REMARKS ON THE RELATION BETWEEN QUANTUM AND CLASSICAL ENTROPY AS PROPOSED BY A. WEHRL

ARMIN UHLMANN

Department of Physics and NTZ, Karl-Marx-University, Leipzig, G.D.R.

(*Received December 4, 1978*)

Using an inequality of Lieb one can compute a probability distribution which is less mixed than any $(z|\omega z)dz$, where z labels the points of phase space, ω is a density matrix, and $(z|$ the appropriate coherent state.

According to Wehrl one associates to every state of a quantum system characterized by a density matrix ω a "classical entropy" build up with help of the probability measure $(z|\omega z)dz$. Here $(z|$ denotes the coherent state labelled by the point z of the phase space, and dz is the Liouville measure. In [1] and [2] Wehrl stated a number of interesting properties of this classical entropy and its relation to quantum entropy. In [3] Lieb proved $S^{\mathrm{cl}} \geqslant 1$ in appropriate units (Boltzmann's constant equal one), an inequality conjectured by Wehrl.

In this note we begin with the technical remark that the map $\omega \to (z|\omega|z)dz$ is dual to a "quantization map" and can be easily extended to singular states. Then we comment on the fact that the mentioned map is mixing enhancing. From this, using Lieb's inequality [3], we arrive at an *a priori* lower bound for F^{cl} where F is any concave function. We conclude with a further remark on singular states.

I. Let us denote the points of the phase space $\boldsymbol{R}^{2n}$ by

$$z = \{q, p\} = \{q_1, ..., q_n, p_1, ..., p_n\}$$

and the Liouville measure (i.e. the Lebesgue measure), multiplied by h^{-n} by dz:

$$dz := h^{-n} d^n q d^n p, \qquad h = \text{Planck's constant.}$$

For the "algebra of observables" we choose the W^*-algebra $\boldsymbol{M}$ of all (equivalence classes of) bounded and dz-measurable functions defined on the phase space $\boldsymbol{R}^{2n}$

$$\boldsymbol{M} := L^{\infty}(\boldsymbol{R}^{2n}, dz)$$

We denote the points of $\boldsymbol{R}^n$ by $x = \{x_1, \ldots, x_n\}$, its Lebesgue measure by dx, and by $\boldsymbol{H}$ the Hilbert space

$$\boldsymbol{H} := L^2(\boldsymbol{R}^n, dx)$$

The "algebra of observables" of the quantum system is

$$\boldsymbol{B} := \boldsymbol{B}(\boldsymbol{H}),$$

the algebra of bounded operators acting on $\boldsymbol{H}$.

Consider now the "quantization map" $\phi\colon \boldsymbol{M} \to \boldsymbol{B}$ defined by

$$\phi f := \int f(z)|z)(z|\,dz \tag{1}$$

where $|z)(z|$ denotes the 1-dimensional projection operator onto the coherent state $|z)$ of $\boldsymbol{H}$. Explicitly, $|z)$ is given as the L^2-function depending on x

$$|z) := (h/2)^{-n/4}\exp\{[-(x-q)^2/2 + ipx]/(2\pi h)\}. \tag{2}$$

Because there is no danger of confusion we denote the identity elements of both, $\boldsymbol{M}$ and $\boldsymbol{B}$, by 1, and the C^*-norms in these algebras (the a.e. supremum and the operator norm) by $\|\cdot\|$. The following properties of ϕ are well known:

(i) ϕ is contracting, i.e. $\|f\| \geqslant \|\phi f\|$.

(ii) $\phi 1 = 1$.

(iii) ϕ is positive, i.e. from $f \geqslant 0$ a.e. follows $\phi f \geqslant 0$ in the operator sense.

(iv) If $f \in \boldsymbol{M}$ is dz-integrable, $f \in L^1(\boldsymbol{R}^{2n}, dz)$, then ϕf is of trace class and $\int f dz = \operatorname{Tr}(\phi f)$.

A *state*, ω, of a W^*-algebra is a linear functional on this algebra which is non-negative for positive elements of the algebra and which takes the value 1 on the identity element of the algebra.

Now if $\omega : a \mapsto \omega(a)$, $a \in \boldsymbol{B}$, is a state of $\boldsymbol{B}$, we can define a state $\phi^*\omega$ of $\boldsymbol{M}$ by

$$(\phi^*\omega)(f) = \omega(\phi f) \qquad \text{for all } f \in \boldsymbol{M}. \tag{3}$$

Due to (ii) and (iii) $\phi^*\omega$ is a state of $\boldsymbol{M}$. Further, one can define ϕ^* in the same manner for weights (in the sense of A. Connes) and especially property (iv) shows that ϕ^* maps the trace of $\boldsymbol{B}$ onto the Liouville measure, $\phi^* : \operatorname{Tr}(\ ) \to \int(\ )dz$.

Before making to the first statement we introduce the following convention: for every density matrix ρ we denote the associated normal state of $\boldsymbol{B}$ also by ρ, thus writing $\operatorname{Tr} b\rho = \rho(b)$ for all $b \in \boldsymbol{B}$.

THEOREM 1. *Let ω be a state of $\boldsymbol{B}$. If ω is singular, then $\phi^*\omega$ is a singular state of $\boldsymbol{M}$. If ω is normal and given by the density matrix ω, then $\phi^*\omega$ is given by the measure* $(z|\omega|z)dz$.

The predual of M is $L^1(R^{2n}, dz)$ and its image under ϕ is in the trace class, i.e. in the predual of $\boldsymbol{B}$, and hence $(\phi^*\omega)(f) = \omega(\phi f) = 0$ if f is in the predual of $\boldsymbol{M}$ and ω is singular on $\boldsymbol{B}$. This proves the first part of the assertion. The second one is well known and is only a rephrasal of

$$\int f(z)(z|\omega|z)dz = \int f(z)\mathrm{Tr}\{\omega|z)(z|\}dz = \mathrm{Tr}\{\omega\cdot\phi f\}.$$

Let us recall for clarity that $\phi^*\omega$ with singular ω is not given by a measure supported on $\boldsymbol{R}^{2n}$. On the other hand, $\phi^*\omega$ with normal ω is not only a probability measure on $\boldsymbol{R}^{2n}$ but even absolutely continuous with respect to dz. This is characteristic of *all* normal states of $\boldsymbol{M}$ because two functions, which differ only on a set of dz-measure zero, are identified.

II. Now we discuss the meaning of the statement "ϕ^* is *mixing enhancing*".

We consider first the interesting case of normal states and then we add some simple remarks concerning the general situation.

Let $s \mapsto F(s)$ be a real-valued function defined on $[0,\infty)$ which is continuous, concave and fulfils $F(0) = 0$. Together with concavity the last condition guarantees that $F(s)/s$ is decreasing for positive s.

For every normal state ρ of $\boldsymbol{B}$, given by the density matrix ρ, we define

$$F^{q}(\rho) := \mathrm{Tr}\,F(\rho) \tag{4}$$

if $F(\rho)$ is trace, class,and $\lim\limits_{s\to 0} F(s)/s$ otherwise. Now going over to F^{cl} we first note the dependence of this construction on the measure dz which is not distinguished by the algebraic structure of $\boldsymbol{M}$ – in contrast to the trace functional on $\boldsymbol{B}$. Hence F^{cl} depends not only on the structure of the state space of $\boldsymbol{M}$ but also on the "additional" given Liouville measure dz. Let us now consider a probability measure, dv, on $\boldsymbol{R}^{2n}$. To define $F^{cl}(dv)$ consider an arbitrary decomposition of $\boldsymbol{R}^{2n}$, $\boldsymbol{R}^{2n} = \bigcup_j N_j$, into non-intersecting measurable subsets $N_1, N_2, \ldots$ of the phase space. Assume further for all k

$$0 < \int_{N_k} dz < \infty.$$

Then $F^{cl}(dv)$ is the infimum of the numbers

$$\sum_j u_j F(v_j/u_j) \qquad \text{with } u_k = \int_{N_k} dz \text{ and } v_k = \int_{N_k} dv.$$

The infimum is taken over all possible decompositions of $\boldsymbol{R}^{2n}$. It is known that these coarse grained quantities decrease (for concave F) if the coarse graining becomes finer.

If dv is absolutely continuous with respect to dz, and thus yields a normal state $f \to \int f dv$ of $\boldsymbol{M}$, then

$$F^{\text{cl}}(dv) = \int F(dv/dz)dz. \tag{5}$$

This applies to $\phi^*\omega$ if ω is a normal state of $\boldsymbol{B}$.

As Wehrl has shown, we have for all normal states ω of $\boldsymbol{B}$

$$F^{\text{cl}}(\phi^*\omega) \geqslant F^{\text{q}}(\omega). \tag{6}$$

This we express by saying that "ϕ^* is mixing enhancing". This property is equivalent to the one in the following

THEOREM 2. *Let ω be a normal state of $\boldsymbol{B}$ and let us denote by $\lambda_1 \geqslant \lambda_2 \geqslant \ldots$ the eigenvalues of the density matrix ω. For every positive real s between the integers m and $m+1$ we have*

$$\int_N (z|\omega|z)dz \leqslant \lambda_1 + \ldots + \lambda_m + (s-m)\lambda_{m+1} \tag{7}$$

if only the Liouville measure of N is smaller than s.

Here we shall not derive this from (6) but prove the theorem directly: we take a measurable function f from $\boldsymbol{M}$ such that $0 \leqslant f \leqslant \mathbf{1}$ (almost everywhere) and $\int f dz \leqslant s$. Then $\int f(z)(z|\omega|z)dz = \omega(\phi f)$. Because of (ii) and (iii) the operator ϕf satisfies $0 \leqslant \phi f \leqslant \mathbf{1}$ and $\operatorname{Tr} \phi f \leqslant s$. However, the right-hand side of (7) is equal to the supremum of the numbers $\omega(b)$ with $b \leqslant \boldsymbol{B}$ running through all the elements fulfilling $0 \leqslant b \leqslant \mathbf{1}$ and $\operatorname{Tr} b \leqslant s$.

At this point we see the importance of the number

$$e_s := \sup_f(\phi f), \quad f \in \{0 \leqslant f \leqslant \mathbf{1}, \int f dz \leqslant s\}. \tag{8}$$

Obviously, there is a set $N \subset \boldsymbol{R}^{2n}$ with dz-measure s such that integrating $(z|\omega|z)\,dz$ over this set, we get exactly e_s. Let us denote the characteristic function of this set N by g. Then $e_s = \int g(z)(z|\omega|z)\,dz$. We try to estimate this integral using Hölder's inequality. Clearly, the L^p-norm of g equals $s^{1/p}$.

Now we use an inequality of E. H. Lieb [3] : if ω is a pure normal state one has

$$\int (z|\omega|z)^q dz \leqslant (1/q)^n, \quad q \geqslant 1. \tag{9}$$

The convexity of $t \to t^q$ for $q \geqslant 1$ makes (9) valid for every density matrix. Applying now Hölder's inequality we get

$$e_s \leqslant s^{1/p}(1/q)^{1/qn} \quad \text{with} \quad 1/p + 1/q = 1.$$

Taking the infimum over $1 \leqslant p \leqslant \infty$ we arrive at

THEOREM 3. *Let ω be a normal state and let the Liouville measure of the subset $N \subset \boldsymbol{R}^{2n}$ be smaller than s. Then*

$$\int_N (z|\omega|z)dz \leqslant s\exp(-s^n/ne) \quad \textit{for } 0 \leqslant s^n \leqslant e; \tag{10}$$

here, $e = 2.7182\ldots$

For pure states (10) is much sharper than (7). It would be nice to get an estimate which sharpens Theorem 2 so that it includes Theorem 3!

Theorem 3 gives a lower bound to every F^{cl} in the following way. We construct a probability measure which is, relative to the Liouville measure, less mixed than all the distributions $(z|\omega|z)dz$. Then if one calculates (5) with such a measure, one gets this lower bound. Doing so we obtain

THEOREM 4. *Let F be concave and continuous on* $[0,1]$ *and* $F(0) = 0$. *Then for every normal state ω of* $\boldsymbol{B}$

$$F^{\mathrm{cl}}(\omega) \geqslant F_0 \tag{11}$$

with

$$F_0 = e^{1/n}\int_0^1 F\left[(1-t^n)\exp(-t^n/n)\right]dt. \tag{12}$$

Remarks. At first fix $n = 1$. (a) For $F = -s^q$, $q \geqslant 1$, our estimates are worse than Lieb's. For example, for $q = 2$ we get only -0.58 instead of Lieb's bound -0.5. This is unavoidable, because the order structure cannot be determined by these concave functions only. (b) We get, however, also estimates for other functions by easy numeric estimations. An example is $F_0 = 1.492$ if $F = s^{1/2}$. (c) Quite another example provides the function $F = 0$ for $s \leqslant r$ and $F(s) = r - s$ otherwise. With r in the unit interval choose v to satisfy $v\exp v = re$. Then $F_0 = er(2 - v - 1/v)$. This, again, is easy to handle numerically.

III. Now let us shortly comment on the total state space. If $\boldsymbol{A}$ is a W^*-algebra, ω one of its states we have a unique decomposition with some $0 \leqslant p \leqslant 1$

$$\omega = p\omega_s + (1-p)\omega_n,$$

where the indices s and n refer to "normal" and "singular". Then we define for concave F

$$F_{\boldsymbol{A}}(\omega) = pF_{\boldsymbol{A}}(\omega_s) + (1-p)F_{\boldsymbol{A}}(\omega_n).$$

In case $\boldsymbol{A} = \boldsymbol{B}$ we identify

$$F_{\boldsymbol{B}}(\omega_n) = F^q(\omega_n), \qquad F_{\boldsymbol{B}}(\omega_s) = \sup_{s\to 0} F(s)/s.$$

Similarly one must have

$$F_M(\omega_n) = F^{\mathrm{cl}}(\omega_n) \quad \text{and} \quad F_M(\omega_s) = \lim_{s\to 0} F(s)/s\,.$$

According to Theorem 1 we then always have $F_B(\omega) \leqslant F_M(\phi^*\omega)$.

This seems to be a tautology. However, one can give a more intrinsic definition for the F-functionals.

One example of such a definition emerges from [4] and from the lower semicontinuity of F_B, which is proved as in [1] for the entropy. This definition is

$$F_B(\omega) = \sup_Q \inf_j \sum F(t_j), \quad \sum t_j = 1,\ 0 \leqslant t_j\,.$$

Here the supremum is taken over all weak neighbourhoods Q of the state ω in the state space, and the infimum runs through all the indicated sums restricted as follows: there are pure normal states $\rho_1, \rho_2, \ldots$ such that the state $\sum t_j \rho_j$ is contained in Q

$$\sum_j t_j \rho_j \in Q\,.$$

Acknowledgement

I wish to thank A. Wehrl and E. Lieb for showing me their papers prior to publication. Valuable discussions with P. Alberti and G. Lassner are acknowledged.

REFERENCES

[1] Wehrl, A.: *Rev. Mod. Phys.* **50** (1978), 221.
[2] —: *On the Relation between Classical and Quantum-Mechanical Entropy* (to appear in Rep. Math. Phys.).
[3] Lieb, E. H.: *Commun. math. Phys.* **62** (1978), 35.
[4] Uhlmann, A.: *Rep. Math. Phys.* **1** (1970), 147.