Entropy in Quantum Probability I [1]

Dénes Petz [2]

Mathematical Institute of the Hungarian Academy of Sciences
H-1364 Budapest, PF. 127, Hungary

The notion of entropy was introduced by Clausius in order to discuss the thermal behavior of physical systems. After the work of Clausius, Boltzmann, Gibbs and others, the beginning of the exact treatment of entropy was due to von Neumann and Shannon. Their starting point was very different, von Neumann was motivated by quantum mechanics and Shannon founded communication theory.

The entropy of a state describing a physical system is a quantity expressing the uncertainty or randomness of the system. Shannon regarded this uncertainty attached to a physical system as the amount of information carried by the system, so that the entropy of a state can be read as the information carried by the state. His idea comes from the following consideration: If a physical system has a large uncertainty and one receives information on the system by some procedure, then the so obtained information is more valuable than that received from a system having less uncertainty.

It is understood in probability theory that the notion of (Shannon or measure theoretic) entropy has successful applications in a variety of subjects because it determines the asymptotic behavior of certain probabilities in the course of independent trials. First we will discuss this phenomena for finite quantum systems.

By a finite quantum system we mean an algebra of matrices which is stable under taking adjoint. (In other words, a finite quantum system is a finite dimensional C*-algebra.) If $\mathcal{A}$ is such an algebra then there is a linear functional Tr which takes the value 1 at each minimal projection. It is "tracial" in the sense that

$$\mathrm{Tr}\, ab = \mathrm{Tr}\, ba \qquad (a, b \in \mathcal{A})$$

Every functional ω on $\mathcal{A}$ is determined by a density operator $D_\omega \in \mathcal{A}$ in the form

$$\omega(a) = \mathrm{Tr}\, D_\omega\, a \qquad (a \in \mathcal{A})\,.$$

The entropy $S(\omega)$ of a functional ω is defined by means of its density operator as

$$S(\omega) = \mathrm{Tr}\, \eta(D_\omega)\,.$$

This notion was introduced by von Neumann in 1927 and we term it von Neumann's entropy or shortly entropy (cf. [1]). In fact, von Neumann defined the entropy such

---

[1] Published in *Quantum Probability and Related Topics VII*, ed. L. Accardi (World Scientific, Singapore, 1992), pp. 275–297.

[2] E-mail: petz@renyi.hu

a way for the algebra of all bounded operators on a given Hilbert space. In this case any two minimal projections could be transformed into each other by means of a unitary operator which is, of course, belongs to our operator algebra. Therefore it is natural that the minimal projections must have equal weights and it is a question of normalization that we choose this weight to be one. However, if the the operator algebra is not a factor, for example, $\mathcal{A} = M_2(\mathbb{C}) \oplus M_3(\mathbb{C})$ then there is no internal connection between minimal projections of the first summand and those of the second one. Now we just note this point which might become important.

Let $\mathcal{A}$ be a finite quantum system with a faithful state ω. The n-fold algebraic tensor product $\mathcal{A}_n = \mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is again a finite quantum system and the product functional $\omega_n = \omega \otimes \ldots \otimes \omega$ is a state of $\mathcal{A}_n$. Using the obvious identifications, the inclusion $(\mathcal{A}_n, \omega_n) \subset (\mathcal{A}_m, \omega_m)$ holds for $n \leq m$ and we set

$$(\mathcal{A}_\infty, \omega_\infty) = \bigcup \{(\mathcal{A}_n, \omega_n) \,:\, n \in \mathbb{N}\}\,.$$

On the $*$-algebra $\mathcal{A}_\infty$ the right shift endomorphism γ is defined for $a_1 \otimes a_2 \otimes \ldots \otimes a_n \in \mathcal{A}_n$ as

$$\gamma(a_1 \otimes a_2 \otimes \ldots \otimes a_n) = I \otimes a_1 \otimes a_2 \otimes \ldots \otimes a_n \in \mathcal{A}_{n+1}$$

and ω_∞ is invariant under γ. Now perform the GNS-construction with the state ω_∞ and arrive at the triplet $(\pi, \mathcal{H}, \Omega)$. We identify $\mathcal{A}_\infty$ through its faithful representation π with a subalgebra of the generated von Neumann algebra $\mathcal{M} = \pi(\mathcal{A}_\infty)'' \subset \mathcal{B}(\mathcal{H})$. The normal state

$$\omega(a) = \langle \Omega, a\Omega \rangle \qquad (a \in \mathcal{M})$$

is an extension of ω_∞ and the endomorphism γ extends to $\mathcal{M}$ such that the relation $\omega \circ \gamma = \omega$ is preserved. (For the sake of simpler notation we do not use a new letter for the extension.)

The following well-known result may be called the weak law of large numbers (for independent finite quantum systems.)

Lemma 1. In the above described situation the following statements hold.

(i) If $a \in \mathcal{M}$ and $\gamma(a) = a$ then $a \in \mathbb{C}\, I$.

(ii) For every $a \in \mathcal{M}$ the sequence $S_n(a) = n^{-1}(a + \gamma(a) + \ldots + \gamma^{n-1}(a))$ converges to $\omega(a)I$ in the strong operator topology.

(iii) If $a \in \mathcal{M}^{sa}$ and $J \subset \mathbb{R}$ is closed interval such that $\omega(a) \notin J$ and p_n is the spectral projection of $S_n(a)$ corresponding to the interval J then $p_n \to 0$ in the strong operator topology.

Let us fix a positive number $\varepsilon < 1$. For a while we say that a projection $Q_n \in \mathcal{A}_n$ is rather sure if $\omega_n(Q_n) \geq 1 - \varepsilon$. On the other hand, the size of Q_n, the cardinality of a maximal pairwise orthogonal family of projections contained in Q_n, is given by $\mathrm{Tr}_n\, Q_n$. (The subscript n in Tr_n indicates that the algebraic trace functional on $\mathcal{A}_n$ is meant here.)

The theorem below says that the von Neumann's entropy of ω governs asymptotically the size of rather sure projections: A rather sure projection in $\mathcal{A}_n$ contains at least $\exp(nS(\omega))$ pairwise orthogonal minimal projection.

Theorem 2. ([2]) Under the above conditions and with the above notation the limit relation

$$\lim_{n\to\infty} \frac{1}{n} \inf\{\log \mathrm{Tr}_n\, Q_n \,:\, Q_n \in \mathcal{A}_n \text{ is a projection, } \omega_n(Q_n) \geq 1-\varepsilon\} = S(\omega)$$

holds.

Proof. If D_n denotes the density of ω_n then one can see easily that

$$-\log D_n = \sum_{i=0}^{n-1} \gamma^i(-\log D_1),$$

where γ stands for the right shift. The sequence $(\gamma^i(-\log D_1))$ behaves as independent identically distributed random variables with respect to the state ω_∞. More precisely, the previous lemma applies for $a = -\log D_1$ and tells

$$\frac{1}{n} \log D_n \to S(\omega) I$$

strongly. Let $P(n,\delta)$ be the spectral projection of the self-adjoint operator $-n^{-1}\log D_n$ corresponding to the interval $(S(\omega)-\delta,\, S(\omega)+\delta)$. According to (iii) of Lemma 1 one has

$$P(n,\delta) \to I \tag{1}$$

strongly for every $\delta > 0$. In particular,

$$\omega(P(n,\delta)) = \langle P(n,\delta)\Omega, \Omega\rangle \to 1$$

as $n\to\infty$ and $P(n,\delta)$ is a rather sure projection if n is large enough. It follows from the definition of $P(n,\delta)$ that

$$D_n\, P(n,\delta) \exp(n\, S(\omega) - n\delta) \leq P(n,\delta) \leq D_n\, \exp(n\, S(\omega) + n\delta) \tag{2}$$

which gives

$$\frac{1}{n} \log\, \mathrm{Tr}_n\, P(n,\delta) \leq S(\omega) + \delta\,.$$

Since $\delta > 0$ was arbitrary, we establish

$$\limsup_{n\to\infty} \frac{1}{n}\, \inf\{\log \mathrm{Tr}_n\, Q_n \,:\, Q_n\,\} \leq S(\omega)\,. \tag{3}$$

To prove that $S(\omega)$ is actually the limit we shall argue by contradiction. Assume that there exist a sequence $n(1) < n(2) < \ldots$ of integers, a number $t > 0$ and projections $Q(n(k)) \in \mathcal{A}_{n(k)}$ $(k = 1, 2, \ldots)$ such that

(i) $\omega_\infty(Q(n(k)) \geq 1 - \varepsilon$,

(ii) $\log \mathrm{Tr}_{n(k)} Q(n(k)) \leq n(k)(S(\omega) - t)$.

The bounded sequence $(Q(n(k)))_k$ has a weak limit point in the von Neumann algebra $\mathcal{M}$, say $T \in \mathcal{M}$. Instead of selecting a subsequence we suppose that $Q(n(k)) \to T$ weakly. It is straightforward to show that from (1) the weak limit

$$Q(n(k))P(n(k),\delta) \to T$$

follows. Consequently,

$$\liminf_{k\to\infty} \omega_\infty(Q(n(k))P(n(k),\delta)) \geq \omega(T) \geq 1 - \varepsilon\,. \tag{4}$$

Using the first part of (2) we estimate

$$\begin{array}{rcl} \mathrm{Tr}\, Q(n(k)) & \geq & \mathrm{Tr}\, Q(n(k))P(n(k),\delta) \\ & \geq & \mathrm{Tr}\, D_{n(k)} Q(n(k))P(n(k),\delta) \exp(n\, S(\omega) - n\delta) \\ & = & \exp(n\, S(\omega) - n\delta)\omega_\infty(Q(n(k))P(n(k),\delta)) \end{array}$$

and

$$\begin{aligned} &\liminf_{k\to\infty} \frac{1}{n(k)} \log \mathrm{Tr}_{n(k)} Q(n(k)) \\ &\qquad \geq S(\omega) - \delta + \lim_{k\to\infty} \frac{1}{n(k)} \log \omega_\infty(Q(n(k))P(n(k),\delta))\,. \end{aligned}$$

The limit term on the right hand side vanishes due to (4) and we arrive at a contradiction with (ii) if $0 < \delta < t$. This proves the theorem. ☐

Opposite to the commutative case the state space of a quantum system is not a Choquet simplex in the sense that states admit several extremal decompositions. For example, for $\mathcal{A} = M_2(\mathbb{C})$ the general form of a density matrix is

$$D = \frac{1}{2}\begin{pmatrix} 1+a & b+\mathrm{i}c \\ b-\mathrm{i}c & 1-a \end{pmatrix}, \tag{5}$$

where a, b, c are real numbers and $a^2 + b^2 + c^2 \leq 1$. Thanks to the affine correspondence $D \longleftrightarrow (a, b, c)$ we can visualize the state space as a ball (of radius 1) and surface points correspond to pure states.

Let φ be a state of a finite quantum system and $\varphi = \sum_i \lambda_i \psi_i$ be an extremal decomposition (that is, every ψ_i is pure). Approaching from information theory one might think that the entropy of φ is $-\sum_i \lambda_i \log \lambda_i$. This, however, would not be satisfactory because the λ_i's are not in general the probabilities of mutually exclusive events. In fact,

$$S(\varphi) \leq -\sum_i \lambda_i \log \lambda_i \tag{6}$$

and the equality holds if and only if the extremal decomposition $\sum \lambda_i \psi_i$ is orthogonal. This was obtained in [3] a long time ago and here it could be deduced by means of the relative entropy as well. The inequality (6) is interpreted as follows. In the sense of information content, the most economical extremal decomposition is the orthogonal one, which is implemented by the density matrix.

The relative entropy is an information measure representing the uncertainty of a state with respect to another state. Hence it indicates a kind of distance between the two states. In information theory the relative entropy $S(\nu, \mu)$ of two finite probability distributions $\nu = (\kappa_1, \kappa_2, \ldots, \kappa_n)$ and $\mu = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ on an n-point space is usually defined by

$$S(\nu, \mu) = \begin{cases} \sum_i \kappa_i (\log \kappa_i - \log \lambda_i) & \text{if } \lambda_i = 0 \text{ implies } \kappa_i = 0 \\ +\infty & \text{otherwise .} \end{cases} \tag{7}$$

In the quantum case ([23]) the entropy of ω with respect to φ is defined by

$$S(\omega, \varphi) = \begin{cases} \operatorname{Tr} D_\omega (\log D_\omega - \log D_\varphi) & \text{if } \operatorname{supp} D_\varphi \ \geq \ \operatorname{supp} D_\omega \\ +\infty & \text{otherwise .} \end{cases} \tag{8}$$

Here $\operatorname{supp} D_\psi$ denotes the smallest projection p such that $\psi(p) = \psi(I)$. In particular, $S(\omega, \varphi)$ is always finite if the density of φ has strictly positive eigenvalues. (Such a φ is called faithful.) When D_φ commutes with D_ω and their eigenvalue lists are $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $(\kappa_1, \ldots, \kappa_n)$ respectively, then $S(. \, , \, .)$ reduces to the classical expression (7) due to Kullback and Leibler.

Although we mostly speak of the relative entropy of states it is convenient to allow ω and φ in the definition of $S(\omega, \varphi)$ to be arbitrary positive functionals.

Now we approach quantum relative entropy axiomatically. Our crucial postulate for the relative entropy includes the notion of conditional expectation. Let us recall that in the setting of operator algebras conditional expectation (or projection of norm one) is defined as a positive unital idempotent linear mapping onto a subalgebra.

Now we list the properties of the relative entropy functional which will be used in an axiomatic characterization:

(i) Conditional expectation property: Assume that $\mathcal{A}$ is a subalgebra of $\mathcal{B}$ and there exists a projection of norm one E of $\mathcal{B}$ onto $\mathcal{A}$ such that $\varphi \circ E = \varphi$. Then for every state ω of $\mathcal{B}$ $S(\omega, \varphi) = S(\omega|\mathcal{A}, \varphi|\mathcal{A}) + S(\omega, \omega \circ E)$ holds.

(ii) Invariance property: For every automorphism α of $\mathcal{B}$ we have $S(\omega, \varphi) = S(\omega \circ \alpha, \varphi \circ \alpha)$.

(iii) Direct sum property: Assume that $\mathcal{B} = \mathcal{B}_1 \oplus \mathcal{B}_2$. Let $\varphi_{12}(a \oplus b) = \lambda \varphi_1(a) + (1 - \lambda)\varphi_2(b)$ and $\omega_{12}(a \oplus b) = \lambda \omega_1(a) + (1 - \lambda)\omega_2(b)$ for every $a \in \mathcal{B}_1, b \in \mathcal{B}_2$ and some $0 < \lambda < 1$. Then $S(\omega_{12}, \varphi_{12}) = \lambda S(\omega_1, \varphi_1) + (1 - \lambda) S(\omega_2, \varphi_2)$.

(iv) Nilpotence property: $S(\varphi, \varphi) = 0$.

(v) Measurability property: The function $(\omega,\varphi)\mapsto S(\omega,\varphi)$ is measurable on the state space of the finite dimensional C*-algebra $\mathcal{B}$ (when φ is assumed to be faithful).

The properties (i)-(v) are well-known properties of the relative entropy functional. Among them the conditional expectation property is the most crucial and it will be discussed below. The quantity

$$S_{\rm BS}(\omega,\varphi)=\operatorname{Tr} D_\omega\,\log(D_\omega^{1/2}D_\varphi^{-1}D_\omega^{1/2}) \tag{9}$$

shares the properties (ii)-(v) and coincides with the relative entropy for commuting densities. (Variant (9) of the relative entropy appeared in [4].) It was obtained in [5] that

$$S_{\rm BS}(\omega,\varphi)\ge S(\omega,\varphi) \tag{10}$$

which is equivalent to the inequality

$$\operatorname{Tr} A\log AB\le \operatorname{Tr} A(\log A+\log B)$$

for positive definite matrices A and B.

The relative entropy $S(\omega,\varphi)$ is the informational divergence of ω from φ. In this spirit the conditional expectation property has a rather natural interpretation. The informational divergence $S(\omega,\varphi)$ has two components. One component is the divergence of ω from φ on the subalgebra $\mathcal{A}$. The other component is coming from the extension procedure of a state on the subalgebra to the whole algebra $\mathcal{B}$. Relative to a state φ (or rather to the conditional expectation E), the natural extension of $\omega|\mathcal{A}$ to $\mathcal{B}$ is obviously $\omega\circ E$. Hence the second component of $S(\omega,\varphi)$ is the informational divergence of ω from $\omega\circ E$. (Note that if the φ-preserving conditional expectation of $\mathcal{B}$ onto $\mathcal{A}$ does not exist then $\omega|\mathcal{A}$ does not have a natural extension to $\mathcal{B}$ relative to φ and in this case our argument does not lead to any conclusion.)

The interpretation of the invariance, direct sum and nilpotence properties is obvious. The measurability is a merely technical axiom.

Theorem 3. ([6]) If a real valued functional $R(\omega,\varphi)$ defined for faithful states φ and arbitrary states ω of finite quantum systems shares the properties (i)-(v), then there exists a constant $c\in\mathbb{R}$ such that

$$R(\omega,\varphi)=c\operatorname{Tr} D_\omega(\log D_\omega-\log D_\varphi)\,.$$

A variant of Theorem 2 holds for the relative entropy. Let us use the setting of Theorem 2 but in addition let φ be a faithful state on the finite quantum system $\mathcal{A}$. We write φ_n and φ_∞ for the corresponding product states of $\mathcal{A}_n$ and $\mathcal{A}_\infty$, respectively. Instead of the trace, φ_n will measure the size of the projections in $\mathcal{A}_n$ while ω_n remains being interpreted as probability. Set

$$R(n,\varepsilon)=\inf\{\log\varphi_n\,(Q_n)\,:\,Q_n\in\mathcal{A}_n \text{ is a projection, } \omega_n(Q_n)\ge 1-\varepsilon\}$$

for positive ε.

Theorem 4. ([5])

$$\lim_{\varepsilon\to+0}\lim_{n\to\infty}\frac{1}{n}R(n,\varepsilon)=-S(\omega,\varphi)$$

The relative entropy may be defined for linear functionals of an arbitrary C*-algebra. The general definition goes through von Neumann algebras and normal states.

Let $\mathcal{M}$ be a von Neumann algebra with normal states φ and ω. The relative entropy $S(\omega,\varphi)$ will be defined by means of the spatial derivative operator. Of course, for finite quantum systems the new definition reduces to the previous one based upon density matrices.

Let $\mathcal{M}$ act on a Hilbert space $\mathcal{H}$ so that $\omega=\omega_\xi$ is a vector state given by a vector $\xi\in\mathcal{H}$. The vector $\xi\in\mathcal{H}$ induces a vector state ω'_ξ on the commutant $\mathcal{M}'$ of $\mathcal{M}$. The spatial derivative $\Delta(\varphi/\omega'_\xi)$ is at our disposal. This is a positive self-adjoint operator with support $[\mathcal{M}\xi]$ supp φ. (Here supp φ, the support projection of φ, belongs to $\mathcal{M}$ and $[\mathcal{M}\xi]$ stands for the orthogonal projection onto the closure of the linear manifold $\mathcal{M}\xi$. The latter projection is an element of $\mathcal{M}'$.) Araki defined the relative entropy as follows ([7]).

$$S(\omega,\varphi)=\begin{cases}+\infty & \text{if } \xi\notin \text{ supp } \varphi\\ -\langle\log\Delta(\varphi/\omega'_\xi)\xi,\xi\rangle & \text{otherwise.}\end{cases}\tag{11}$$

Note that $\xi\in$ supp φ is equivalent to supp $\omega\le$ supp φ. Let

$$\Delta(\varphi/\omega'_\xi)=\int_0^\infty\lambda dE_\lambda$$

be the spectral decomposition. More precisely, (11) is meant as

$$-\langle\log\Delta(\varphi/\omega'_\xi)\xi,\xi\rangle=-\int_0^1\log\lambda d\langle E_\lambda,\xi,\xi\rangle-\int_1^\infty\log\lambda d\langle E_\lambda\xi,\xi\rangle\,.\tag{12}$$

Since $\log\lambda\le\lambda$ and

$$\int_0^\infty\lambda d\langle E_\lambda\xi,\xi\rangle=\|\Delta(\varphi/\omega'_\xi)^{1/2}\xi\|^2=\varphi(I)$$

we see that the second term in the right hand side of (12) is always finite. So $S(\omega,\varphi)$ is finite or $+\infty$ depending on the integral $\int_0^1\log\lambda d\langle E_\lambda\xi,\xi\rangle$. Till now it is not clear whether this definition is independent of the auxiliary vector ξ.

It is desirable to get rid of the domain problem in (11) caused by the logarithmic function. The following equivalent definition is essentially due to Uhlmann who embedded it into a quadratic interpolation machinery ([8]).

$$S(\omega,\varphi)=-\lim_{t\to+0}t^{-1}(\|\Delta(\varphi/\omega'_\xi)^{t/2}\xi\|^2-\|\xi\|^2)\,.\tag{13}$$

Since for a given $\lambda > 1$ $(0 < \lambda < 1)$

$$\lim_{t\to+0} t^{-1}(\lambda^t - 1) = \log\lambda \tag{14}$$

increasingly (decreasingly), the monotone convergence theorem of integrals ensures that (11) and (13) are equivalent.

The spatial derivative operator involved in the definition of the relative entropy is a representation dependent self-adjoint operator. It is useful to have at our disposal a formula expressing the relative entropy of two states by means of operator algebraic terms in a representation independent way. We recall that the Connes' cocycle is a one-parameter family of contractions. For normal states φ and ω of the von Neumann algebra $\mathcal{M}$ acting on the Hilbert space $\mathcal{H}$ the Connes' cocycle is

$$[D\varphi, D\omega]_t = \Delta(\varphi/\psi')^{it}\ \Delta(\omega/\psi')^{-it} \qquad (t\in\mathbb{R}),$$

where ψ' is a faithful normal functional on the commutant $\mathcal{M}'$.

Theorem 5. ([8]) If φ and ω are normal state on a von Neumann algebra such that $S(\varphi,\omega)$ is finite then

$$S(\omega,\varphi) = \mathrm{i}\lim_{t\to+0} t^{-1}(\omega([D\varphi, D\omega]_t) - 1)\,.$$

In principle Theorem 5 gives possibility to compute relative entropy. On injective von Neumann algebras the finite dimensional approximation allows a convenient determination of the relative entropy. Let $\mathcal{M}$ be an injective von Neumann algebra with normal states φ and ω. Then $S(\omega,\varphi)$ is the supremum of all the quantities $S(\omega\circ\alpha, \varphi\circ\alpha)$, where α runs over all completely positive unital mappings from a finite dimensional algebra into $\mathcal{M}$. (See [9] for the detailed proof.) Next we discuss another formula due to Kosaki, which is too complicated for computational purposes but, on the other hand, it shows all the main properties of the relative entropy.

Theorem 6. ([10]) Assume that N is a linear subspace of a von Neumann algebra $\mathcal{M}$ and that I is included in N. If N is dense in the strong*- operator topology, then for any $\varphi,\omega\in\mathcal{M}_*^+$ we obtain

$$S(\omega,\varphi) = \sup\,\sup\,\{\omega(I)\log n - \int_{1/n}^{\infty}[\omega(y(t)^*y(t)) + t^{-1}\varphi(x(t)x(t)^*]t^{-1}dt\},$$

where the first sup is taken over all natural numbers n, the second one is over all step functions $x : (1/n,\infty)\to N$ with finite range and $y(t) = I - x(t)$.

Kosaki's formula contains very important, nevertheless not characteristic, properties of the relative entropy. We say so because the formula is based upon operator convexity of the function $-\log x$ and several other entropy like quantities share similar properties. Now we turn to the conditional expectation property (used already for axiomatic determination of the relative entropy).

Theorem 7. ([11]) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{M}_1$ its von Neumann subalgebra. Assume that there exists a faithful normal conditional expectation E of $\mathcal{M}$ onto $\mathcal{M}_1$. If φ_1 and ω are states of $\mathcal{M}_1$ and $\mathcal{M}$, respectively then

$$S(\omega, \varphi_1 \circ E) = S(\omega|\mathcal{M}_1, \varphi_1) + S(\omega, \omega \circ E).$$

The simplest definition of the relative entropy has been given by statistical operators, in the more general situation statistical operators are not available (due to the lack of trace functional) and the spatial derivative operator is the crucial object in the definition. For positive functionals of an arbitrary C*-algebra $\mathcal{A}$, the relative entropy may be defined through the enveloping von Neumann algebra $\mathcal{A}^{**}$. Let us recall that the second dual $\mathcal{A}^{**}$ of $\mathcal{A}$ is the double commutant of $\mathcal{A}$ in its universal representation. So every state ψ of $\mathcal{A}$ in its universal representation. So every state ψ of $\mathcal{A}$ admits a unique normal extension $\tilde{\psi}$ to $\mathcal{A}^{**}$. We set

$$S(\psi_1, \psi_2) = S(\tilde{\psi}_1, \tilde{\psi}_2),$$

where the right hand side is defined for the normal states $\tilde{\psi}_1$ and $\tilde{\psi}_2$. Kosaki's formula supplies us with an equivalent definition.

$$S(\nu, \psi) = \sup\sup\{\nu(I)\log n - \int_{1/n}^{\infty} [\nu(y(t)^* y(t) - t^{-1}\psi(x(t)x(t)^*)]t^{-1}dt\}\,.$$

Here the double supremum is taken as it was described in Theorem 7. The GNS-construction may be used to have a third equivalent. Let $(\mathcal{H}, \Phi, \pi)$ stand for the GNS-triplet for the unital C*-algebra $\mathcal{A}$ and the positive functional φ of $\mathcal{A}$. Let ω be a positive functional on $\mathcal{A}$. We write $\bar{\omega}$ for the normal state of $\pi(\mathcal{A})''$ such that

$$\bar{\omega}(\pi(a)) = \omega(a) \qquad (a \in \mathcal{A})$$

if a normal functional with this property exists. (When $\bar{\omega}$ exists one says that ω is quasi-contained in φ.) We check that

$$S(\omega, \varphi) = \begin{cases} S(\bar{\omega}, \bar{\varphi}) & \text{if } \bar{\omega} \text{ exists,} \\ +\infty & \text{otherwise.} \end{cases} \tag{16}$$

If $\bar{\omega}$ exists then $S(\omega, \varphi) = S(\bar{\omega}, \bar{\varphi})$ follows obviously from Kosaki's formula. Now assume that ω is not quasi contained in φ. One knows that in this case the central support z_ω of $\tilde{\omega}$ is not smaller than the central support z_φ of $\tilde{\varphi}$. Therefore there exists a projection p in $\mathcal{A}^{**}$ such that $\tilde{\varphi}(p) = 0$ and $\tilde{\omega}(p) \neq 0$. This yields $S(\tilde{\omega}, \tilde{\varphi}) = +\infty$). Properties of the relative entropy functional of states of a C*-algebra are easily deduced from the von Neumann algebra case. We summarize the most important ones.

Theorem 8. The relative entropy of positive functionals of a C*-algebra shares the following properties.

(i) $(\omega, \varphi) \mapsto S(\omega, \varphi)$ is convex and weakly lower semi-continuous.

(ii) $\|\varphi - \omega\|^2 \le 2S(\omega, \varphi)$ if $\varphi(I) = \omega(I) = 1$.

(iii) $S(\omega, \varphi_1) \ge S(\omega, \varphi_2)$ if $\varphi_1 \le \varphi_2$.

(iv) For a unital Schwarz map $\alpha : \mathcal{A}_0 \to \mathcal{A}_1$ the relation $S(\omega \circ \alpha, \varphi \circ \alpha) \le S(\omega, \varphi)$ holds.

(v) For $\omega = \sum_i \omega_i$ we have $S(\omega, \varphi) + \sum_{i=1}^n S(\omega_i, \omega) = \sum_{i=1}^n S(\omega_i, \varphi)$.

These properties have been obtained in different papers at different levels of generality. (i) is related to [10] and [12], (ii) is essentially from [13], (iv) is a result from [8] and (v) is from [15].

It is noteworthy that one can infer easily the strict convexity of the relative entropy in the second variable. Let $\varphi, \omega_1, \omega_2$ be states and assume that $S(\varphi, \omega_1)$ and $S(\varphi, \omega_2)$ are finite. Then the Donald identity (v) gives

$$\begin{aligned}&\lambda S(\omega_1, \varphi) + (1-\lambda)S(\omega_2, \varphi) - S(\lambda\omega_1 + (1-\lambda)\omega_2, \varphi)\\ &\quad = \lambda S(\omega_1, \lambda\omega_1 + (1-\lambda)\omega_2) + (1-\lambda)S(\omega_2, \lambda\omega_1 + (1-\lambda)\omega_2)\,.\end{aligned}$$

If for $0 < \lambda < 1$ the left hand side vanishes then $\omega_1 = \omega_2$ according to property (ii). The strict convexity is clear also from the following estimate.

Theorem 9. Let φ and ω_i be states on a C*-algebra (i=1,2,…,n). Then

$$0 \ge S(\sum_{i=1}^n \lambda_i \omega_i, \varphi) - \sum_{i=1}^n \lambda_i S(\omega_i, \varphi) \ge \sum_{i=1}^n \lambda_i \log \lambda_i$$

provided that $\lambda_i \ge 0$ and $\sum_{i=1}^n \lambda_i = 1$.

Proof. We concentrate on the second inequality. By Donald's identity

$$S(\sum_{i=1}^n \lambda_i \omega_i, \varphi) = \sum_{i=1}^n \lambda_i S(\omega_i, \varphi) - \sum_{j=1}^n \lambda_j S(\omega_j, \sum_{i=1}^n \lambda_i \omega_i)\,.$$

Since $S(\omega_j, \sum_{i=1}^n \lambda_i \omega_i) \le S(\omega_j, \lambda_j \omega_j) = -\log \lambda_j$ the statement is proved. □

The next result appeared in connection with the definition of the dynamical entropy in [14] and could be called marginal inequality.

Theorem 10. Let $(\varphi_{j,k})$ be a finite family of positive functionals on a C*-algebra. If $\varphi = \sum_{j,k} \varphi_{j,k}$, $\varphi_j^{(1)} = \sum_k \varphi_{j,k}$ and $\varphi_k^{(2)} = \sum_j \varphi_{j,k}$ then

$$\sum_{j,k} S(\varphi_{j,k}, \varphi) \ge \sum_j S(\varphi_j^{(1)}, \varphi) + \sum_k S(\varphi_k^{(2)}, \varphi)\,.$$

Theorem 11. ([11) Let $\mathcal{K}$ be a closed convex set in the state space of the C*-algebra $\mathcal{A}$ and φ be a state of $\mathcal{A}$. If $F : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous convex functional and

$$d = \inf\{F(\omega) + S(\omega, \varphi) : \omega \in \mathcal{K}\}$$

is finite, then there exists a unique $\psi \in \mathcal{K}$ such that $F(\psi) + S(\varphi, \psi) = d$.

Proof. We choose a sequence $(\omega_n) \subset \mathcal{K}$ such that $\lim_{n\to\infty}\{F(\omega_n) + S(\omega_n, \varphi)\} = d$. From the convexity of F and the identity (v) of Theorem 8 we infer

$$\begin{aligned} & F(\omega_n) + S(\omega_n, \varphi) + F(\omega_m) + S(\omega_m, \varphi) \\ & \geq 2F\left(\frac{\omega_n + \omega_m}{2}\right) + 2S\left(\frac{\omega_n + \omega_m}{2}, \varphi\right) \\ & \quad + S\left(\omega_m, \frac{\omega_n + \omega_m}{2}\right) + S\left(\omega_n, \frac{\omega_n + \omega_m}{2}\right) \\ & \geq 2d + S\left(\omega_n, \frac{\omega_n + \omega_m}{2}\right) + S\left(\omega_m, \frac{\omega_n + \omega_m}{2}\right). \end{aligned}$$

Hence

$$S\left(\omega_n, \frac{\omega_n + \omega_m}{2}\right) \leq \varepsilon$$

if n and m are big enough. The inequality (ii) of Proposition 5.22 guarantees that (ω_n) is a Cauchy sequence and it converges to a state $\psi \in \mathcal{K}$ (in norm). From lower semi-continuity

$$F(\psi) + S(\psi, \varphi) \leq \liminf_n \{F(\omega_n) + S(\omega_n, \varphi\} \leq d$$

and ψ is a minimizer. Since any minimizing sequence is convergent the minimizer must be unique. □

Theorem 12. Let $\varphi_1 \otimes \varphi_2$ and ω_{12} be normal states of the tensor product von Neumann algebra $\mathcal{M}_1 \otimes \mathcal{M}_2$ and let $\omega_i = \omega_{12}|\mathcal{M}_i$ $(i = 1, 2)$. Then

$$S(\omega_{12}, \varphi_1 \otimes \varphi_2) = S(\omega_1, \varphi_1) + S(\omega_{12}, \omega_1 \otimes \varphi_2) \tag{16}$$

Proof. The equality holds trivially if the support of ω_{12} is strictly larger then that of $\varphi_1 \otimes \varphi_2$. Hence we may assume that $\varphi_1 \otimes \varphi_2$ is faithful. The mapping $a \otimes b \mapsto \varphi_2(b)a$ extends to a faithful normal conditional expectation E which preserves the state $\varphi_1 \otimes \varphi_2$. Hence Theorem 7 applies and $\omega_{12} \circ E = \omega_1 \otimes \varphi_2$ yields the statement. □

According to the previous theorem

$$S(\omega_{12}, \omega_1 \otimes \varphi_2) = S(\omega_2, \varphi_2) + S(\omega_{12}, \omega_1 \otimes \omega_2) \tag{17}$$

and a combination of (16) and (17) is the identity

$$S(\omega_{12}, \varphi_1 \otimes \varphi_2) = S(\omega_1, \varphi_1) + S(\omega_2, \varphi_2) + S(\omega_{12}, \omega_1 \otimes \omega_2). \tag{18}$$

We claim that

$$S(\omega_{12}, \varphi_1 \otimes \varphi_2) \geq S(\omega_1, \varphi_1) + S(\omega_2, \varphi_2) \tag{19}$$

with the above notation. Moreover, if the equality holds and the relative entropies are finite, then $\omega = \omega_1 \otimes \omega_2$. The proof of (19) is nothing else but a reference to the monotonicity of the relative entropy.

Now we turn to the C*-algebra version of (19). Let $\varphi_1 \otimes \varphi_2$ and ω_{12} be states of the (projective) tensor product $\mathcal{A}_1 \otimes \mathcal{A}_2$ of the C*-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ and let $\omega_i = \omega_{12}|\mathcal{A}_i$ $(i = 1, 2)$ as before. If ω_{12} is not quasi-contained in $\varphi_1 \otimes \varphi_2$ then (19) holds trivially with $S(\omega_{12}, \varphi_1 \otimes \varphi_2) = +\infty$. If ω_{12} is quasi-contained in $\varphi_1 \otimes \varphi_2$ then we may pass to the GNS-representation and may refer to the von Neumann algebraic version of (19).

Theorem 13. Let φ be a positive normal functional of the von Neumann algebra $\mathcal{M}$ and let $t \in \mathbb{R}$. Then the set

$$\mathcal{K}(\varphi, t) = \{\omega \in \mathcal{M}_+^* : S(\omega, \varphi) \leq t\}$$

consists of normal functional and it is a convex compact set with respect to the $\sigma(\mathcal{M}_*, \mathcal{M})$ topology.

Proof. If $S(\omega, \varphi)$ is finite then ω is quasi-contained in φ and so ω is normal. Therefore the $\sigma(\mathcal{M}_*, \mathcal{M})$ and the $\sigma(\mathcal{M}^*, \mathcal{M})$ topologies coincide on $\mathcal{K}(\varphi, t)$. To show compactness it suffices to note that $\mathcal{K}(\varphi, t)$ is closed and bounded. Boundedness follows from

$$\omega(I)(\log \omega(I) - \log \varphi(I)) \leq S(\omega, \varphi)$$

and $\mathcal{K}(\varphi, t)$ is closed and convex as a consequence of lower semi-continuity and convexity of the relative entropy.

For a finite quantum system $\mathcal{A}$ with Hamiltonian $H \in \mathcal{A}^{sa}$ the free energy functional at inverse temperature β is defined as

$$F(\omega) = \omega(H) - \frac{1}{\beta} S(\omega)$$

for a state ω. The canonical state φ minimizing the free energy functional possesses the density

$$\frac{e^{-\beta H}}{\operatorname{Tr} e^{-\beta H}}.$$

The perturbed state $[\varphi^h]_\beta$ is the canonical state for the Hamiltonian $H + h$. Hence $[\varphi^h]_\beta$ represents the equilibrium state of the perturbed physical system in which the energy of each state σ has been increased by $\sigma(h)$. The density of $[\varphi^h]_\beta$ is

$$\frac{e^{-\beta(H+h)}}{\operatorname{Tr} e^{-\beta(H+h)}}.$$

In the finite quantum case $[\varphi^h]_\beta$ minimizes the functional

$$\omega \mapsto \beta^{-1} S(\varphi, \omega) + \omega(h)$$

and following Donald ([15]) we use this property as a definition in the general case. Due to the simple transformation

$$[\varphi^h]_\beta = [\varphi^{\beta h}]_1$$

the value $\beta = 1$ will be fixed mostly and we write $[\varphi^h]$ instead of $[\varphi^h]_1$.

Let $\mathcal{M}$ be a von Neumann algebra acting on a Hilbert space $\mathcal{H}$. The normal state space of $\mathcal{M}$ will be denoted by $\Sigma_*(\mathcal{M})$. We are going to deal with lower-bounded self-adjoint operators affiliated with $\mathcal{M}$. Let us recall that a self-adjoint operator is affiliated with $\mathcal{M}$ if all its spectral projections belong to $\mathcal{M}$. It will be allowed that $+\infty$ is an eigenvalue of the operator. More formally, an extended-valued lower-bounded (self-adjoint) operator (affiliated with $\mathcal{M}$) can be defined by a spectral decomposition.

$$\int_c^\infty \lambda dE_\lambda + \infty p\,,$$

where $c \in \mathbb{R}$, p is a certain projection in $\mathcal{M}$ and E_λ is a spectral measure with values in the von Neumann algebra $p^\perp \mathcal{M} p^\perp$ and given on the interval $[c, \infty)$.

Now we are ready to define $[\varphi^h]$ for a normal state φ of $\mathcal{M}$ and for a generalized operator $h \in \mathcal{M}^{\mathrm{ext}}$. If

$$d \equiv \inf\{F(\omega) \equiv S(\omega, \varphi) + h(\omega) : \omega \in \Sigma_*(\mathcal{M})\}$$

is finite then due to Theorem 11 there is a unique state $[\varphi^h]$ of $\mathcal{M}$ satisfying

$$S([\varphi^h], \varphi) + h([\varphi^h]) = d.$$

Hence the perturbed state $[\varphi^h]$ is determined as the unique minimizer of the weakly lower semi-continuous functional F. (If $d = \infty$ then $[\varphi^h]$ is not defined.) Note that if $h(\varphi)$ is finite then $[\varphi^h]$ exists and thanks to $S([\varphi^h], \varphi) < +\infty$ the state $[\varphi^h]$ is always normal. Moreover

$$\operatorname{supp}[\varphi]^h \leq \operatorname{supp}\varphi.$$

It follows also from the definition that

$$[\varphi]^h = [\varphi]^{qhq} \tag{20}$$

whenever q is a projection in $\mathcal{M}$ with $q \geq \operatorname{supp}\varphi$. Property (20) has the consequence that one can easily pass to an algebra on which φ is faithful. One should only replace h by $(\operatorname{supp}\varphi)h(\operatorname{supp}\varphi)$.

Let us discuss first a simple example about state perturbation which is related to the paper [16]. Let $\mathcal{A}$ be a finite quantum system and let $p \in \mathcal{A}$ be a projection. If $k \in \mathcal{A}^{sa}$ such that $\operatorname{supp} k \perp p$ then $k + \infty p$ is a lower-bounded extended-valued operator. Set e^{-H} for the density of a state φ of $\mathcal{A}$ and define

$$\psi(a) = \frac{\operatorname{Tr} aqe^{-qHq-k}}{\operatorname{Tr} qe^{-qHq-k}} \qquad (a \in \mathcal{A}),$$

where q makes shorter $I - p$. We are going show that $[\varphi^{k+\infty p}] = \psi$. One computes that

$$S(\psi, \varphi) = -\log \operatorname{Tr} qe^{-qHq-k} - \psi(k)\,.$$

Hence it suffices to show the inequality

$$S(\omega,\varphi)+\omega(k)+\infty\omega(p)\geq -\log \operatorname{Tr} qe^{-qHq-k} \tag{21}$$

for every state ω of $\mathcal{A}$. This trivially holds if $\omega(p)\neq 0$, so we may assume $\operatorname{supp}\omega \leq q$. We have

$$S(\omega,\varphi)+\omega(h)\geq -\log \operatorname{Tr} e^{-H-h} \tag{22}$$

for every $h\in\mathcal{A}^{sa}$. Now replace h here by $-H+qHq+k+np$ ($n\in\mathbb{N}$) and (22) becomes

$$S(\omega,\varphi)+\omega(k)\geq -\log \operatorname{Tr} e^{-qHq-k}e^{-np} = -\log \operatorname{Tr}\,(qe^{-qHq-k}+e^{-n}p).$$

Taking the limit $n\to\infty$ we conclude (21). Notice in this example that $\operatorname{supp}[\varphi^{k+\infty p}]=q$ and it needs not to be equal to $\operatorname{supp}\varphi$.

Although the perturbed state was formulated in the the setting of von Neumann algebras, the same definition works if ψ is an arbitrary state of a C*-algebra $\mathcal{A}$ and $h=h^*\in\mathcal{A}$. Observe that $[\psi^h]=\psi$ for a pure state ψ.

We define

$$c_\beta(\varphi,h)=\inf\{\tfrac{1}{\beta}S(\omega,\varphi)+h(\omega):\omega\in\Sigma_*(\mathcal{M})\}$$

given a state $\varphi\in\Sigma_*(\mathcal{M})$ and an extended-valued lower-bounded operator h affiliated with the von Neumann algebra $\mathcal{M}$. Then $c_\beta(\varphi,h)<+\infty$ is equivalent to the claim that $[\varphi^h]_\beta$ exists, and in this case

$$c_\beta(\varphi,h)=\tfrac{1}{\beta}S([\varphi^h],\varphi)+[\varphi^h](h).$$

We write $c(\varphi,h)$ with the understanding that $\beta=1$.

The relation $[\varphi^{h+k}]=[[\varphi^h]^k]$ is very much expected and termed chain rule.

Theorem 14. ([15]) Let $k=k^*\in\mathcal{M}$ and h be an extended-valued lower-bounded self-adjoint operator affiliated with the von Neumann algebra $\mathcal{M}$. Let φ be a normal state of $\mathcal{M}$ such that $c(\varphi,h)<\infty$. Then $[\varphi^{h+k}]=[[\varphi^h]^k]$, $c(\varphi,h+k)=c([\varphi^h],k)+c(\varphi,h)$ and for every $\sigma\in\Sigma_*(\mathcal{M})$ the equality

$$S(\sigma,[\varphi^h])+\sigma(k)+c(\varphi,h)=c(\varphi,h+k)+S(\sigma,[\varphi^{h+k}]) \tag{23}$$

holds.

Let $\mathcal{A}$ be a C*-algebra with a fixed state φ for a while. The relation between the functional $\omega\mapsto S(\omega,\varphi)$ defined on the state space and the function $c\mapsto c(\varphi,k)$ defined on the self-adjoint part of $\mathcal{A}$ may be clarified by means of some standard convex analysis. Let V and U two linear spaces placed into duality by a pairing $<\,.\,,\,.\,>$. The spaces V and U will be endowed with the topologies $\sigma(V,U)$ and $\sigma(U,V)$, respectively. Let F be a function of V into $\mathbb{R}\cup\{+\infty\}$. Then the formula

$$F^*(u)=\sup\{<v,u>-F(v):\,v\in V\} \tag{24}$$

defines a function F^* of U into $\mathbb{R} \cup \{+\infty\}$ which is called the conjugate function of F. (F^* is called the Legendre transform of F as well.) It is obvious that in (24) we can confine ourselves to those v such that $F(v)$ is finite. This process can be repeated and leads to the second conjugate F^{**}, which is a function of V into $\mathbb{R} \cup \{+\infty\}$:

$$F^{**}(v) = \sup\{\langle v, u\rangle - F^*(u) : \, u \in U\} \tag{25}$$

A basic result in convex analysis says that if F is a lower semi-continuous convex function then $F = F^{**}$.

Choose now $U = \mathcal{A}^{sa}$ and $V = \mathcal{A}^*_h$, the real linear space of hermitian functionals of $\mathcal{A}$. With respect to the duality

$$\langle \nu, k\rangle = \nu(k) \quad (k \in \mathcal{A}^{sa} \text{ and } \nu \in \mathcal{A}^*_h)$$

the conjugate function of

$$F(\omega) = \begin{cases} \beta^{-1} S(\omega, \varphi) & \text{if } \omega \text{ is a state} \\ +\infty & \text{otherwise.} \end{cases}$$

is nothing else but $-c_\beta(\varphi, k)$. Since F is lower semi-continuous and convex we have

$$\beta^{-1} S(\omega, \varphi) = \sup\{-\omega(k) + c_\beta(\varphi, k) : k \in \mathcal{A}^{sa}\}. \tag{26}$$

In mathematical physics one often meets the problem of determining the asymptotic behavior, for large n, a sequence

$$\mathrm{Tr}\,(\exp H_n) \qquad\qquad (n = 1, 2, \ldots),$$

where (H_n) is a sequence of self-adjoint operators acting on some Hilbert space. The problem arises, for example, in statistical mechanics of quantum lattice systems. In the one dimensional case to each $n \in \mathbb{Z}$ a copy $\mathcal{B}_n$ of a C*-algebra is associated. If $\mathcal{J} \subset \mathbb{Z}$ then $\mathcal{B}_\mathcal{J}$ denotes $\otimes_{n\in\mathcal{J}} \mathcal{B}_n$ and $\mathcal{B}$ will be written for $\mathcal{B}_\mathbb{Z}$. The n-fold tensor product $\mathcal{B}_{[1,n]}$ represents n identical systems interacting with each other. If the local Hamiltonian $H_n \in \mathcal{B}_{[1,n]}$ can be diagonalized then

$$\lim_{n\to\infty} \frac{1}{n\beta} \log \mathrm{Tr}\, e^{-\beta H_n} \tag{27}$$

may be treated by powerful probabilistic methods. Roughly speaking, the sequence (H_n) is modelled by a sequence of random variables and the large deviation technique is used.

Let (μ_n) be a sequence of measures on a (sufficiently regular) topological space X. The sequence is said to obey the large deviation principle with constants (V_n) and rate function $\mathcal{L} : X \to \mathbb{R}^+ \cup \{\infty\}$ if the following conditions are satisfied:

(i) The level sets $\{x \in X : \mathcal{L}(x) \le b\}$ are compact for each $b \in \mathbb{R}^+$.

(ii) $\limsup_n V_n^{-1} \log \mu_n(F) \le -\inf\{\mathcal{L}(x) : x \in F\}$ whenever $F \subset X$ is a closed set.

(iii) $\liminf_n V_n^{-1} \log \mu_n(G) \geq -\inf\{\mathcal{L}(x) : x \in G\}$ whenever $G \subset X$ is an open set.

The strong relation of the large deviation principle to the thermodynamic limit (27) will be transparent from the next result, due to Varadhan ([22]). The large deviation principle is equivalent to the limit relation

$$\lim_{n\to\infty} \frac{1}{V_n} \int e^{V_n f(x)} d\mu_n(x) = \sup\{f(x) - \mathcal{L}(x) : x \in X\} \tag{28}$$

for every continuous bounded function $f : X \to \mathbb{R}$. The mean field interaction is one of the simplest example when the large deviation theory is not suitable for the determination of the limit (27). Let the local Hamiltonian be of the form

$$H_{[m,n]} = \sum_{i=m}^{n} h_i + \frac{1}{n-m+1} \sum_{i,j=m}^{n} x_i x_j \,, \tag{29}$$

where $h_i \in \mathcal{B}_i$ and $x_i \in \mathcal{B}_i$ are copies of some self-adjoint operators $h_1 \in \mathcal{B}_1$ and $x_1 \in \mathcal{B}_1$, respectively. The expression mean field reflects the second term of (29). If the single system is described by a finite dimensional C*-algebra $\mathcal{B}_1$ then the local free energy density at the inverse temperature β is given us

$$F_n(\beta) = \frac{1}{n} \log \mathrm{Tr}\, \exp\left(-\beta \sum_{i=1}^{n} h_i - \frac{\beta}{n} \sum_{i,j=1}^{n} x_i x_j\right). \tag{30}$$

Let φ be the state of $\mathcal{B}_1$ which possesses the density

$$D_\varphi = \frac{\exp(-\beta h)}{\mathrm{Tr}\, \exp(-\beta h)} \tag{31}$$

and let φ_∞ be the corresponding product state of $\mathcal{B}$. In an other notation we have

$$F_n(\beta) = -\frac{\beta}{n} c(\varphi_\infty, \frac{1}{n} \sum x_i x_j, \beta). \tag{32}$$

It is not a real restriction if we put $\beta = 1$ in the sequel. In order to make contact with probability theory, let us assume for a while that $\mathcal{B}$ is abelian, or equivalently $[h, x] = 0$. Then φ_∞ corresponds to a product measure which is the joint distribution of the sequence (x_i) of identically distributed independent random variables. In this probabilistic translation (30) becomes

$$\frac{1}{n} \log \int \exp(-n(\frac{1}{n} \sum_{i=1}^{n} x_i)^2) d\mu. \tag{33}$$

If μ_n is the distribution of $\frac{1}{n}\sum_{i=1}^{n} x_i$ then the sequence (μ_n) obeys the large deviation principle with the rate function $\mathcal{L}$ given below.

$$L(u) = \int \exp u x_1 d\mu(u) \quad \text{and} \quad \mathcal{L}(x) = \sup\{ux - \log L(u) : u \in \mathbb{R}\}. \tag{34}$$

(The large deviation result for the means of identically distributed independent random variables is a reformulation of a classical theorem proved by Cramér.) Typically $\mathcal{L}$ is a convex function which vanishes at the expectation value m of x_1. One knows from the law of large numbers that if $F \subset \mathbb{R}$ is a closed set such that $m \notin F$ then

$$\mathrm{Prob}\left(\frac{1}{n}\sum_{i=1}^{n} x_i \in F\right) \to 0$$

as $n \to \infty$. Condition (ii) of the large deviation principle tells us that this convergence is exponentially fast and its speed is shown by the rate function.

From the above discussion we want to conclude that thermodynamics of mean field interactions is strongly related to the large deviation principle in the case $[h, x] = 0$. All this will serve as a motivation for us to develop an analogue theory in the pure quantum case $[h, x] \neq 0$. The functional c from (32) will take the place of "$\log \int \exp$" here and the guideline will be version (28) of the large deviation principle.

Let γ denote the right shift automorphism of the C*-algebra $\mathcal{B}$. The limit of the means

$$s_n(a) \equiv \frac{1}{n}\sum_{i=1}^{n-1} \gamma^i(a) \qquad\qquad (a \in \mathcal{B}^{sa}, \rangle n \in \mathbb{N})$$

is the subject of ergodic theorems. The C*-algebraic ergodic theorem we are going to prove needs the following lemma.

Lemma 15. Let $\mathcal{L}$ be a polynomial and $K \in \mathbb{R}^+$. Then for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(A) - f(B)| < \varepsilon$$

whenever $A, B \in \mathcal{B}^{sa}$, $||A - B|| < \delta$ and $||A||, ||B|| \leq K$.

Theorem 16. Let $g : [s,t] \to R$ be a continuous function and $b \in \mathcal{B}^{sa}$ such that $\mathrm{Sp}\, b \subset [s,t]$. Then

$$\lim_{n\to\infty} \omega(g(s_n(b))) \tag{35}$$

exists for every γ-invariant state ω of $\mathcal{B}$.

Proof. It is a plain consequence of the Weierstrass approximation theorem that it suffices to prove the existence of the limit (35) when g is a polynomial. A further possibility of reduction is based on Lemma 15. Since the local elements are norm dense in the algebra $\mathcal{A}$, we may assume that a is local, that is, there exists $k \in \mathbb{N}$ such that $b \in \mathcal{B}_{[1,k]}$.

First we consider the case $k = 1$. Then the sequence $s_n(b)$ is in the abelian C*-algebra $\ldots \otimes \mathcal{C} \otimes \mathcal{C} \otimes \ldots$, where $\mathcal{C}$ is the subalgebra generated by a in $\mathcal{B}_1$. According to the representation theorem of abelian C*-algebras, $\mathcal{C}$ may be viewed as an algebra of continuous functions on the spectrum of a and $\omega | \ldots \otimes \mathcal{C} \otimes \mathcal{C} \otimes \ldots$ corresponds to an integration with respect to a measure μ. By the individual ergodic theorem $s_n(b)$

converges μ-almost everywhere and so does $g(s_n(b))$ for any measurable bounded function g. The Lebesgue theorem tells us that

$$\int g(s_n(b))d\mu \equiv \omega(g(s_n(b))$$

converges as well us $n \to \infty$. This completes the proof for $b \in \mathcal{A}_1^{sa}$.

The general case $b \in \mathcal{B}_{[1,k]}^{sa}$ will be reduced to the above discussed abelian case by changing the localization. For an integer $\ell \geq k$ we set

$$b_\ell \equiv \sum_{i=0}^{\ell-k} \gamma^i(b) \in \mathcal{B}_{[1,\ell]}$$

and write the arbitrary n in the form $u \cdot \ell + r$, where u and r are integers and $0 \leq r < \ell$. Then

$$\Big\| \sum_{j=0}^{n-1} \gamma^j(k) - \sum_{j=0}^{n-1} \gamma^{j\ell}(b_\ell) \Big\| \leq (uk + r)||b||. \tag{36}$$

This gives that

$$\Big\| s_n(k) - \frac{1}{n} \sum_{j=1}^{n-1} \gamma^{j\ell}(b_\ell/\ell) \Big\|$$

can be arbitrary small if ℓ is big enough (with respect to k) and n is big too. According to Lemma 15

$$||f(s_n(b)) - f(\frac{1}{n} \sum_{j=1}^{n-1} \gamma^{j\ell}(b_\ell/\ell)||$$

is small as well. On the other hand, for fixed ℓ

$$\omega(f(\frac{1}{n} \sum_{j=0}^{n-1} \gamma^{j\ell}(b_\ell/\ell))$$

is a Cauchy sequence. Indeed, the sequence $\gamma^{j\ell}(b_\ell/\ell)$ is built from pairwise commuting operators and the above abelization argument works. In this way we are able to conclude that $\omega(f(s_n(b)))$ is a Cauchy sequence which was to be proven.

Theorem 17. ([17]) For $b \in \mathcal{B}_1$ and a continuous function f the variational formula

$$\lim_{n\to\infty} \frac{1}{n} c(\varphi_\infty, nf(s_n(b))) = \inf\{f(\psi(b)) + S(\psi, \varphi_1) \ : \ \psi \in \Sigma(\mathcal{B}_1)\}$$

holds.

Theorem 17 is a typical example of what we shall call perturbational limit theorem. Let $\mathcal{A}$ and $\mathcal{C}$ be C*-algebras abd ψ a state of $\mathcal{C}$. Assume that positive unital mappings $\alpha_n : \mathcal{A} \to \mathcal{C}$ $(n \in \mathbb{N})$ are given and that $\psi \circ \alpha_n$ is independent of n. Motivated by

Theorem 17 (as well as large deviation theory, in particular Varadhan's theorem) the perturbational limit principle is said to be held if

$$\lim_{n\to\infty}\frac{1}{n}c(\psi, nf(\mathcal{A}_n(a))) = \inf\{E^f_\omega(a) + I(\omega)\ :\ \omega \in \Sigma(\mathcal{A})\} \tag{37}$$

for every $a = a* \in \mathcal{A}$, for every continuous function $f : \mathbb{R} \to \mathbb{R}$ and for a certain weak* lower semi-continuous functional $I : \Sigma(\mathcal{A}) \to \mathbb{R}^+ \cup \{\infty\}$. The latter functional I will be called rate functional.

To view Theorem 17 as a perturbational limit theorem one chooses $\mathcal{A} = \mathcal{B}_1$, $\mathcal{C} = \mathcal{B}$, $\psi = \varphi_\infty$ and

$$\mathcal{A}_n = n^{-1}(\mathrm{id} + \gamma + \ldots + \gamma^{n-1})|\mathcal{B}_1 \quad (n \in \mathbb{N}).$$

The rate function is the relative entropy in this case.

The object $(\mathcal{B},\ \mathcal{B}_1,\ \gamma,\ \varphi_\infty)$ in Theorem 17 is an independent process in the setting of C*-algebras which might be viewed at different levels. When one is interested in the collective behavior of the perturbational limit for every $b \in \mathcal{B}$ then the algebra $\mathcal{A}$ must be chosen larger and the rate functional will be defined on the state space of the enlarged algebra. (In the terminology of large deviation theory this means result of level III, see [22].)

Let ω be a translation invariant state of $\mathcal{B}$. We benefit from the superadditivity (19) of the relative entropy as follows.

$$\begin{aligned}
&S(\omega|\mathcal{B}_{[1,n+m]}, \varphi_\infty|\mathcal{B}_{[1,n+m]})\\
&\quad \geq S(\omega|\mathcal{B}_{[1,n]}, \varphi_\infty|\mathcal{B}_{[1,n]}) + S(\omega|\mathcal{B}_{[n+1,n+m]}, \varphi_\infty|\mathcal{B}_{[n+1,n+m]})\\
&\quad = S(\omega|\mathcal{B}_{[1,n]}, \varphi_\infty|\mathcal{B}[1,n]) + S(\omega|\mathcal{B}_{[1,m]}, \varphi_\infty|\mathcal{B}_{[1,m]}).
\end{aligned}$$

This yields that the numbers $t_n = S(\omega|\mathcal{B}_{[1,n]}, \varphi_\infty|\mathcal{B}_{[1,n]})$ form a superadditive sequence (that is, $t_{n+m} \geq t_n + t_m$) and so

$$S_M(\omega, \varphi_\infty) \equiv \lim_{n\to\infty}\frac{1}{n}S(\omega_n, \varphi_n) = \sup\{\frac{1}{n}S(\omega_n, \varphi_n) : n \in \mathbb{N}\}. \tag{38}$$

Since $S(\omega_n, \varphi_n)$ is a weak*-lower semi-continuous function of ω, it follows immediately that $S_M(\omega, \varphi_\infty)$ is lower semi-continuous as well. The quantity $S_M(\infty, \varphi_\infty)$ will be called mean relative entropy (or relative entropy density).

Theorem 18. The mean relative entropy $\omega \mapsto S_M(\omega, \varphi_\infty)$ is a weak*-lower semi-continuous affine functional on $\Sigma^\gamma(\mathcal{B})$.

Proof. It is remained to show the affinity. Theorem 9 tells us that

$$0 \geq S(\lambda\omega^1_n + \mu\omega^2_n, \varphi_n) - \lambda S(\omega^1_n, \varphi_n) - \mu S(\omega^2_n, \varphi_n) \geq \lambda\log\lambda + \mu\log\mu$$

if $\lambda + \mu = 1$. Dividing by n and letting $n \to \infty$ we obtain that

$$S(\lambda\omega^1 + \mu\omega^2, \varphi_\infty) = \lambda S(\omega^1, \varphi_\infty) + \mu S(\omega^2, \varphi_\infty).$$

At this point we can anticipate what kind of analogue of the variational formula (28) must hold in the quantum case. Let $b \in \mathcal{B}^{sa}_{[1,k]}$ be a local operator in the quasi-local algebra $\mathcal{B}$. By the definition of the function c we have

$$\begin{aligned} & c(\varphi_\infty, nf(s_n(b))) = c(\varphi_\infty|\mathcal{B}_{[1,n+k]}, nf(s_n(k))) \\ & \leq \omega(nf(s_n(a))) + S(\omega|\mathcal{B}_{[1,n+k]}, \varphi_\infty|\mathcal{B}_{[1,n+k]}) \end{aligned}$$

for any state ω of $\mathcal{B}$. Dividing this inequality by n and letting $n \to \infty$ we arrive at the following relation.

$$\limsup_{n\to\infty} \frac{1}{n} c(\varphi_\infty, nf(s_n(b))) \leq \lim_{n\to\infty} c(\varphi_\infty, f(s_n(b))) + S_M(\omega, \varphi_\infty) \tag{39}$$

which holds for every translation invariant state ω. It is our aim to show that the form of (39) is an analogue of Varadhan's theorem. While in Theorem 17 the perturbational limit was stated for self-adjoint elements of $\mathcal{B}_1$ in the next theorem all quasi-local elements are included. Correspondingly, the variational expression is over a larger set.

Theorem 19. For $b \in \mathcal{B}^{sa}$ and for a continuous real function f the variational formula

$$\begin{aligned} & \lim_{n\to\infty} \frac{1}{n} c(\varphi_\infty, f(s_n(b))) \\ & \qquad = \inf\{ \lim_{n\to\infty} \omega(f(s_n(b)) + S_M(\omega, \varphi_\infty) : \omega \in \Sigma^\gamma(\mathcal{B})\} \end{aligned}$$

holds, where the notation $\Sigma^\gamma(\mathcal{B})$ stands for the translation invariant states of $\mathcal{B}$.

Theorem 19 supplies us with a new example of Legendre transformation. Choose $U = \mathcal{B}^{sa}$ and let V be the real linear space of hermitian γ-invariant functionals of $\mathcal{B}$. With respect to the duality

$$\langle \nu, b\rangle = \lim_{n\to\infty} \nu(f(s_n(b)))$$

the conjugate function of

$$F(\nu) = \begin{cases} S_M(\nu, \varphi_\infty) & \text{if } \nu \text{ is a } \gamma\text{-invariant state} \\ +\infty & \text{otherwise.} \end{cases}$$

is

$$F^*(b) = -\lim_{n\to\infty} \frac{1}{n} c(\varphi_\infty, nf(s_n(b)))$$

Improving the method of [17] variatinal formulas were obtained in order to establish the Gibbs variational principle for different kind of mean field models in [19] and [24], see also the contribution of R.F. Werner in this volume.

Since the selection of the results concerning quantum entropy follows the interest of the present author, readers are suggested to look at other reviews of the subject, for example [20] and [21].

Acknowledgment. This paper was written during the author's stay at the University of Heidelberg. The financial support of SFB 123 and the very kind hospitality of Professor W. von Waldenfels are acknowledged.

References

[1] Neumann, J. von (1927): “Thermodynamik quantenmechanischer Gesamheiten”, Gött. Nachr. pp. 273–291.

[2] Ohya, M., Petz, D. (1996): “Notes on quantum entropy”, Studia Sci. Math. Hungar. **31**, 423–430.

[3] Jaynes, E.T. (1956): “Information theory and statistical mechanics. II”, Phys. Rev. **108**, pp. 171–190.

[4] Belavkin, V. P., Staszewski, P. (1982): “C*-algebraic generalization of relative entropy and entropy”, Ann. Inst. Henri Poincaré, Sec. A **37**, pp. 51–58.

[5] Petz, D. (1992): “Characterization of the relative entropy of states of matrix algebras”, Acta Math. Hungar. **59**, 449–455.

[6] Hiai, F., Petz, D. (1991): “The proper formula for relative entropy and its asymptotics in quantum probability”, Commun. Math. Phys. **143**, pp. 99–114.

[7] Araki, H. (1976): “Relative entropy for states of von Neumann algebras, Publ. PIMS, Kyoto Univ. **11**, pp. 809–833.

[8] Uhlmann, A. (1977): “Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory”, Commun. Math. Phys. **54**, pp. 21–32.

[9] Petz, D. (1986): “Properties of the relative entropy of states of a von Neumann algebra”, Acta Math. Hung. **47**, pp. 65–72.

[10] Kosaki, H. (1986): “Relative entropy for states: a variational expression”, J. Operator Theory **16**, pp. 335–348.

[11] Petz, D. (1991): “On certain properties of the relative entropy of states of operator algebras”, Math. Zeitsch., **206**, pp. 351–361.

[12] Donald, M.J. (1985): “On the relative entropy”, Commun. Math. Phys. **105**, pp. 13–34.

[13] Hiai, F., Ohya, M. Tsukada, M. (1981): “Sufficiency, KMS condition and relative entropy in von Neumann algebras”, Pacific J. Math. **96**, pp. 99–109.

[14] Connes, A. (1985): “Entropie de Kolmogoroff-Sinai et mècanique statistique quantique”, C.R. Acad. Sci. Paris Ser. I. Math. **301**, pp. 1–6.

[15] Donald, M.J. (1990): “Relative Hamiltonian which are not bounded from above”, J. Functional Anal. **91**, pp. 143–173

[16] Raggio, G.A., Werner, R.F. (1990): “Minimizing entropy in a face”, Lett. Math. Phys. **19**, pp. 7–14.

[17] Petz, D., Raggio, G.A., Verbeure, A (1989): “Asymptotics of Varadhan-type and the Gibbs variational principle”, Commun. Math. Phys. **121**, pp. 271–282.

[18] Petz, D. (1990): “First steps towards a Donsker and Varadhan theory in operator algebras”, in *Quantum probability and applications V* ed. by L. Accardi and W. von Waldenfels (Lecture Notes in Math. **1442**, Springer), pp. 311–319.

[19] Raggio, G.A., Werner, R.F. (1989): “Quantum statistical mechanics of general mean field systems, Helvetica Phys. Acta **62**, pp. 980–1003

[20] Wehrl, A. (1978): “General properties of entropy”, Rev. Mod. Physics **50**, pp. 221–260

[21] Araki, H. (1987): “Recent progress on entropy and relative entropy”, in *VIIIth Int. Cong. on Math. Phys.* (World Sci. Publishing, Singapore) pp. 354-365

[22] Ellis, R.S. (1985): *Entropy, large deviations, and statistical mechanics* (Springer)

[23] Umegaki, H. (1962): “Conditional expectations in an operator algebra IV (entropy and information)”, Kodai Math. Sem. Rep. **14**, pp. 59–85

[24] Raggio, G.A., Werner, R.F. (1991): “The Gibbs variational principle for inhomogeneous mean field systems”, Helv. Phys. Acta **64**, pp. 633–667.