KMU - QFT - 1979/8

On Entropy in Quantum Statistics

Armin Uhlmann
Karl-Marx-University
Dep. of Physics and NTZ
GDR-701 Leipzig

To appear in:
Proceedings of the "International Symposium on Fundamental Problems of Theoretical and Mathematical Physics",
Dubna, USSR, August 23 - 27, 1979

1. Generalities.

About 10 years ago a book [1] has been edited in honour of Nikolai N. Bogolubov.

In this book F. Baumann and R. Jost described their work on an entropy inequality, the so-called strong subadditivity property conjectured by O.E. Lanford, D. Robinson, and D. Ruelle [2], [3], and on the Yanase-Wigner-Dyson conjecture. E. Lieb [4] was the first, partially together with B. Ruskai [5], who was able to solve both problems which, indeed, are connected one-to-another. Thus, I could add only two further variants of proving this. One [6] makes use of the concavity of $A \rightarrow \mathrm{Tr.}\,\exp(B + \ln A)$ discovered by E. Lieb [4] and H. Epstein [7], the other variant relies on an interpolation theory [8] based on an observation of W. Pusz and S. L. Woronowicz [9]. In trying to analyse the strong subadditivity conjecture in the quantum case (- it is almost trivial for classical discrete distributions -), as a by-product, the order structure of states [10] appeared, a tool to describe some aspects of irreversibility (though not the very problem of irreversibility).

In asking about differences in the general properties of entropy in classical and in quantum physics one finds astonishingly only very few. Indeed, at the first glance the main difference is in the much harder proofs of the quantum entropy properties. But there is (at least) one striking difference between classical and quantum entropy: the failure of monotonicity for the latter (see [11] , [12]).

Though mathematically rather simple, if not trivial, it is physically remarkable: The entropy of a subsystem can be strictly larger than the entropy of the total system !

One does not mention, usually, this fact for several reasons. At first, in characterizing subsystems by spatial separation and then performing the thermodynamical limit, the entropy (per volume) becomes additive and monotonicity holds. Secondly, monotonicity can be destroyed only if the observables characterizing the total system do not commute with those of the subsystem.

This may well be compared with a measurement in the sense of von Neumann. To recognize something as a subsystem of a total system "reduction" of the density matrix. This reduction is implies a accompanied by an extra increase in entropy. This extra increase can overcompensate the usual decrease we have in situations equivalent with classical ones.

In view of the non-monotonicity of quantum entropy it is important and really not self-evident that most of the properties of entropy known in classical statistics remain valid in the quantum case also.

In the following we shortly discuss some of those properties of entropy which have been under investigation in the last decade. We restrict ourselves to the technical simplest possible assumptions, i.e. we handle only finite-dimensional density matrices, Hilbert spaces,... Some more complete discussions are given in W. Thirring [13] and A. Wehrl [12] .

2. Systems and Subsystems, states and reduced states.

Let us assume (for simplicity) that the observables of a certain physical system can bee choosen as the set of Hermitian n-by-n-matrices. We then call the matrix algebra $\underline{\underline{M}}_n$ of all n-by-n-matrices the "algebra of observables".

A "state", ω, is uniquely determined by the family of expectation values, $\omega(A)$, where $A \in \underline{\underline{M}}_n$ is any matrix. One requires linearity in $A \rightarrow \omega(A)$, $\omega(A) \geqslant 0$ if $A \geqslant 0$ (positivity), and the normalization $\omega(E) = 1$ with the unity matrix E . These requirements are satisfied if and only if there is with a density matrix ω

$$\omega(A) \quad = \quad \text{Tr.}\, \omega A \qquad \text{for all} \quad A \tag{1}$$

(We use the same symbol for the state functional $A \rightarrow \omega(A)$ and its density matrix. Let us recall that a density matrix is a positive semidefinite matrix of trace one.)

$\underline{\underline{M}}_n$ can be considered as the algebra $\underline{\underline{B}}(\underline{\underline{H}})$ of linear operators acting on a n-dimensional Hilbert space $\underline{\underline{H}}$. If then $x \in \underline{\underline{H}}$ is a normed vector, it defines a pure state $A \rightarrow \pi(A) = (x,Ax)$. If π is the projection operator from $\underline{\underline{H}}$ onto the subspace spanned by x we have $(x,Ax) = \text{Tr.}\, \pi A$ and π is a density matrix. As it is known a density matrix π represents a pure state iff it is a one-dimensional projection.

One of the most satisfying ways to define "subsystems" is by observables. We have to select all those observables of the total system which do not discriminate outside the subsystem. Clearly, a state of the subsystem should be characterized by

its expectation values with respect to these observables.
A correct and sufficiently general definition is to associate the concept of a subsystem with that of an algebra of observables which is a <u>subalgebra</u> $\underline{\underline{A}}$ of the algebra $\underline{\underline{M}}_n$. For the possibility of a correct interpretation the subalgebra $\underline{\underline{A}}$ should fulfil two conditions: $A \in \underline{\underline{A}}$ should imply $A^* \in \underline{\underline{A}}$, and $E \in \underline{\underline{A}}$ where E is the unit element of $\underline{\underline{M}}_n$.
Every linear, positive, and normed functional on $\underline{\underline{A}}$ is called "<u>state of</u> $\underline{\underline{A}}$", i.e. state of the subsuystem under consideration.
If we have a state of $\underline{\underline{M}}_n$ i.e. of our total system, then its restriction onto $\underline{\underline{A}}$ is called the state <u>reduced to the subsystem given by</u> $\underline{\underline{A}}$.
The subalgebras in question are well known in its structure.
We shall consider here only those which are themselves isomorphic to full matrix algebras. This is the case if and only if the centre of $\underline{\underline{A}}$ consists of the multiples of the unit element only ("trivial centre"). Technically, $\underline{\underline{A}}$ is then called a <u>factor</u>.
Let $\underline{\underline{A}}$ be a factor of $\underline{\underline{M}}_n \simeq \underline{\underline{B}}(\underline{\underline{H}})$. Then there is a decomposition $\underline{\underline{H}} = \underline{\underline{H}}_1 \otimes \underline{\underline{H}}_2$ such that $\underline{\underline{A}}$ consists of all matrices (operators) of the form $A_1 \otimes E_2$ where A_1 acts on $\underline{\underline{H}}_1$, E_2 is the identity on $\underline{\underline{H}}_2$ and $\otimes$ is the (Kronecker) tensor product. If we change the role of $\underline{\underline{H}}_1$ and $\underline{\underline{H}}_2$ we get again a factor. It is denoted by $\underline{\underline{A}}'$ and consists of all those matrices of $\underline{\underline{M}}_n$ which commute with every matrix of $\underline{\underline{A}}$. $\underline{\underline{A}}'$ describes the subsystem <u>orthogonal</u> to $\underline{\underline{A}}$.
Let now ω be a state of $\underline{\underline{M}}_n$ and $\omega_{\underline{\underline{A}}}$ its reduction onto $\underline{\underline{A}}$.
Then

$$A_1 \longrightarrow \omega_{\underline{\underline{A}}} (A_1 \otimes E_2) = \omega (A_1 \otimes E_2) \,, \quad A_1 \in \underline{\underline{B}}(\underline{\underline{H}}_1) \qquad (2)$$

is a state of $\underline{\underline{B}}(\underline{\underline{H}}_1)$ and hence given by a density matrix ω_1.
ω_1 is usually called <u>reduced density matrix</u>.

Sometimes it is useful to lift the reduced density matrix to a density matrix of the total system such that it carries as few as possible information about the subsystem given by $\underline{\underline{A}}'$. This is done by the map

$$\omega \rightarrow T_{\underline{\underline{A}}}\,\omega = (1/n_2)\, \omega_1 \otimes E_2 \tag{3}$$

where n_2 is the dimension of $\underline{\underline{H}}_2$. The right hand side of (3) can be obtained by invariant integration of $U^{-1} \omega U$ over the group of unitary matrices contained in $\underline{\underline{A}}'$. Let us now assume $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ denote complete orthonormal systems of $\underline{\underline{H}}_1$ and $\underline{\underline{H}}_2$ respectively. The reduced density matrices ω_1, ω_2 of both subsystems considered with respect to the pure state π characterized by the vector

$$\sum c_{ik}\, x_i \otimes y_k \tag{4a}$$

have matrix elements with respect to the orthonormal systems

$$i,j \rightarrow \sum_k c_{ik}\, \bar{c}_{jk} \quad \text{and} \quad k,l \rightarrow \sum_i c_{ik}\, \bar{c}_{il} \tag{4b}$$

Here we see explicitely the reason for the non-monotonicity of entropy: If $\dim \underline{\underline{H}}_1 \gg \dim \underline{\underline{H}}_2$, <u>every</u> state of $\underline{\underline{A}}$ can be obtained by reducing a suitable <u>pure</u> state of the total system to the given subsystem! In classical statistics, in contrat to this, a pure state always reduces to pure states in subsystems.

There is, of course, no "total system" in Nature. Every system can be considered as a subsystem of another larger or more detailed described system. (<u>Feynman</u> in is "lectures" remarked: "But it is true that if we look at a glass of wine closely enough we see the entire universe.") Hence the concept of pureness of a state entirely relates to a given observable algebra. By enlarging a system, by a more detailled analysis,

a mixed state (mixture a la von Neumann and Gibbs) can become a pure one and vice vera.
Unfortunatly, the connection of this pecularity (and of similar ones) of quantum statistics to the phenomenon of irreversibility is not yet fully analysed.

3. Entropy.

Entropy is a state functional of the following type: Let $s \to f(s)$ denote a real-valued function on the unit intervall and define

$$S_f(\omega) = k_B \,\mathrm{Tr}.f(\omega) \tag{5}$$

(k_B is Boltzmann's constant.) In the case $f(s) = -s \ln s$ we write simply $S(\omega)$ and call it "entropy of ω".
There is a more intrinsic definition for S and, more general, for S_f if $f'' \leq 0$ (concavity), which applies for alle (here finite dimensional) algebras $\underline{\underline{A}}$: The set of all states of $\underline{\underline{A}}$ is naturally a convex set with respect to the performing of mixtures. Its extremal elements are called "pure states". (For $\underline{\underline{M}}_n$ this is a statement, for general $\underline{\underline{A}}$ it is a definition.)
Now $S_f(\omega)$ is the infimum of all numbers $k_B \sum_j f(c_j)$ where the sequences $c_1, c_2, \ldots$ runs over all possibilities to get $\omega = \sum c_i \pi_i$ with $c_i \geq 0$ and suitably choosen pure states π_i.
The entropy of a state is zero iff it is a pure state. Because we can get pure states after enlarging a given system, entropy does not necessarily shows monotonicity in going from a system to a subsystem or vice vera.
A very usual example is the observation of one pecular atom in a molecule, the molecule being in its ground state. The reduced density matrix of the atom observed is higly mixed. Exotic ex-

amples are (i) the remark of Lieb [11] that the entropy of every (meta-)galaxis may steadily increase though the universum's total entropy remain zero and (ii) the absorption (going beyond the horizon of forgetting of a black hole) of one component of a particle-antiparticle system originally in a pure state. If the body is in a Gibbsian canonical state then the escaping particle is in a Gibbsian canonical ensemble of the same temperature as the body is (Hawking).

These affairs are controlled by some inequalities which, apart from some finer technical points, characterize S uniquely within the state functionals (not only within all S_f). If we consider $\underline{\underline{A}}$ and $\underline{\underline{A}}'$ of $\underline{\underline{M}}_n$ belongig to a decomposition $\underline{\underline{H}} = \underline{\underline{H}}_1 \otimes \underline{\underline{H}}_2$, a state ω and its reductions ω_1 and ω_2 to $\underline{\underline{A}}$ and $\underline{\underline{A}}'$ accordingly, we have

$$S(\omega) \leq S(\omega_1) + S(\omega_2) \quad \underline{\text{subadditivity}} \tag{6}$$

$$\underline{\text{triangel inequality:}} \quad |S(\omega_1) - S(\omega_2)| \leq S(\omega) \tag{7}$$

Subadditivity is long and well known, but (7) is discovered by Araki and Lieb [14].

The already mentioned strong subadditivity is a considerable sharpening of (6). Here we consider the case $\underline{\underline{H}} = \underline{\underline{H}}_1 \otimes \underline{\underline{H}}_2 \otimes \underline{\underline{H}}_3$ and the corresponding to $\underline{\underline{H}}_j$ factors $\underline{\underline{A}}_j$. For any permutation (i,j,k) of the integers $(1,2,3)$ the factor $\underline{\underline{A}}_{jk}$ corresponding to $\underline{\underline{H}}_{jk}$ satisfies $\underline{\underline{A}}_{jk} = \underline{\underline{A}}_i' = \underline{\underline{A}}_j \otimes \underline{\underline{A}}_k$. The reduction of a state ω onto a subsystem $\underline{\underline{A}}_i$ or $\underline{\underline{A}}_{jk}$ is denoted by ω_i and ω_{jk} strong subadditivity may then be expressed by

$$S(\omega) + S(\omega_i) \leq S(\omega_{ik}) + S(\omega_{ij}) \tag{8}$$

One proof of (8) can be sketched as follows: We at first notice that (3) implies $S(T_{\underline{\underline{A}}}\omega) = S(\omega_1) + \ln n_2$.

Denoting by T_i , T_{jk} , the maps corresponding to the factors $\underline{\underline{A}}_i$, $\underline{\underline{A}}_{jk}$, one easily finds the equivalence of (8) with

$$S(\omega) + S(T_i\omega) \leq S(T_{ik}\omega) + S(T_{ij}\omega) \tag{8a}$$

because the logarithmic terms cancel.
This inequality can be handled directly by the Epstein-Lieb result mentioned in §1. Another way is to introduce Umegaki's relative entropy $S(\varrho,\omega) = \mathrm{Tr.}\,(\omega \ln \omega - \omega \ln \varrho)$ and to mention the nice relation

$$S(T_{ik}\omega, \omega) = -S(\omega) + S(T_{ik}\omega) \tag{9}$$

Now one finds $T_{ij}\,T_{ik} = T_i$ and we are done if we could prove for $T = T_{ij}$ and $\varrho = T_{ik}\,\omega$ the general relation

$$S(\varrho, \omega) \geq S(T\varrho, T\omega) \tag{10}$$

for a class of maps T which contain T_{ik} . (10) is true for all maps of the form ($t_j \geq 0$, $\sum t_i = 1$)

$$\sigma \rightarrow T\sigma = \sum t_i\, U_i\, \sigma\, U_i^{-1} \quad , \quad U_j \text{ unitary} \tag{11}$$

to which set the map T_{ik} belongs. The proof is by noting

$$S(\varrho,\omega) = -(d/ds)_{s=o}\ \mathrm{Tr.}\ \omega^{1-s}\,\varrho^{s} \tag{12}$$

and proving the assertion for $\mathrm{Tr.}\ \omega^{1-s}\,\varrho^{s}$.
The latter statement, however, is a consequence of Lieb's concavity theorem [4] for ($0 < s < 1$)

$$\varrho, \omega \rightarrow \mathrm{Tr.}\ X^{*}\,\omega^{1-s}\,X\,\varrho^{s} \tag{13}$$

Lieb has shown his theorem by analytical interpolation. Another interpolation theory [8] shows the validity of (10) and the corresponding inequality for the functional (13) for every map T which is linear, and chaos enhancing (see below), and for which T^{*} fulfils a Kadison inequality.

4. The order structure of states.

Some properties of entropy are shared by many other state functionals: Let $f'' \leq 0$, $t_i \geq 0$, $\sum t_j = 1$.
Then for arbitrary choosen states we have

$$S_f(\sum t_i \rho_i) \geq \sum t_j S_f(\rho_j) \tag{14}$$

which is complemented by

$$\sum S_f(t_i \rho_i) \geq S_f(\sum t_j \rho_j) \tag{15}$$

According to (14) every S_f is increasing in the "direction" of performing Gibbs-von Neumann mixtures - but every S_f does this in its own way. To get rid of this arbitrariness one considers a partial order within the set of all states, the "order structure" of states [10], [12], [13] in defining

$$\omega \succ \rho \quad \text{iff} \quad S_f(\omega) \geq S_f(\rho) \quad \text{for } \underline{\text{all}} \ f'' \leq 0 \tag{16}$$

Verbally, $\omega \succ \rho$ is expressed by saying ω is more chaotic (or "more mixed", or "less pure") than ρ.
A map T is called chaos enhancing if always $T\rho \succ \rho$
Every map (3) is of that kind. Other examples are described in the quoted literature.
Generally, $\succ$ characterizes relaxations, diffusions and other processes which point, in a very strict sense, "in the direction of irreversibility".
The mathematical technique will be described in [15].

References:

1. Problems of Theoretical Physics. NAUKA, Moscow 1969
2. Lanford, O.E., and D. Robinson, J.Math.Phys. 9, (1968) 1120
3. Robinson, D.W., and D. Ruelle, Comm.Math.Phys. 5 (1967) 288
4. Lieb, E., Adv.Math. 11 (1973) 267
5. Lieb, E., and M.B. Ruskai, J.Math.Phys. 14 (1973) 1938
6. Uhlmann, A., Wiss.Z. Karl-Marx-Univ. Leipzig, 22 (1973) 139
7. Epstein, H., Comm.Math.Phys. 31 (1973) 317
8. Uhlmann, A., Comm.Math.Phys. 54 (1977) 21
9. Pusz, W., and S.L. Woronowicz, Rep.Math.Phys. 8 (1975) 159
10. Uhlmann, A., Wiss.Z. Karl-Marx-Univ. Leipzig, 20 (1971) 633 and 21 (1972) 427;
    Alberti, P.M., thesis, Leipzig 1973
    Wehrl, A., Rep.Math.Phys. 6 (1974) 15
11. Lieb, E., Bull.Am.Math.Soc. 81 (1975) 1
12. Wehrl, A., Rev.Mod.Phys. 50 (1978) 221
13. Thirring, W., Lehrbuch der Mathematischen Physik IV, Springer-Verlag Wien New York, to appear.
14. Araki, H., and E. Lieb, Comm.Math.Phys. 18 (1970) 160
15. Alberti, P.M., and Uhlmann, A., "Stochsticity and Partial Order". (in preparation).