Elliott H. Lieb
Mary Beth Ruskai

Proof of the Strong Subadditivity of Quantum-Mechanical Entropy

Les rencontres physiciens-mathématiciens de Strasbourg - RCP25, 1973, tome 19
« Conférences de : P. Lelong, A. Lichnerowicz, E. Lieb et C. Stanojevic », , exp. n° 5,
p. 36-55
<http://www.numdam.org/item?id=RCP25_1973__19__A5_0>

PROOF OF THE STRONG SUBADDITIVITY OF QUANTUM-MECHANICAL ENTROPY.

Elliott H. LIEB[*+]

I . H . E . S .

91- Bures-sur-Yvette

France

:-:-:-:

Mary Beth RUSKAI[*§]

Department of Mathematics

M . I . T .

Cambridge, Mass.02139

U . S . A .

:-:-:-:

ABSTRACT - We prove several theorems about quantum-mechanical entropy ; in particular, that it is strongly subadditive.

---

*Work partially supported by U.S. National Science Foundation Grant GP-31674 X.

+On leave from Department of Mathematics, M.I.T.,Cambridge, Mass.02139, U.S.A. . Work partially supported by a Guggenheim Memorial Foundation fellowship.

§Presently, N.R.C. fellow at Department of Physics, University of Alberta, Edmonton 7, Canada.

I.- INTRODUCTION.

In this paper we prove several theorems about quantum mechanical entropy, in particular, that it is strongly subadditive (SSA). These theorems were announced in an earlier note[1], to which we refer the reader for a discussion of the physical significance of SSA and for a review of the historical background. We repeat here a bibliography of relevant papers[2-9].

The setting for these theorems is this :

a) Given a separable Hilbert space H and a positive, trace-class operator, ρ , on H (i.e. $\rho \geq 0$ means $(\Psi, \rho\Psi) \geq 0$ for all Ψ in H) , the entropy of ρ is defined to be

$$S(\rho) \equiv -\mathrm{Tr}\ \rho \ln \rho = -\sum_{i=1}^{\infty} \lambda_i \ln \lambda_i \quad , \tag{1.1}$$

where Tr means trace, the λ_i are the eigenvalues of ρ , $0 \ln 0 \equiv 0$, and we permit the possibility $S(\rho) = \infty$. In physical applications one also requires that $\mathrm{Tr}\ \rho = 1$, in which case ρ is called a density matrix.

b) If $H_{12} = H_1 \otimes H_2$ is the tensor product of two Hilbert spaces and ρ_{12} is a positive, trace-class operator on H_{12} , we can define a positive, trace-class operator, ρ_1 , on H_1 by the partial trace, i.e.

$$\rho_1 \equiv \mathrm{Tr}_2\ \rho_{12} \tag{1.2}$$

by which we mean

$$(\varphi, \rho_1 \Psi) = \sum_{i=1}^{\infty} (\varphi \otimes e_i \, , \, \rho_{12} [\Psi \otimes e_i]) \qquad (1.3)$$

for all φ , Ψ in H_1 and $\{e_i\}_{i=1}^{\infty}$ any orthonormal basis in H_2 . We shall denote $S(\rho_1)$ by S_1 etc... In like manner one can have $H_{123} = H_1 \otimes H_2 \otimes H_3$, and ρ_{123} a positive, trace-class operator on H_{123} , and define ρ_{12} on $H_{12} \equiv H_1 \otimes H_2$, ρ_1 on H_1 , etc... by partial traces. When no confusion arises, we shall frequently use the symbol ρ_1 to denote the operator $\rho_1 \otimes \mathbb{1}_2$ on H_{12}.

Our main results are the following two theorems.

<u>Theorem 1</u> : Let $H_{12} = H_1 \otimes H_2$. Then the function

$$\rho_{12} \longmapsto S_1 - S_{12} \qquad (1.4)$$

is convex on the set of positive, trace-class operators on H_{12} .

<u>Theorem 2 - (Strong Subadditivity)</u> : Let H_{123} and ρ_{123} be defined as in (b) above. Then

(i) $S_{123} + S_2 - S_{12} - S_{23} \leq 0$ (1.5)

and

(ii) $S_1 + S_3 - S_{12} - S_{23} \leq 0$ (1.6)

In the next section we prove these theorems in the finite-dimensional case. In section III we elucidate the connection between these two theorems and give some related results. Section IV contains the proofs for the infinite-dimensional case and is based on the appendix kindly contributed by B. Simon, to whom we are most grateful.

II.- PROOFS OF THEOREMS 1 AND 2 IN THE FINITE-DIMENSIONAL CASE.

Proof of Theorem 1 : The theorem states that

$$(S_1 - S_{12})(\rho_{12}) \leq \alpha\,(S_1 - S_{12})(\rho'_{12}) + (1-\alpha)\,(S_1 - S_{12})(\rho''_{12}) \qquad (2.1)$$

where $\rho_{12} = \alpha\,\rho'_{12} + (1-\alpha)\,\rho''_{12}$, $0 \leq \alpha \leq 1$, and ρ'_{12} and ρ''_{12} are any positive, trace-class operators on H_{12}. We shall assume that both ρ'_{12} and ρ''_{12} are strictly positive and appeal to continuity of $\rho \longmapsto S(\rho)$ in the semi-definite case. Letting

$$\Delta = \alpha\,\mathrm{Tr}_{12}\;\rho'_{12}(-\ln\rho'_{12} + \ln\rho'_1 + \ln\rho_{12} - \ln\rho_1)\quad ,$$

and

$$\Gamma = (1-\alpha)\,\mathrm{Tr}_{12}\;\rho''_{12}\,(-\ln\rho''_{12} + \ln\rho''_1 + \ln\rho_{12} - \ln\rho_1)\quad ,$$

one sees that (2.1) is equivalent to $\Delta + \Gamma \leq 0$. We now use Klein's inequality [7,10] :

$$\mathrm{Tr}\,(-A\ln A + A\ln B) \leq \mathrm{Tr}\,(B - A)\quad . \qquad (2.2)$$

(Alternatively, one could use the Peierls - Bogoliubov inequality in a similar way[2]) . We first apply (2.2) to Δ with $A = \rho'_{12}$ and $B = \exp\,[\ln\rho'_1 + \ln\rho_{12} - \ln\rho_1]$ and then similarly to Γ . Then

$$\begin{aligned}\Delta + \Gamma &\leq \alpha\,\mathrm{Tr}_{12}[\exp(\ln\rho'_1 + \ln\rho_{12} - \ln\rho_1) - \rho'_{12}]\\ &\quad + (1-\alpha)\,\mathrm{Tr}_{12}\,[\exp(\ln\rho''_1 + \ln\rho_{12} - \ln\rho_1) - \rho''_{12}] \qquad (2.3)\\ &\leq \mathrm{Tr}_{12}\,[\exp(\ln\rho_1 + \ln\rho_{12} - \ln\rho_1) - \rho_{12}] = 0\quad .\end{aligned}$$

The second inequality in (2.3) follows from the concavity [11] of $C \longmapsto \mathrm{Tr}[\exp(K + \ln C)]$ for positive C applied to $\rho_1^{*} = \alpha\,\rho_1' + (1-\alpha)\,\rho_1''$ with $K = \ln \rho_{12} - \ln \rho_1$. Q.E.D.

<u>Proof of Theorem 2</u> : It has already been pointed out[2] that (1.5) and (1.6) are equivalent ; however, we shall prove each statement separately.

<u>(i) Proof of (1.5)</u> : We use Klein's inequality, (2.2), with $A = \rho_{123}$ and $B = \exp[-\ln \rho_2 + \ln \rho_{12} + \ln \rho_{23}]$. One finds

$$F(\rho_{123}) \equiv S_{123} + S_2 - S_{12} - S_{23} \leq \mathrm{Tr}_{123}\,[\exp(\ln \rho_{12} - \ln \rho_2 + \ln \rho_{23}) - \rho_{123}].$$

We now apply a generalization[11] of the Golden-Thompson inequality, i.e.

$$\mathrm{Tr}[\exp(\ln B - \ln C + \ln D)] \leq \mathrm{Tr} \int_0^\infty B\,(C + x\,\mathbb{1})^{-1}\,D(C + x\mathbb{1})^{-1}\,dx. \qquad (2.4)$$

Thus

$$F(\rho_{123}) \leq \mathrm{Tr}_{123}\Big[\int_0^\infty \rho_{12}(\rho_2 + x\mathbb{1})^{-1}\,\rho_{23}(\rho_2 + x\mathbb{1})^{-1}\,dx - \rho_{123}\Big]$$

$$= \mathrm{Tr}_2 \int_0^\infty \rho_2\,(\rho_2 + x\mathbb{1})^{-1}\,\rho_2(\rho_2 + x\mathbb{1})^{-1}\,dx - \mathrm{Tr}_{123}\,\rho_{123}$$

$$= \mathrm{Tr}_2\,\rho_2 - \mathrm{Tr}_{123}\,\rho_{123} = 0 \quad . \qquad \text{Q.E.D.}$$

(ii) Proof of (1.6) : Call the left side of (1.6) $G(\rho_{123})$. Note that $S_1 - S_{12}$ is convex in ρ_{12} by Theorem 1 ; since ρ_{12} is linear in ρ_{123} , $S_1 - S_{12}$ is convex in ρ_{123} . Thus, $G(\rho_{123})$ is convex in ρ_{123} . In the convex cone of positive matrices, the extremal rays consist of matrices of the form $\rho = \alpha P$ where $\alpha \geq 0$ and P is a one-dimensional projection. If ρ_{123} is extremal, then (see Ref.2, lemma 3) $S_1 = S_{23}$ and $S_3 = S_{12}$, so that $G(\rho_{123}) = 0$. Every positive matrix ρ_{123} can be written as a convex combination of extremal matrices ; it then follows from the convexity of G that $G(\rho_{123}) \leq 0$. Q.E.D.

III.- REMARKS AND RELATED RESULTS.

We have already noted in the proof of (1.6) that Theorem 1 implies Theorem 2. We now note that the converse is also true and give several alternate proofs of Theorems 1 and 2 . We then show that $F(\rho_{123})$ is not convex and give a corollary to Theorem 1.

A) To show Theorem 2 implies Theorem 1 it suffices to note that (apart from the trivial interchange of the subscripts 1 and 2 in (2.1)) (1.5) is identical to (2.1) for a special choice of ρ_{123}, i.e. $\rho_{123} = \alpha\rho'_{12} \otimes E_3 + (1 - \alpha)\, \rho''_{12} \otimes F_3$ where H_3 is chosen to be two-dimensional and E_3 and F_3 are orthogonal, one-dimensional projections on H_3 .

B) Uhlmann[9] has shown that (1.5) follows from the concavity of $C \longmapsto \mathrm{Tr}\, \exp(K + \ln C)$. This has been shown to be true by Lieb[11] , and an alternate proof was later found by Epstein[12] . Therefore, Uhlmann's remark gives an alternate proof of (1.5) .

C) The proof of (1.6) shows that Theorem 1 implies Theorem 2 . However, (1.6) is not equivalent to (1.5) in other contexts[13]. (In fact, (1.6) is false in the classical continuous case[6]) . Therefore, it is instructive to note that one can show that Theorem 1 implies (1.5) directly without using (1.6). Baumann and Jost[3,5] have shown that a special choice of ρ'_{12} and ρ''_{12} in (2.1) implies that $\mathrm{Tr} \int_0^\infty A^*(C+x\mathbb{1})^{-1} A(C+x\mathbb{1})^{-1}\, dx$ is jointly convex in (A,C) where A and C are matrices with $C > 0$. Lieb has then shown[11] that this implies $C \longmapsto \mathrm{Tr}\, \exp(K + \ln C)$ is concave in C . The last statement was used to prove[11] (2.4) which, as we have already seen, implies (1.5). Alternatively, we have already noted in (B) above that concavity of $C \longmapsto \mathrm{Tr}\, \exp[K + \ln C]$ implies (1.5).

D) We have already shown that the left side of (1.6), $G(\rho_{123})$, is convex. One might wonder, therefore, if the left side of (1.5), $F(\rho_{123})$, is also convex. In fact, it is not. If it were, one could choose H_2 to be one-dimensional so that

$$F(\rho_{123}) = S_{13} - S_1 - S_3 \equiv E(\rho_{13}) ,$$

would have to be a convex function of ρ_{13}. Take H_1 and H_3 to be two-dimensional and choose ρ'_{13} and ρ''_{13} to be the following orthogonal, one-dimensional projections :

$$\rho'_{13}(i_1, i_3 ; j_1, j_3) = \tfrac{1}{2} \delta(i_1, i_3) \delta(j_1, j_3)$$

and

$$\rho''_{13}(i_1, i_3 ; j_1, j_3) = \tfrac{1}{2} [1 - \delta(i_1,i_3)][1 - \delta(j_1,j_3)] ,$$

where δ is the Kronecker delta. Then $\rho'_1 = \rho''_1 = \frac{1}{2} \mathbb{1}_1$, $\rho'_3 = \rho''_3 = \frac{1}{2} \mathbb{1}_3$, and $E(\rho'_{13}) + E(\rho''_{13}) - 2 E(\frac{1}{2} \rho'_{13} + \frac{1}{2} \rho''_{13}) = - 2 \ln 2 < 0$, which is a contradiction.

E) It was pointed out in Ref. 11 that if $f(A)$ is a convex function from the set of positive matrices into $\mathbb{R}$, and if it is also homogenous (i.e. $f(\lambda A) = \lambda f(A)$ for all $\lambda > 0$), then

$$\frac{d}{dx} f(A + x B)\Big|_{x=0} \equiv \lim_{x \downarrow 0} x^{-1} [f(A + x B) - f(A)] \leq f(B) , \qquad (3.1)$$

whenever A,B are positive matrices and the above limit exists. The function $(S_1 - S_{12})(\rho_{12})$ has these properties. To apply (3.1) we compute :

$$\frac{d}{dx} S(\rho + x\gamma) = -\frac{d}{dx} \operatorname{Tr}[(\rho + x\gamma)\ln(\rho + x\gamma)]$$

$$= -\operatorname{Tr} \gamma \ln(\rho + x\gamma) - \operatorname{Tr} \gamma .$$

Using this in (3.1) we conclude :

<u>Corollary</u> : Let γ_{12} and ρ_{12} be positive, trace-class matrices on H_{12} . Then

$$\operatorname{Tr}_{12} \gamma_{12} \ln \rho_{12} - \operatorname{Tr}_1 \gamma_1 \ln \rho_1 \leq \operatorname{Tr}_{12} \gamma_{12} \ln \gamma_{12} - \operatorname{Tr}_1 \gamma_1 \ln \gamma_1 , \quad (3.2)$$

i.e. for each fixed γ_{12} , the left side of (3.2) achieves its maximum when $\rho_{12} = \gamma_{12}$.

IV.- EXTENSION TO INFINITE-DIMENSIONS.

We can use Theorem A.3 to extend Theorems 1 and 2 to infinite - dimensions. For simplicity, we confine our discussion to Theorem 1 where $H_{12} = H_1 \otimes H_2$. The extension of Theorem 2 is similar and we point out the necessary changes at the end of this section.

Let E_i^n ($i = 1,2$ and $n = 1,2,\ldots$) be sequences of increasing, finite-dimensional projections on H_i , converging strongly to the identity, and define

$$E^n = E_1^n \otimes E_2^n ,$$

$$\rho_{12}^n = E^n \rho_{12} E^n , \text{ and}$$

$$\rho_1^n = \mathrm{Tr}_2\, \rho_{12}^n = E_1^n(\mathrm{Tr}_2\, E_2^n \rho_{12} E_2^n) E_1^n . \tag{4.1}$$

Since the spaces $E_i^n H_i$ are finite dimensional, Theorem 1 is satisfied by ρ_{12}^n on $E_1^n H_1 \otimes E_2^n H_2$ for each n . Thus, it suffices to show that the sequences of matrices $\{\rho_{12}^n\}_{n=1}^\infty$ and $\{\rho_1^n\}_{n=1}^\infty$ satisfy the hypotheses of Theorem A.3 so that, e.g. $\lim_{n\to\infty} S(\rho_{12}^n) = S(\rho_{12}) = S_{12}$.

To show that $\{\rho_{12}^n\}_{n=1}^\infty$ satisfies Theorem A.3, we first note that $E^n \xrightarrow{s} \mathbb{1}_{12}$. If[14] the sequences $A_n \xrightarrow{s} A$ and $B_n \xrightarrow{s} B$, then $A_n B_n \xrightarrow{s} AB$. Consequently, ρ_{12}^n converges to ρ_{12} strongly, and therefore weakly. It follows from the Ritz principle (see Proposition A.1) that $\rho_{12}^n = E^n \rho_{12} E^n \prec E^{n+1} \rho_{12} E^{n+1} \prec \rho_{12}$, with $\prec$ as defined in the Appendix. Therefore, the hypotheses of Theorem A.3 are satisfied and

$$\lim_{n\to\infty} S(\rho_{12}{}^{n}) = S_{12} \tag{4.2}$$

To show that $\{\rho_1{}^n\}_{n=1}^{\infty}$ also satisfies Theorem A.3, define $\tilde{\rho}_1{}^n = \mathrm{Tr}_2\, E_2{}^n\, \rho_{12}\, E_2{}^n$. Then $\rho_1{}^n = E_1{}^n\, \tilde{\rho}_1{}^n\, E_1{}^n$. To show that $\rho_1{}^n$ converges to ρ_1 weakly, it suffices to show that $\tilde{\rho}_1{}^n$ converges to $\rho_1{}^n$ strongly. (In fact, it converges uniformly). To do this we can assume, without loss of generality, that $E_2{}^n$ projects on the space spanned by $e_i \ldots. e_n$ where $\{e_i : i = 1 \ldots \infty\}$ is an orthonormal basis in H_2 . Then

$$(\Psi, \tilde{\rho}_1{}^n\, \Psi) = \sum_{i=1}^{n} (\Psi \otimes e_i\ ,\ \rho_{12}\, \Psi \otimes e_i)$$

for all Ψ in H_1 , and it follows that

$$\tilde{\rho}_1{}^n \leq \tilde{\rho}_1{}^{n+1} \quad , \quad \text{and} \tag{4.3}$$

$$\lim_{n\to\infty} (\Psi\ ,(\rho_1 - \rho_1{}^n)\ \Psi) = \lim_{n\to\infty} \sum_{n+1}^{\infty} (\Psi \otimes e_i,\ \rho_{12}\ \Psi \otimes e_i) = 0\ . \tag{4.4}$$

Since $\tilde{\rho}_1{}^n$ is a monotone sequence of positive operators, (4.4) implies that $\tilde{\rho}_1{}^n \xrightarrow{\ s\ } \rho_1$ and therefore $\rho_1{}^n \xrightarrow{\ s\ } \rho_1$. Further, it follows from (4.3), i.e. the monotonicity of $\tilde{\rho}_1{}^n$, that

$$\rho_1{}^n \ntriangleleft E_1{}^{n+1}\ \tilde{\rho}_1{}^n\ E_1{}^{n+1}$$

$$\leq E_1{}^{n+1}\ \tilde{\rho}_1{}^{n+1}\ E_1{}^{n+1} = \rho_1{}^{n+1} \ntriangleleft \rho_1 \quad .$$

Thus, Theorem A.3 implies $\lim_{n\to\infty} S(\rho_1{}^n) = S(\rho_1) = S_1$.

The analysis for Theorem 2 is similar. One defines

$$E^n = E_1{}^n \otimes E_2{}^n \otimes E_3{}^n \quad ,$$

$$\rho^n_{123} = E^n \rho_{123} E^n \quad , \text{ and}$$

$$\rho^n_{12} = Tr_3 \rho^n_{123} \quad , \text{ etc...}$$

ACKNOWLEDGEMENTS.

We thank D. Ruelle for generating and encouraging our interest in this problem. We are also grateful to R. Jost, O. Lanford, and D. Robinson for their encouragement and for helpful conversations. One author (E.L) would like to thank the Chemistry Laboratory III, University of Copenhagen, where part of this work was done.

APPENDIX : CONVERGENCE THEOREMS FOR ENTROPY.

by B. Simon*, Princeton University.

We discuss a variety of convergence theorems which are useful in extending entropy inequalities from finite dimensional matrices to infinite dimensional operators on a Hilbert space.

Definition : Let A be a positive compact operator. $\mu_k(A)$ denotes the k th largest eigenvalue of A counting multiplicity.

Definition : Let $s(x)$ be the function on $[0,\infty)$ given by

$$s(x) = \begin{cases} -x \ln x & \text{if } x \geq 0 \\ 0 & \text{if } x = 0 \end{cases} .$$

If A is positive and compact, we set

$$S(A) = \sum_{k=1}^{\infty} s(\mu_k(A)) ,$$

the value infinity being allowed.

Definition : Let A and B be positive, compact operators. We write $A \lhd B$ if and only if $\mu_k(A) \leq \mu_k(B)$ for all k.

Definition : Let $\{A_n\}_{n=1}^{\infty}$ and A be positive, compact operators. We write $A_n \xrightarrow{\mu} A$ if and only if $\mu_k(A_n) \longrightarrow \mu_k(A)$ for each fixed k.

Remarks : 1) The topology defined by μ-convergence is, of course, non-Hausdorff.

*) A. Sloan Fellow

2) The order $\prec$ is useful because of the following consequence of the Ritz principle:

Proposition A.1 : Let A be a positive, compact operator and let P be a projection. Then $PAP \prec A$. In particular, if P and Q are projections and $P \leq Q$, then $PAP \prec QAQ$.

The above is false if $\prec$ is replaced by $\leq$.

Theorem A.2 : (Basic Convergence Theorem). Let B be a positive, compact operator with $S(B) < \infty$. Suppose $\{A_n\}$ and A are given positive, compact operators with

(1) $A_n \xrightarrow{\mu} A$

(2) $A_n \prec B$ for each n .

Then $\lim_{n \to \infty} S(A_n) = S(A)$.

Proof : The proof is based on the fact that s is monotone in $[o,e^{-1}]$. Since B is compact, $\mu_k(B) \to 0$. Suppose $\mu_N(B) \leq e^{-1}$. By (1) and the continuity of s, $s(\mu_k(A_n)) \longrightarrow s(\mu_k(A))$, each k, and by (2) and the monotonicity of s in $[0,e^{-1}]$, $s(\mu_k(A_n)) \leq s(\mu_k(B))$ for $k \geq N$, each n . Thus by the dominated convergence theorem for sums, $\sum_{k\geq N} s(\mu_k(A_n)) \longrightarrow \sum_{k\geq N} s(\mu_k(A))$. Since $\sum_{k\leq N-1} s(\mu_k(A_n))$ certainly converges , the theorem is proven . Q.E.D.

For applications of theorem A.2, it is convenient to have statements expressed in a more usual form than μ-convergence.

<u>Theorem A.3</u> : Let $\{A_n\}$ and A be positive, compact operators. If

(1) $\text{w-}\lim_{n\to\infty} A_n = A$ and

(2) $A_n \nless A$ for all n ,

then $\lim_{n\to\infty} S(A_n) = S(A)$.

<u>Proof</u> : We first prove that $A_n \xrightarrow{\mu} A$. Fix k and ϵ . By weak convergence and the min-max principle, it is easy to find a k-dimensional space, V , and an N such that

$$(\Psi, A_n \Psi) \geq (\mu_k(A) - \epsilon) \|\Psi\|^2$$

if $\Psi \in V$ and $n \geq N$. But then $\mu_k(A_n) \geq \mu_k(A) - \epsilon$ if $n \geq N$. Since $\mu_k(A) \geq \mu_k(A_n)$ by (2) , this means $|\mu_k(A) - \mu_k(A_n)| < \epsilon$ if $n \geq N$ and hence $A_n \xrightarrow{\mu} A$. If $S(A) < \infty$, the theorem then follows from Theorem A.2 . If $S(A) = \infty$, for any M we can find an L such that $\sum_{k=1}^{L} s(\mu_k(A)) > M$. However, for L sufficiently large, $S(A_n) \geq \sum_{k=1}^{L} s(\mu_k(A_n))$ and, since $\mu_k(A_n) \longrightarrow \mu_k(A)$, the latter sum can be made arbitrarily close to M . Thus $S(A_n) \longrightarrow \infty$. Q.E.D.

<u>Theorem A.4</u> : (<u>Dominated Convergence Theorem for Entropy</u>) : Let $\{A_n\}$, A and B be positive, compact operators and suppose that :

(1) $S(B) < \infty$

(2) $\text{w-}\lim_{n \to \infty} A_n = A$

(3) $A_n \leq B$ (operator inequality!).

Then, $\lim_{n \to \infty} S(A_n) = S(A)$.

<u>Proof</u> : Since B is compact, for any $\epsilon > 0$ we can find a finite-dimensional subspace $K \subset H$ such that $(u, B u) = \| B^{\frac{1}{2}} u \| < \epsilon \, \|u\|$ for $u \in L$, where L is the orthogonal complement of K . Since $A_n \leq B$, $\|A_n^{\frac{1}{2}} u\| = (u, A_n u) \leq (u, B u) \leq \epsilon \, \|u\|$ for all u in L . Since $A_n \xrightarrow{w} A$, $A \leq B$ and $\|A^{\frac{1}{2}} u\| \leq \epsilon \, \|u\|$ for all u in L also. We now show $A_n \to A$ uniformly. Recall that $\|A_n - A\| = \sup \{ |(\varphi , (A_n - A) \Psi)| : \varphi, \Psi \in H, \|\varphi\| = \|\Psi\| = 1 \}$. Now write $\varphi = f + u$, $\Psi = g + v$ where f, g are in K and u, v in L . Then

$$
\begin{aligned}
(\varphi, (A_n - A) \Psi) &= ((f + u), (A_n - A) (g + v)) \\
&\leq (f, (A_n - A)g) + \|A_n^{\frac{1}{2}} f\|^{\frac{1}{2}} \|A_n^{\frac{1}{2}} v\|^{\frac{1}{2}} \\
&\quad + \|A^{\frac{1}{2}} f\|^{\frac{1}{2}} \|A^{\frac{1}{2}} v\|^{\frac{1}{2}} + \|A_n^{\frac{1}{2}} u\|^{\frac{1}{2}} \|A_n^{\frac{1}{2}} g\|^{\frac{1}{2}} \\
&\quad + \|A^{\frac{1}{2}} u\|^{\frac{1}{2}} \|A^{\frac{1}{2}} g\|^{\frac{1}{2}} + \|A_n^{\frac{1}{2}} u\|^{\frac{1}{2}} \|A_n^{\frac{1}{2}} v\|^{\frac{1}{2}} \\
&\quad + \|A^{\frac{1}{2}} u\|^{\frac{1}{2}} \|A^{\frac{1}{2}} v\|^{\frac{1}{2}} ,
\end{aligned}
$$

which can be arbritrarily small since $A_n \longrightarrow A$ uniformly on K, $A_n^{\frac{1}{2}}$ and $A^{\frac{1}{2}}$ are bounded on K, $\|A_n^{\frac{1}{2}} u\| < \epsilon$, $\|A^{\frac{1}{2}} u\| < \epsilon$, etc..., and $\|f\| \leq \|\varphi\|$ etc... Thus $|(\varphi, (A_n - A)\Psi)|$ can be made arbitrarily small independent of φ, Ψ (for all φ, Ψ with $\|\varphi\| = \|\Psi\| = 1$) and thus $\|A_n - A\| \longrightarrow 0$. By the min-max principle, $|\mu_k(A_n) - \mu_k(A)| \leq \|A_n - A\|$. Thus $A_n \xrightarrow{\mu} A$, and (1) implies that Theorem A.2 is applicable. Q.E.D.

<u>Example</u> : Let $\{A_n\}$, A and B be the following operators on H, where $\{\varphi_n\}$ is an orthonormal basis for H :

$$A \varphi_k = 0 \text{ , each } k$$

$$A_n \varphi_k = \delta_{nk} e^{-1} \varphi_n$$

$$B = A_1 \quad .$$

Then $A_n \not\leq B$, $A_n \longrightarrow A$ strongly,but $S(A_n)$ does not converge to $S(A)$. This example shows that $\leq$ and not $\not\leq$ is needed in Theorem A.4.

REFERENCES.

[1] E.H. Lieb and M.B. Ruskai, Phys. Rev. Letters 30 , 434 (1973).

[2] H. Araki and E.H. Lieb, Commun. Math. Phys. 18, 160 (1970).

[3] F. Bauman and R. Jost, in Problems of Theoretical Physics ; Essays Dedicated to N.N. Bogoliubov, 285, Moscow, Nauka (1969).

[4] R. Jost, in "Quanta" -Essays in Theoretical Physics Dedicated to Gregor Wentzel, P.G.O. Freund, C.J. Goebel and Y. Nambu eds., 13, University of Chicago Press, Chicago (1970).

[5] F. Baumann, Helv. Phys. Acta 44, 95 (1971).

[6] D.W. Robinson and D. Ruelle, Commun. Math. Phys. 5 , 288 (1967).

[7] O. Lanford III and D.W. Robinson, J. Math. Phys. 9, 1120 (1968).

[8] E.P. Wigner and M.M. Yanase, Proc. Nat. Acad. Sci. 49, 910 (1963); Canad. J. Math. 16, 397 (1964).

[9] A. Uhlmann, "Endlich Dimensionale Dichtematrizen, II". Wiss. Z. Karl-Marx-Univ. Leipzig, Math-Naturwiss. R., to appear.

[10] D. Ruelle, Statistical Mechanics : Rigorous Results. Benjamin, New York (1969), Theorem 2.5.2.

[11] E.H. Lieb, "Convex Trace Functions and Proof of the Wigner-Yanase - Dyson Conjecture". To appear in Advances in Mathematics.

[12] H. Epstein, "Remark on Two Theorems of E. Lieb", Preprint

[13] M.B. Ruskai, "A Generalization of the Entropy Using Traces on von Neumann Algebras". Preprint.

[14] O. Lanford III, in Statistical Mechanics and Quantum Field Theory, C. De Witt and R. Stora eds., 174, Gordon and Breach, New York (1971).

I.H.E.S./P/73/046.