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N(F(X_k^N)|P^N) = F(x_k) \tag{2.1}$$

for all $F \in C(\mathbb{R})$ and $k \in K$; we have used the notation

$$\mathrm{tr}^N(\cdot|P^N) = \frac{\mathrm{Tr}^N(P^N \cdot P^N)}{\mathrm{Tr}^N(P^N)} = \frac{\mathrm{Tr}^N(P^N \cdot)}{\mathrm{Tr}^N(P^N)} \tag{2.2}$$

for the normalized trace state on $P^N \mathcal{H}^N$. To indicate that a sequence of projections is concentrating at x we use the shorthand $P^N \overset{\mathrm{mc}}{\to} x$.

2. Noncommutative functions

The previous lines, in formula (2.1), consider functions of a single observable. By properly defining the joint functions of two or more operators that do not mutually commute, the concentration property extends as follows.

Let $\mathcal{I}_K$ denote the set of all finite sequences from K, and consider all maps $G:\mathcal{I}_K \to \mathbb{C}$ such that

$$\sum_{m\geq 0} \sum_{(k_1,\ldots,k_m)\in\mathcal{I}_K} |G(k_1, \ldots, k_m)| \prod_{i=1}^m r_{k_i} < \infty \tag{2.3}$$

for some fixed $r_k > \sup_N \|X_k^N\|, k \in K$. Slightly abusing the notation, we also write

$$G(X^N) = \sum_{m\geq 0} \sum_{(k_1,\ldots,k_m)\in\mathcal{I}_K} G(k_1, \ldots, k_m) X_{k_1}^N \cdots X_{k_m}^N \tag{2.4}$$

defined as norm-convergent series. We write $\mathcal{F}$ to denote the algebra of all these maps G, defining noncommutative "analytic" functions on the multidisc with radii $(r_k), k \in K$.

Proposition 2.1: Assume that $P^N \overset{\mathrm{mc}}{\to} x$. *Then, for all* $G \in \mathcal{F}$,

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N[G(X^N)|P^N] = G(x). \tag{2.5}$$

Remark 2.2: In particular, the limit expectations on the left-hand side of (2.5) coincide for all classically equivalent noncommutative functions. As example, for any complex parameters $\lambda_k, k\in R$ with R a finite subset of K and for $P^N \overset{\mathrm{mc}}{\longrightarrow} x$,

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N(e^{\sum_{k\in R}\lambda_k(X_k^N-x_k)}|P^N) = \lim_{N\uparrow+\infty} \mathrm{tr}^N\left(\prod_{k\in R} e^{\lambda_k(X_k^N-x_k)}|P^N\right) = 1$$

no matter in what order the last product is actually performed.

Proof of Proposition 2.1: For any monomial $G(X^N)=X_{k_1}^N\cdots X_{k_m}^N$, $m\geqslant 1$, we prove the statement of the proposition by induction, as follows. Using the shorthands $Y^N=X_{k_1}^N\cdots X_{k_{m-1}}^N$ and $y=x_{k_1}\cdots x_{k_{m-1}}$, the induction hypothesis reads $\lim_{N\uparrow+\infty}\mathrm{tr}^N(Y^N|P^N)=y$ and we get

$$|\mathrm{tr}^N(Y^N X_{k_m}^N - y x_{k_m}|P^N)|$$

$$=|\mathrm{tr}^N(Y^N(X_{k_m}^N - x_{k_m})|P^N) + x_{k_m}\mathrm{tr}^N(Y^N - y|P^N)|$$

$$\leqslant \|Y^N\|\{\mathrm{tr}^N((X_{k_m}^N - x_{k_m})^2|P^N)\}^{\frac{1}{2}} + |x_{k_m}||\mathrm{tr}^N(Y^N - y|P^N)| \to 0$$

since $P^N \overset{\mathrm{mc}}{\longrightarrow} x$ and (Y^N) are uniformly bounded. That readily extends to all noncommutative polynomials by linearity, and finally to all uniform limits of the polynomials by a standard continuity argument. □

3. H-function

Only the concentrating sequences of projections on the subspaces of the largest dimension become candidates for noncommutative variants of macrostates associated with $x=(x_k)_{k\in K}$, and that maximal dimension yields the (generalization of) Boltzmann's H-function. More precisely, to any macroscopic value $x=(x_k)_{k\in K}$ we assign

$$H^{\mathrm{mc}}(x) = \limsup_{P^N \overset{\mathrm{mc}}{\to} x} \frac{1}{N}\log \mathrm{Tr}^N[P^N], \tag{2.6}$$

where $\limsup_{P^N \overset{\mathrm{mc}}{\to} x} = \sup_{P^N \overset{\mathrm{mc}}{\to} x}\limsup_{N\uparrow+\infty}$ is the maximal limit point over all sequences of projections concentrating at x. By construction, $H^{\mathrm{mc}}(x)\in\{-\infty\}\cup[0,+\infty]$ and we write Ω to denote the set of all $x\in\mathbb{R}^K$ for which $H^{\mathrm{mc}}(x)\geqslant 0$; these are all admissible macroscopic configurations. Slightly abusing the notation, any sequence $P^N \overset{\mathrm{mc}}{\longrightarrow} x$, $x\in\Omega$ such that $\limsup_N(1/N)\log \mathrm{Tr}^N[P^N]=H^{\mathrm{mc}}(x)$, will be called a microcanonical macrostate at x.

4. Example

Take a spin system of N spin-1/2 particles for which the magnetization in the α-direction, $\alpha=1,2,3$, is given by

$$X_\alpha^N = \frac{1}{N}\sum_{i=1}^N \sigma_i^\alpha \tag{2.7}$$

in terms of (copies of) the Pauli matrices σ^α.

Let δ_N be a sequence of positive real numbers such that $\delta_N \downarrow 0$ as $N \uparrow +\infty$. For $\vec{m} = (m_1, m_2, m_3) \in [-1,1]^3$, let $\vec{e} \| \vec{m}$ be a unit vector for which $\vec{m} = m\vec{e}$ with $m \geqslant 0$. Consider $Y^N(\vec{m}) = \Sigma_{\alpha=1}^3 m_\alpha X_\alpha^N$ and its spectral projection $Q^N(\vec{m})$ on $[m-\delta_N, m+\delta_N]$. One easily checks that if $N^{1/2}\delta_N \uparrow +\infty$, then $(Q^N(\vec{m}))_N$ is a microcanonical macrostate at $\vec{m}$, and

$$H^{\mathrm{mc}}(\vec{m}) = \begin{cases} -\dfrac{1-m}{2}\log\dfrac{1-m}{2} - \dfrac{1+m}{2}\log\dfrac{1+m}{2} & \text{for } m \leqslant 1 \\ -\infty & \text{otherwise.} \end{cases}$$

B. Canonical setup

The concept of macrostates as above and associated with projections on certain subspaces on which the selected macroscopic observables take sharp values is physically natural and restores the interpretation of "counting microstates." Yet, sometimes it is not very suitable for computations. Instead, at least when modeling thermal equilibrium, one usually prefers canonical or grand-canonical ensembles, and one relies on certain equivalence of all these ensembles.

1. Concentrating states

For building the ensembles of quantum statistical mechanics, one does not immediately encounter the problem of noncommutativity. One requires a certain value for a number of macroscopic observables and one constructs the density matrix that maximizes the von Neumann entropy.

We write $\omega^N \overset{1}{\rightarrow} x$ for a sequence of states (ω^N) on $\mathcal{H}^N$ whenever $\lim_{N\uparrow+\infty}\omega^N(X_k^N) = x_k$ (convergence in mean).

That construction and that of the concentrating sequences of projections of Sec. II A 1 still has other variants. We say that a sequence of states (ω^N) is concentrating at x and we write $\omega^N \to x$, when

$$\lim_{N\uparrow+\infty} \omega^N(G(X^N)) = G(x) \tag{2.8}$$

for all $G \in \mathcal{F}$. The considerations of Proposition 2.1 apply also here and one can equivalently replace the set of all noncommutative analytic functions with functions of a single variable.

2. Gibbs-von Neumann entropy

The counting entropy of Boltzmann extends to general states such as the von Neumann entropy which is the quantum variant of the Gibbs formula, both being related to the relative entropy defined with respect to a trace reference state. Analogous to (2.6), we define

$$H^{\mathrm{can}}(x) = \limsup_{\omega^N \to x} \frac{1}{N}\mathcal{H}(\omega^N), \tag{2.9}$$

where $\mathcal{H}(\omega^N) \geqslant 0$ is, upon identifying the density matrix σ^N for which $\omega^N(\cdot) = \mathrm{Tr}^N(\sigma^N \cdot)$,

$$\mathcal{H}(\omega^N) = -\mathrm{Tr}[\sigma^N \log \sigma^N]. \tag{2.10}$$

Second, we consider

$$H_1^{\mathrm{can}}(x) = \limsup_{\omega^N \xrightarrow{1} x} \frac{1}{N}\mathcal{H}(\omega^N). \tag{2.11}$$

Obviously, H_1^{can} is the analog of the canonical entropy in thermostatics and the easiest to compute, see also under Sec. II B 3. To emphasize that, we call any sequence of states (ω^N), $\omega^N \xrightarrow{1} x$ such that $\limsup_N (1/N)\mathcal{H}(\omega^N) = H_1^{\mathrm{can}}(x)$ a canonical macrostate at x.

Another generalization of the H-function is obtained when replacing the trace state (corresponding to the counting) with a more general reference state $\rho = (\rho^N)_N$. In that case we consider the H-function as derived from the relative entropy, and differing from the above-used convention by the sign and an additive constant:

$$H_1^{\mathrm{can}}(x|\rho) = \liminf_{\omega^N \xrightarrow{1} x} \frac{1}{N}\mathcal{H}(\omega^N|\rho^N). \tag{2.12}$$

Here, defining σ^N and σ_0^N as the density matrices such that $\omega^N(\cdot) = \mathrm{Tr}[\sigma^N \cdot]$ and $\rho^N(\cdot) = \mathrm{Tr}[\sigma_0^N \cdot]$,

$$\mathcal{H}(\omega^N|\rho^N) = \mathrm{Tr}[\sigma^N(\log \sigma^N - \log \sigma_0^N)]. \tag{2.13}$$

Remark that this last generalization enables one to cross the border between closed and open thermodynamic systems. Here, the state (ρ^N) can be chosen as a nontrivial stationary state for an open system, and the above-defined H-function $H_1^{\mathrm{can}}(x|\rho)$ may lose natural counting and thermodynamic interpretations. Nevertheless, its monotonicity properties under dynamics satisfying suitable conditions justify this generalization, see Sec. IV.

3. Canonical macrostates

The advantage of the canonical formulation of the variational problem for the H-function as in (2.11) is that it can often be solved in a very explicit way. A class of general and well-known examples of canonical macrostates have the following Gibbsian form.[3]

If $\lambda = (\lambda_1, \ldots, \lambda_n)$ are such that the sequence of states (ω_λ^N), $\omega_\lambda^N(\cdot) = \mathrm{Tr}^N(\sigma_\lambda^N \cdot)$ defined by

$$\sigma_\lambda^N = \frac{1}{\mathcal{Z}_\lambda^N} e^{N\Sigma_k \lambda_k X_k^N}, \quad \mathcal{Z}_\lambda^N = \mathrm{Tr}^N\left(e^{N\Sigma_k \lambda_k X_k^N}\right) \tag{2.14}$$

satisfies $\lim_{N\uparrow+\infty} \omega_\lambda^N(X_k^N) = x_k$, $k = 1, \ldots, n$, then (ω_λ^N) is a canonical macrostate at x, and

$$H_1^{\mathrm{can}}(x) = \limsup_N \frac{1}{N} \log \mathcal{Z}_\lambda^N - \sum_k \lambda_k x_k. \tag{2.15}$$

III. EQUIVALENCE OF ENSEMBLES

A basic intuition of statistical mechanics is that adding those many new concentrating states in the variational problem, as done in Sec. II B, does not actually change the value of the H-function. In the same manner of speaking, one would like to understand the definitions (2.9) and (2.11) in counting-terms. In what sense do these entropies represent a dimension (the size) of a (microscopic) subspace?

Trivially, $H^{\mathrm{mc}} \leqslant H^{\mathrm{can}} \leqslant H_1^{\mathrm{can}}$, and $H^{\mathrm{can}}(x) = H_1^{\mathrm{can}}(x)$ iff some canonical macrostate $\omega^N \xrightarrow{1} x$ is actually concentrating at x, $\omega^N \to x$. We give general conditions under which the full equality can be proven. We have again a sequence of observables X_k^N with spectral measure given by the projections $Q_k^N(\mathrm{d}z), k \in K$.

Theorem 3.1: *Assume that for a sequence of density matrices $\sigma^N > 0$, the corresponding $(\omega^N)_N$ is a canonical macrostate at x and that the following two conditions are verified*:

(i) (Exponential concentration property.)

For every $\delta > 0$ *and* $k \in K$ *there are* $C_k(\delta) > 0$ *and* $N_k(\delta)$ *so that*

$$\int_{x_k-\delta}^{x_k+\delta} \omega^N(Q_k^N(\mathrm{d}z)) \geqslant 1 - e^{-C_k(\delta)N} \tag{3.1}$$

for all $N > N_k(\delta)$.

(ii) (Asymptotic equipartition property.)
For all $\delta > 0$,

$$\lim_{N\uparrow+\infty} \frac{1}{N} \log \int_{-\delta}^{\delta} \omega^N(\tilde{Q}^N(\mathrm{d}z)) = 0, \tag{3.2}$$

where $\tilde{Q}^N$ *denotes the projection operator-valued measure of the operator* $(1/N)(\log \sigma^N - \omega^N(\log \sigma^N))$.

Then, $H^{\mathrm{mc}}(x) = H^{\mathrm{can}}(x) = H_1^{\mathrm{can}}(x) \geqslant 0$.

Theorem 3.1 evidently expresses that the microcanonical and the canonical ensembles are equivalent. Results of that kind are well-known in the literature, see e.g., Ref. 14 or 7. An example of a similar type of reasoning for the quantum case is given in Ref. 11. Theorem 3.1 is, however, slightly different from these results in the following aspects,

(1) When considering the quantum microcanonical ensemble, one usually starts out with spectral projections P^N associated with one macroscopic observable. That at least is the approach in Ref. 11 and it is also sketched at the very beginning of Sec. II. Our approach is, however, not limited to one macroscopic observable. Indeed, remember that the $(X_k^N)_k$ need not commute (Sec. II A).
(2) Results on equivalence of ensembles, including those contained in, e.g., Refs. 14, 7, and 11 are mostly dealing solely with translation-invariant lattice spin systems. We do not have that limitation here; instead we have the assumptions (3.2) and (3.1).
(3) Even within the context of translation-invariant lattice spin systems, the results in Refs. 14, 7, and 11 do not yield Theorem 3.1. In these references the microcanonical state is defined as the average of projections P^N, translated over all lattice vectors. That lattice average is translation-invariant by construction (and hence technically easier to handle), but of course it is itself not longer a projection and hence it is not a microcanonical state in the sense of the present paper.

Remarks on the conditions of Theorem 3.1: Whether one can prove the assumptions of Theorem 3.1, depends heavily on the particular model.

The exponential concentration property (3.1) is not trivial even for quantum lattice spin systems, and not even in their one-phase region. Let us mention one criterion under which (3.1) can be checked, which indicates its deep relation to the problem of quantum large deviations. Consider the generating functions

$$\psi_k(t) = \lim_{N\uparrow+\infty} \frac{1}{N} \log \omega^N(e^{tNX_k^N}), \quad k \in K. \tag{3.3}$$

Their existence together with their differentiability at $t=0$ imply by an exponential Chebyshev inequality that ω^N exponentially concentrates at $x = (\psi_k'(0); k \in K)$. However, to our knowledge, the differentiability of $\psi_k(t)$ has only been proven so far for lattice averages over local observables for quantum spin lattice systems in a "high-temperature regime," see Ref. 12, Theorem 2.15 and Remark 7.13, where a cluster expansion technique has been used. The existence of the generating functions (3.3) has also been studied in Ref. 10.

The asymptotic equipartition property (3.2) is easier. The terminology, originally in information theory, comes from its immediate consequence (3.7) below, where P^N projects on a "high probability" region: as in the classical case, the Gibbs-von Neumann entropy measures in some sense the size of the space of "sufficiently probable" microstates. For (3.2) it is enough to prove that the state ω^N is concentrating for the observable

$$A^N = \frac{1}{N} \log \sigma^N. \tag{3.4}$$

Explicitly, it is enough to show that for all $F \in C(\mathbb{R})$,

$$\lim_{N\uparrow+\infty} [\omega^N(F(A^N)) - F(\omega^N(A^N))] = 0. \tag{3.5}$$

In particular, if (ω^N), $\omega^N = \omega_\lambda^N$ is given by formula (2.14), a sufficient condition for the asymptotic equipartition property to be satisfied is that the pressure $p(\lambda)$ defined as

$$p(\lambda) = \lim_{N\uparrow+\infty} \frac{1}{N} \log \mathcal{Z}_\lambda^N \tag{3.6}$$

exists and is continuously differentiable at $\lambda = \lambda(x)$.

Remark that for ergodic states of spin lattice systems, the asymptotic equipartition as expressed by (3.2) and (3.7) follows from the quantum Shannon-McMillan theorem, see Ref. 2, and the references therein. An interesting variant of that result, which touches the problem of quantum large deviations, is the quantum Sanov theorem, proven for i.i.d. processes in Ref. 1. In contrast, our result focuses on the intimate relation of the asymptotic equipartition property to the problem of equivalence of ensembles in the noncommutative context, and Theorem 3.1 formulates sufficient conditions under which such an equivalence follows. An advantage of this approach is that it is not restricted to the framework of spin lattice models with its underlying quasilocal structure.

As $H^{\text{mc}} \leqslant H^{\text{can}} \leqslant H_1^{\text{can}}$, we only need to establish that there is a concentrating sequence of projections for which its H-function equals the Gibbs-von Neumann entropy. Hence, the proof of Theorem 3.1 follows from the following lemma:

Lemma 3.2: If a sequence of states (ω^N) *satisfies conditions* (i) *and* (ii) *of Theorem 3.1, then there exists a sequence of projections* (P^N) *exponentially concentrating at x and satisfying*

$$\lim_{N\uparrow+\infty} \frac{1}{N}(\log \mathrm{Tr}^N(P^N) - \mathcal{H}(\omega^N)) = 0. \tag{3.7}$$

Proof: There exists a sequence $\delta_N \downarrow 0$ such that when substituted for δ, (3.2) is still satisfied. Take such a sequence and define $P^N = \int_{-\delta_N}^{\delta_N} \mathrm{d}\tilde{Q}^N(z)$. By construction,

$$e^{N(h_N-\delta_N)} P^N \leqslant (\sigma^N)^{-1} P^N \leqslant e^{N(h_N+\delta_N)} P^N \tag{3.8}$$

for any $N = 1, 2, \ldots$, with the shorthand $h_N = (1/N)\mathcal{H}(\omega^N)$. That yields the inequalities

$$\mathrm{Tr}^N(P^N) = \omega^N((\sigma^N)^{-1} P^N) \leqslant e^{N(h_N+\delta_N)} \omega^N(P^N) \tag{3.9}$$

and

$$\mathrm{Tr}^N(P^N) \geqslant e^{N(h_N-\delta_N)} \omega^N(P^N). \tag{3.10}$$

Using that $\lim_{N\uparrow+\infty}(1/N)\log \omega^N(P^N) = 0$ proves (3.7).

To see that (P^N) is exponentially concentrating at x, observe that for all $Y^N \geqslant 0$,

$$\begin{aligned}
\omega^N(Y^N) &= \mathrm{Tr}^N((\sigma^N)^{1/2} Y^N (\sigma^N)^{1/2}) \geqslant \mathrm{Tr}^N(P^N(\sigma^N)^{1/2} Y^N (\sigma^N)^{1/2} P^N) \\
&= \mathrm{Tr}^N((Y^N)^{1/2} P^N \sigma^N (Y^N)^{1/2}) \\
&\geqslant e^{N(h_N-\delta_N)} \mathrm{Tr}^N(P^N) \mathrm{tr}^N(Y^N|P^N) \geqslant e^{-2N\delta_N} \omega^N(P^N) \mathrm{tr}^N(Y^N|P^N),
\end{aligned} \tag{3.11}$$

where we used inequalities (3.8)–(3.10). By the exponential concentration property of (ω^N), inequality (3.1), for all $k \in K$, $\epsilon > 0$, and $N > N_k(\epsilon)$,

$$\int_{\mathbb{R}\setminus(x_k-\epsilon,x_k+\epsilon)} \mathrm{tr}^N(\mathrm{d}Q_k^N(z)|P^N) \leqslant e^{-(C_k(\epsilon)-2\delta_N)N}(\omega^N(P^N))^{-1}. \tag{3.12}$$

Choose $N_k'(\epsilon)$ such that $\delta_N \leqslant C_k(\epsilon)/8$ and $(1/N)\log \omega^N(P^N) \geqslant -C_k(\epsilon)/4$ for all $N > N_k'(\epsilon)$. Then (3.12) $\leqslant \exp[-C_k(\epsilon)N/2]$ for all $N > \max\{N_k(\epsilon), N_k'(\epsilon)\}$. □

IV. *H*-THEOREM FROM MACROSCOPIC AUTONOMY

When speaking about an H-theorem or about the monotonicity of entropy one often refers, and even more so for a quantum setup, to the fact that the relative entropy verifies the contraction inequality

$$\mathcal{H}(\omega^N\tau^N|\rho^N\tau^N) \leqslant \mathcal{H}(\omega^N|\rho^N) \tag{4.1}$$

for all states ω^N, ρ^N on $\mathcal{H}^N$ and for all completely positive maps τ^N on $\mathcal{B}(\mathcal{H}^N)$. That is true classically, quantum mechanically and for all small or large N. When the reference state ρ^N is invariant under τ^N, (4.1) yields the contractivity of the relative entropy with respect to ρ^N. However tempting, such inequalities should not be confused with second law or with H-theorems; note in particular that $\mathcal{H}(\omega^N)$ defined in (2.10) is constant whenever τ^N is an automorphism: $\mathcal{H}(\omega^N\tau^N) = \mathcal{H}(\omega^N)$.

In contrast, an H-theorem refers to the (usually strict) monotonicity of a quantity on the macroscopic trajectories as obtained from a microscopically defined dynamics. Such a quantity is often directly related to the fluctuations in a large system and its extremal value corresponds to the equilibrium or, more generally, to a stationary state.

In the previous section we have obtained how to represent a macroscopic state and constructed a candidate H-function. Imagine now a time-evolution for the macroscopic values, always referring to the same set of (possibly noncommuting macroscopic) observables X_k^N. To prove an H-theorem, we need basically two assumptions: macroscopic autonomy and the semigroup property, or that there is a first-order autonomous equation for the macroscopic values. A classical version of this study and more details can be found in Ref. 5.

A. Microcanonical setup

Assume a family of automorphisms $\tau_{t,s}^N$ is given as acting on the observables from $\mathcal{B}(\mathcal{H}^N)$ and satisfying

$$\tau_{t,s}^N = \tau_{t,u}^N\tau_{u,s}^N, \quad t \geqslant u \geqslant s. \tag{4.2}$$

It follows that the trace Tr^N is invariant for $\tau_{t,s}^N$.

Recall that $\Omega \subset \mathbb{R}^K$ is the set of all admissible macroscopic configurations, $H^{\mathrm{mc}}(x) \geqslant 0$. On this space we want to study the emergent macroscopic dynamics.

Autonomy condition. There are maps $(\phi_{t,s})_{t\geqslant s\geqslant 0}$ on Ω and there is a microcanonical macrostate (P^N), $P^N = P^N(x)$ for each $x \in \Omega$, such that for all $G \in \mathcal{F}$ and $t \geqslant s \geqslant 0$,

$$\lim_{N\uparrow+\infty} \mathrm{tr}^N(\tau_{t,s}^N G(X^N)|P^N) = G(\phi_{t,s}x). \tag{4.3}$$

Semigroup property. The maps are required to satisfy the semigroup condition,

$$\phi_{t,u}\phi_{u,s} = \phi_{t,s} \tag{4.4}$$

for all $t \geqslant u \geqslant s \geqslant 0$.

Theorem 4.1: *Assume that the autonomy condition* (4.3) *and the semigroup condition* (4.4) *are both satisfied. Then, for every* $x \in \Omega$, $H^{\mathrm{mc}}(x_t)$ *is nondecreasing in* $t \geqslant 0$ with $x_t = \phi_{t,0}x$.

Proof: Given $x \in \Omega$, fix a microcanonical macrostate $P^N \stackrel{\text{mc}}{\rightarrow} x$ and $t \geqslant s \geqslant 0$. Using that $(\tau_{t,s}^N)^{-1}$ is an automorphism and $\text{Tr}^N((\tau_{t,s}^N)^{-1}\cdot) = \text{Tr}^N(\cdot)$, the identity

$$\text{tr}^N(\tau_{t,s}^N G(X^N)|P^N) = \frac{\text{Tr}^N(G(X^N)(\tau_{t,s}^N)^{-1}P^N)}{\text{Tr}^N((\tau_{t,s}^N)^{-1}P^N)} = \text{tr}^N(G(X^N)|(\tau_{t,s}^N)^{-1}P^N)$$

yields $(\tau_{t,s}^N)^{-1}P^N \stackrel{\text{mc}}{\rightarrow} \phi_{t,s}x$ due to autonomy condition (4.3). Hence,

$$H^{\text{mc}}(\phi_{t,s}x) \geqslant \limsup_{N\uparrow+\infty} \frac{1}{N} \log \text{Tr}^N((\tau_{t,s}^N)^{-1}P^N) = H^{\text{mc}}(x).$$

In particular, one has that $x_s = \phi_{s,0}x \in \Omega$. The statement then follows by the semigroup property (4.3):

$$H^{\text{mc}}(x_t) = H^{\text{mc}}(\phi_{t,0}x) = H^{\text{mc}}(\phi_{t,s}x_s) \geqslant H^{\text{mc}}(x_s).$$

□

It is important to realize that a macroscopic dynamics, even autonomous in the sense of (4.3), need not satisfy the semigroup property (4.1). In that case one actually does not expect the H-function to be monotone; see Ref. 4 and below for an example. As obvious from the proof, without that semigroup property of $(\phi_{t,s})$, (4.3) only implies $H(x_t) \geqslant H(x)$, $t \geqslant 0$. Or, in a bit more generality, it implies that for all $s \geqslant 0$ and $x \in \Omega$ the macrotrajectory $(x_t)_{t\geqslant s}$, $x_t = \phi_{t,s}(x)$ satisfies $H(x_t) \geqslant H(x_s)$ for all $t \geqslant s$.

Remark that while the set of projections is invariant under the automorphisms $(\tau_{t,s}^N)$, this is not true any longer for more general microscopic dynamics defined as completely positive maps, and describing possibly an open dynamical system interacting with its environment. In the latter case the proof of Theorem 4.1 does not go through and one has to allow for macrostates described via more general states, as in Sec. II B. The revision of the argument for the H-theorem within the canonical setup is done in the next section.

B. Canonical setup

We have completely positive maps $(\tau_{t,s}^N)_{t\geqslant s\geqslant 0}$ on $\mathcal{B}(\mathcal{H}^N)$ satisfying

$$\tau_{t,s}^N = \tau_{t,u}^N \tau_{u,s}^N, \quad t \geqslant u \geqslant s \geqslant 0 \tag{4.5}$$

and leaving invariant the state ρ^N; they represent the microscopic dynamics. The macroscopic dynamics is again given by maps $\phi_{t,s}$.

As a variant of autonomy condition (4.3), we assume that the maps $\phi_{t,s}$ are reproduced along the time-evolution in the mean. Namely, see definition (2.12), for every $x \in \Omega_1(\rho) = \{x; H_1^{\text{can}}(x|\rho) < \infty\}$ we ask that a canonical macrostate $\omega^N \stackrel{1}{\rightarrow} x$ exists such that, for all $t \geqslant s \geqslant 0$,

$$\phi_{t,s}x = \lim_{N\uparrow+\infty} \omega^N(\tau_{t,s}^N X^N). \tag{4.6}$$

At the same time, we still assume the semigroup condition (4.4).

Theorem 4.2: *Under conditions* (4.6) and (4.4), *the function* $H_1^{\text{can}}(\phi_{t,0}x|\rho)$ *is nonincreasing in* $t \geqslant 0$ *for all* $x \in \Omega_1(\rho)$.

Proof: If $\omega^N \stackrel{1}{\rightarrow} x$ is a canonical macrostate at x then, by the monotonicity of the relative entropy,

$$H_1^{\text{can}}(x|\rho) = \liminf_{N\uparrow+\infty} \frac{1}{N} \mathcal{H}(\omega^N|\rho^N) \geqslant \liminf_{N\uparrow+\infty} \frac{1}{N} \mathcal{H}(\omega^N \tau_{t,s}^N|\rho^N).$$

On the other hand, by (4.6), the sequence $(\omega^N \tau_{t,s}^N)$ is concentrating in the mean at $\phi_{t,s}(x)$, yielding

$$H_1^{\mathrm{can}}(x|\rho) \geqslant H_1^{\mathrm{can}}(\phi_{t,s}x|\rho).$$

Using (4.4), the proof is now finished as in Theorem 4.1. □

C. Example: The quantum Kac model

A popular toy model to illustrate and to discuss essential features of relaxation to equilibrium has been introduced by Mark Kac.[9] Here we review an extension that can be called a quantum Kac model, we described it extensively in Ref. 4, to learn only later that essentially the same model was considered by Max Dresden and Frank Feiock in Ref. 6. However, there is an interesting difference in interpretation to which we return at the end of the section.

At each site of a ring with N sites there is a quantum bit $\psi_i \in \mathbb{C}^2$ and a classical binary variable $\xi_i = \pm 1$ (which we also consider to be embedded in $\mathbb{C}^2$). The microstates are thus represented as vectors $(\psi;\xi) = (\psi_1, \ldots, \psi_N; \xi_1, \ldots, \xi_N)$, being elements of the Hilbert space $\mathcal{H}^N = \mathbb{C}^{2N} \otimes \mathbb{C}^{2N}$. The time is discrete and at each step two operations are performed: a right shift, denoted below by S^N and a local scattering or update V^N. The unitary dynamics is given as

$$U^N = S^N V^N, \quad U_t^N = (U^N)^t \quad \text{for } t \in \mathbb{N} \tag{4.7}$$

with the shift

$$S^N(\psi;\xi) = (\psi_N, \psi_1, \ldots, \psi_{N-1}; \xi) \tag{4.8}$$

and the scattering

$$V^N(\psi;\xi) = \left(\frac{1-\xi_1}{2} V_1 \psi_1 + \frac{1+\xi_1}{2} \psi_1, \ldots, \frac{1-\xi_N}{2} V_N \psi_N + \frac{1+\xi_N}{2} \psi_N; \xi \right) \tag{4.9}$$

extended to an operator on $\mathcal{H}^N$ by linearity. Here, V is a unitary 2×2 matrix and V_i its copy at site $i = 1, \ldots, N$.

We consider the family of macroscopic observables

$$X_0^N = \frac{1}{N} \sum_{i=1}^N \xi_i, \quad X_\alpha^N = \frac{1}{N} \sum_{i=1}^N \sigma_i^\alpha, \quad \alpha = 1,2,3,$$

where $\sigma_i^1, \sigma_i^2, \sigma_i^3$ are the Pauli matrices acting at site i and embedded to operators on $\mathcal{H}^N$. We fix macroscopic values $x = (\mu, m_1, m_2, m_3) \in [-1, +1]^4$ and we construct a microcanonical macrostate (P^N) in x in the following way.

Let δ_N be a positive sequence in $\mathbb{R}$ such that $\delta_N \downarrow 0$ and $N^{1/2}\delta_N \uparrow +\infty$ as $N \uparrow +\infty$. For $\mu \in [-1, 1]$, let $Q_0^N(\mu)$ be the spectral projection associated to X_0^N, on the interval $[\mu - \delta_N, \mu + \delta_N]$. For $\vec{m} = (m_1, m_2, m_3) \in [-1, 1]^3$, we already constructed a microcanonical macrostate $Q^N(\vec{m})$ in Sec. II A 4. Obviously, $Q_0^N(\mu)$ and $Q^N(\vec{m})$ commute and the product $P^N = Q_0^N(\mu) Q^N(\vec{m})$ is a projection. It is easy to check that P^N is a microcanonical macrostate at $x = (\mu, \vec{m})$.

The construction of the canonical macrostate is standard along the lines of Sec. II B 3. The corresponding H-functions are manifestly equal:

$$H^{\mathrm{mc}}(x) = H_1^{\mathrm{can}}(x) = \eta\left(\frac{1+m}{2}\right) + \eta\left(\frac{1-m}{2}\right) + \eta\left(\frac{1+\mu}{2}\right) + \eta\left(\frac{1-\mu}{2}\right) \tag{4.10}$$

with $\eta(x) = -x \log x$ for $x \in (0,1]$ and $\eta(0) = 0$, otherwise $\eta(x) = -\infty$.

We now come to the conditions of Theorem 4.1. The construction of the macroscopic dynamics and the proof of its autonomy was essentially done in Ref. 4. The macroscopic equation $\xi_t = \xi$ is obvious and the equation for $\vec{m}_t$ can be written, associating $\vec{m}_t$ with the reduced 2×2 density matrix $\nu_t = (1 + \vec{m}_t \cdot \vec{\sigma})/2$, in the form $\nu_t = \Lambda_\mu^t \nu$, $t = 0, 1, \ldots$, where $\Lambda_\mu^t = (\Lambda_\mu)^t$ and

$$\Lambda_\mu(\nu) = \frac{1-\mu}{2} V\nu V^* + \frac{1+\mu}{2}\nu. \tag{4.11}$$

The semigroup condition (4.4) is then also automatically checked.

In order to understand better the necessity of the semigroup property for an H-theorem to be true, compare the above with another choice of macroscopic variables. Assume we had started out with

$$X_0^N = \frac{1}{N}\sum_{i=1}^{N} \xi_i, \quad X_1^N = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^1$$

as the only macroscopic variables, as was done in Ref. 6. A microcanonical macrostate can again be easily constructed by setting $Q_0^N(\mu)$ the spectral projection associated to X_0^N on the interval $[\mu-\delta_N, \mu+\delta_N]$ and $Q_1^N(\vec{m})$ the spectral projection for X_1^N on $[\mu-\delta_N, \mu+\delta_N]$, and finally $P^N = Q_0^N(\mu)Q_1^N(\vec{m})$ as before. The sequence (P^N) defines a microcanonical macrostate at $(\mu,\vec{m})$ and the autonomy condition (4.3) is satisfied. However, the macroscopic evolution does not satisfy the semigroup property (4.4) and, in agreement with that, the corresponding H-functions are not monotonous in time (see Ref. 4).

ACKNOWLEDGMENT

We dedicate this paper to André Verbeure on the occassion of his 65th birthday. W.D.R. thanks the FWO (Flemish Research Fund) for financial support. We thank André Verbeure for useful discussions and for his ongoing interest in this line of research. K.N. acknowledges the support from Project No. AVOZ10100520 in the Academy of Sciences of the Czech Republic.

[1] Bjelaković, I., Deuschel, J. D., Krüger, T., Seiler, S., Siegmund-Schultze, R., and Szkoła, A., "A quantum version of Sanov's theorem," Commun. Math. Phys. **260**, 659–671 (2005).
[2] Bjelaković, I., Deuschel, J. D., Krüger, T., Seiler, S., Siegmund-Schultze, R., and Szkoła, A., "The Shannon-McMillan theorem for ergodic quantum lattice systems," Invent. Math. **155**(1), 203–202 (2004).
[3] Bratelli, O. and Robinson, D. W., *Operator Algebras and Quantum Statistical Mechanics* (Springer, Berlin, 1979), Vol. 2.
[4] De Roeck, W., Jacobs, T., Maes, C., and Netočný, K., "An extension of the Kac ring model," J. Phys. A **36**, 11547–11559 (2003).
[5] De Roeck, W., Maes, C., and Netočný, K., "H-theorems from autonomous equations," Cond-mat/0508089 (2005), J. Stat. Phys. (to be published).
[6] Dresden, M. and Feiock, F., "Models in nonequilibrium quantum statistical mechanics," J. Stat. Phys. **4**, 111–173 (1972).
[7] Georgii, H.-O., *Gibbs Measures and Phase Transitions* (de Gruyter, New York, 1988).
[8] Goderis, D., Verbeure, A., and Vets, P., "Dynamics of fluctuations for quantum lattice systems," Commun. Math. Phys. **128**, 533–549 (1990).
[9] Kac, M., *Probability and Related Topics in Physical Sciences* (Interscience, New York, 1959), Chap. 14.
[10] Lenci, M. and Rey-Bellet, L., "Large deviations in quantum lattice systems: One-phase region," J. Stat. Phys. **119**, 715–746 (2005).
[11] Lima, R., "Equivalence of ensembles in quantum lattice systems: States," Commun. Math. Phys. **24**, 180–192 (1972).
[12] Netočný, K. and Redig, F., "Large deviations for quantum spin systems," J. Stat. Phys., **117**, 521–547 (2004).
[13] Pauli, W., "Uber das H-Theorem vom Anwachsen der Entropie vom Standpunkt der neuen Quantenmechanik," in *Probleme der Moderne Physik* (Arnold Sommerfeld zum 60, Geburtsdag, 1928). Reprinted in *Collected Scientific Papers by Wolfgang Pauli*, Vol. 1, edited by R. Kronig and V. Weisskopf (Interscience, New York, 1964), 549–564.
[14] Simon, B., *The Statistical Mechanics of Lattice Gases* (Princeton University Press, Princeton, 1993).
[15] Tolman, R. C., *The Principles of Statistical Mechanics* (Clarendon Press, Oxford, 1938).
[16] von Neumann, J., *Mathematical Foundations of Quantum Mechanics*, translated form the German edition by R. T. Beyer (Princeton University Press, Princeton, 1955).