## Entropy of states of a gage space

ARTHUR LIEBERMAN

Let $(H, A, m)$ be a regular gage space. Let $\varrho, \sigma$, and $\psi=\lambda \varrho+(1-\lambda) \sigma, 0<\lambda<1$, be regular states. The density operator $D_{e}$ of a regular state is a non-negative (possibly unbounded) self-adjoint measurable operator. Let $F$ be a continuous convex function on $[0, \infty)$ and define the entropy of $\varrho$ by $e(\varrho)=m\left(F\left(D_{\varrho}\right)\right)$. Conditions are obtained, in terms of $e(\varrho)$ and $e(\sigma)$, for $e(\psi)$ to be $-\infty$, finite, $\infty$, or undefined. If both $\varrho$ and $\sigma$ have finite entropy, then $\psi$ has finite entropy and $e(\psi) \geqq \lambda e(\varrho)+$ $+(1-\lambda) e(\sigma)$; if $A=B(H), F$ is strictly convex, and $\varrho \neq \sigma$, then strict inequality is obtained. These results are restated as inequalities concerning the trace of a convex function of an operator.

## 1. Introduction

We work in the context of a regular gage space ( $H, A, m$ ); $H$ is a Hilbert space, $A$ is a von Neumann algebra on $H$, and $m$ is a faithful semi-finite normal trace on $A$. (See [4] for definitions and notation.) A regular state of $A$ is a positive linear functional $\varrho$ on $A$ with $\varrho(I)=1$, where $I$ is the identity operator on $H$, which is strongly continuous on the unit ball of $A$. If $\varrho$ is a regular state of $A$, then by [4] Theorem 14 there is a unique operator $D_{e} \in L^{1}(H, A, m)$ with $D_{e} \geqq 0, m\left(D_{e}\right)=1$, and $\varrho(T)=m\left(D_{Q} T\right)$ for all $T \in A ; D_{e}$ is called the density operator of $\varrho$.

The entropy of a regular state $\varrho$ is usually defined by $e(\varrho)=m\left(-D_{\varrho} \ln D_{\varrho}\right)$, cf. [3] Chapter V and [5]. Both von Neumann and Segal suggested defining the entropy by $e(\varrho)=m\left(F\left(D_{\varrho}\right)\right)$, where $F$ is an arbitrary continuous convex function on $[0, \infty)$; we use this definition for the remainder of this paper. The results basically say that the mixing of states cannot reduce entropy.

Bendat and Sherman [1] determined when a continuous convex function defined on an interval is operator convex; i.e., when $F(\lambda K+(1-\lambda) L) \geqq \lambda F(K)+$ $+(1-\lambda) F(L)$ holds for bounded self-adjoint operators $K$ and $L$ whose spectra

[^0]are contained in the domain of $F$. Below we show that $m(F(\lambda K+(1-\lambda) L)) \geqq$ $\geqq \lambda m(F(K))+(1-\lambda) m(F(L))$ holds under suitable hypotheses for self-adjoint measurable operators $K$ and $L$; this is merely a restatement of the fact that mixing of states cannot reduce entropy.

## 2. Statement of the results

Theorem 1. Let $(H, A, m)$ be a gage space with regular states $\varrho$ and $\sigma$. Let $0<\lambda<1$, and $\psi=\lambda \varrho+(1-\lambda) \sigma$. Assume $\lim \inf _{x \rightarrow \infty} F(x) / F(k x)>0$ for each $k>1$. Then:
A. $e(\psi)$ is defined iff both $e(\varrho)$ and $e(\sigma)$ are defined and $\{e(\varrho), e(\sigma)\} \neq\{-\infty, \infty\}$.
B. $e(\psi)$ is finite iff both $e(\varrho)$ and $e(\sigma)$ are finite.
C. $e(\psi)=\infty$ iff $\{\infty\} \subseteq\{e(\varrho), e(\sigma)\} \subseteq R \cup\{\infty\}$, where $R$ is the set of real numbers.
D. $e(\psi)=-\infty$ iff $\{-\infty\} \subseteq\{e(\varrho), e(\sigma)\} \subseteq\{-\infty\} \cup R$.

Corollary 1. Let $(H, A, m)$ be a gage space with regular states $\varrho$ and $\sigma$. Let $0<\lambda<1$, and $\psi=\lambda \varrho+(1-\lambda) \sigma$. Then
A. $e(\psi)$ is defined if both $e(\varrho)$ and $e(\sigma)$ are defined and $\{-\infty, \infty\} \neq\{e(\varrho), e(\sigma)\}$.
B. $e(\psi)$ is finite if both $e(\varrho)$ and $e(\sigma)$ are finite.
C. $e(\psi)=\infty$ if $\{\infty\} \subseteq\{e(\varrho), e(\sigma)\} \subseteq R \cup\{\infty\}$, and $\lim _{x \rightarrow \infty} F(x)=-\infty$.

Theorem 2. Let $(H, A, m)$ be a gage space with regular states $\varrho$ and $\sigma$. Let $0<\lambda<1$ and $\psi=\lambda \varrho+(1-\lambda) \sigma$. If $e(\varrho)$ and $e(\sigma)$ are finite, then $e(\psi)$ is finite and $e(\psi) \geqq \lambda e(\varrho)+(1-\lambda) e(\sigma)$. If. $A=B(H)=$ all bounded operators on $H, \varrho \neq \sigma$, and the function $F$ is strictly convex, then $e(\psi)>\lambda e(\varrho)+(1-\lambda) e(\sigma)$.

Corollary 2. Let $(H, A, m)$ be a gage space. Let $K, L \in L^{1}(H, A, m)$. Assume that either $K \geqq 0$ and $L \geqq 0$ or $m(I)<\infty$ and $K$ and $L$ are both bounded from below (or from above). Let $F$ be a continuous convex function defined on an interval which includes the spectra of $K$ and $L$ and let $0<\lambda<1$. If $F(K), F(L) \in L^{1}(H, A, m)$, then $\quad F(\lambda K+(1-\lambda) L) \in L^{1}(H, A, m)$, and $\quad m(F(\lambda K+(1-\lambda) L)) \geqq \lambda m(F(K))+$ $+(1-\lambda) m(F(L))$. If $A=B(H), \quad K \neq L, \quad$ and $F$ is strictly convex, then $m(F(\lambda K+(1-\lambda) L))>\lambda m(F(K))+(1-\lambda) m(F(L))$.

Remark. In Theorem 2 and Corollary 2, the restriction that $A=B(H)$ in order to have strict inequality seems unnecessary; this was first suggested by Segal [5]. We know of no example which requires this extra hypothesis, but are unable to prove strict inequality without it.

## 3. Proof of the results

Corollaries 1 and 2 are restatements of Theorems 1 and 2 and require no proof. We now introduce some notation. The self-adjoint operator $T$ has spectral decomposition $T=\int_{-\infty}^{\infty} \alpha d P_{T}(\alpha)$; the function $P_{T}$ is continuous from the left. If $S$ is a Borel measurable set of real numbers, then $P_{T}(S)$ is the spectral projection of $T$ for the set $S$. The spectral distribution function $\Lambda_{T}$ is defined by $\Lambda_{T}(x)=$ $\sup \left\{\lambda: m\left(P_{T}[\lambda, \infty)\right) \geqq x\right\}$; the domain of $\Lambda_{T}$ is $(0, m(I)]$ if $m(I)<\infty$ and $(0, \infty)$ if $m(I)=\infty . \Lambda_{T}(x)$ is a nonincreasing function of $x$ and is continuous from the left. $m\left(P_{\dot{T}}\left(\Lambda_{T}(x), \infty\right)\right)=x$ if $P$ has no point mass at $\Lambda_{T}(x)$ and $T \in L^{1}(H, A, m)$. The properties of the spectral distribution function are developed in [2]. To simplify the notation, we will frequently write $P_{\mathbf{e}}$ for $P_{D_{\boldsymbol{e}}}$ and $\Lambda_{\mathbf{q}}$ for $\Lambda_{D_{\mathbf{u}}}$.

Lemma 1. Let $(H, A, m)$ be a gage space, let $K \in L^{1}(H, A, m)$ with $K \geqq 0$, and let $F$ be a continuous function on $(r, \infty)$, where $P_{K}\{r\}=0$. Then $\int_{r}^{\infty} F(\lambda) d m\left(P_{K}(\lambda)\right)=$ $=\int_{0}^{m\left(P_{K}[r, \infty)\right)} F\left(\Lambda_{K}(x)\right) d x$ in the sense that if either integral is defined, then both integrals are defined and are equal. In addition, if $F$ is continuous on $[0, \infty)$, then

$$
\int_{[0, \infty)} F(\lambda) d m\left(P_{K}(\lambda)\right)=\int_{0}^{m(l)} F\left(\Lambda_{K}(x)\right) d x
$$

Proof. Let $s>r$ with $P_{K}\{s\}=0$. We will show below that $\int_{\boldsymbol{r}}^{s} F(\lambda) d m\left(P_{K}(\lambda)\right)=$ $=\int_{m\left(P_{K}[s, \infty)\right)}^{m\left(P_{K}[r ; \infty)\right)} F\left(\Lambda_{K}(x)\right) d x$. The first conclusion of the theorem will follow by taking the limit as $s \rightarrow \infty$. The second conclusion then follows by taking the limit as $r \rightarrow 0$. Let $P=\left\{x_{1}, x_{2}, \ldots, x_{n+1}\right\}$ be a partition of $[r, s]$ with $m\left(P_{K}\left\{x_{i}\right\}\right)=0$ for

$$
\begin{aligned}
& 1 \leqq i \leqq n+1 \text {. Then } \int_{r}^{s} F(\lambda) d m\left(P_{K}(\lambda)\right) \sim \sum_{i=1}^{n} F\left(\Lambda_{K}\left(m\left(P_{K}\left[x_{i}, \infty\right)\right)\right)\right) m\left(P_{K}\left[x_{i}, x_{i+1}\right]\right)= \\
& =\sum_{i=1}^{n} F\left(\Lambda_{K}\left(m\left(P_{K}\left[x_{i}, \infty\right)\right)\right)\right)\left(m\left(P_{K}\left[x_{i}, \infty\right)\right)-m\left(P_{K}\left[x_{i+1}, \infty\right)\right)\right) \sim \int_{m\left(P_{K}[s, \infty)\right)}^{m\left(P_{K}[r ; \infty)\right)} F\left(\Lambda_{K}(x)\right) d x .
\end{aligned}
$$

Note that, although $m\left(P_{K}\left[x_{i}, \infty\right)\right)-m\left(P_{K}\left[x_{i+1}, \infty\right)\right)$ may be large due to the spectrum of $K$ having point masses in the interval $\left(x_{i}, x_{i+1}\right), F\left(\Lambda_{K}(\alpha)\right)$ is nearly constant on the interval $m\left(P_{K}\left[x_{i+1}, \infty\right)\right)<\alpha \leqq m\left(P_{K}\left[x_{i}, \infty\right)\right)$ since for $\alpha$ in this interval, $x_{i} \leqq \Lambda_{K}(\alpha)<x_{i+1}$.

Proof of Theorem 1. There are essentially four different non-trivial possibilities for $F$ :
A.

$$
F^{\prime}(0)>0, \lim _{x \rightarrow \infty} F(x)=-\infty
$$

B.

$$
F^{\prime}(0)>0, \lim _{x \rightarrow \infty} F(x)=\infty
$$

C.

$$
F^{\prime}(0)>0, \lim _{x \rightarrow \infty} F(x)=k, \quad \text { where } 0<k<\infty .
$$

D.

$$
F^{\prime}(0) \leqq 0, \lim _{x \rightarrow \infty} F(x)=-\infty
$$

Theorem 1 will be proved for case A since this is the most difficult case; the proofs for the other cases are trivial modifications and parts of the results are vacuous in the other cases. For the sake of simplicity, we assume $F(0)=0$; if $F(0) \neq 0$, little change is needed if $m(I)$ is finite and the results become essentially vacuous if $m(I)=\infty$. We further assume that $F$ has a relative maximum at $x=1$, $F(1)=1$, and that $F(2)=0$. We will prove the "if" parts of B, C and D. The remainder of the proof is essentially redundant.

Assume now that $e(\varrho)$ and $e(\sigma)$ are both finite. $e(\psi)$ can be infinite in two ways: $\psi$ can be highly concentrated so that $D_{\psi}$ is unbounded and $e(\psi)=-\infty$, or $\psi$ can be so spread out that $D_{\psi}$ has very large support and $e(\psi)=\infty$.

Let $\alpha>0$ and $x \in H, x \neq 0$. If $P_{\psi}[\alpha, \infty) x=x$, then $\lambda\left(D_{e} x, x\right)+(1-\lambda)\left(D_{\sigma} x, x\right) \geqq$ $\geqq \alpha\|x\|^{2}$, so that either $P_{e}[\alpha, \infty) x \neq 0$ or $P_{\sigma}[\alpha, \infty) x \neq 0$. By [2, lemma 2], $m\left(P_{\psi}[\alpha, \infty)\right) \leqq m\left(P_{\ell}[\alpha, \infty)\right)+m\left(P_{\sigma}[\alpha, \infty)\right)$. Then

$$
\int_{2}^{\infty} F(\alpha) d m\left(P_{\psi}(\alpha)\right) \geqq \int_{2}^{\infty} F(\alpha) d m\left(P_{e}(\alpha)\right)+\int_{2}^{\infty} F(\alpha) d m\left(P_{\sigma}(\alpha)\right)>-\infty,
$$

have finite entropy.
Now let $0<\alpha<1$. $m\left(P_{\psi}(\alpha, 1]\right)=m\left(P_{\psi}(\alpha, \infty)\right)-m\left(P_{\psi}(1, \infty)\right) \leqq m\left(P_{e}(\alpha, \infty)\right)+$ $+m\left(P_{\sigma}(\alpha, \infty)\right)-m\left(P_{\psi}(1, \infty)\right)=m\left(P_{\ell}(\alpha, 1]\right)+m\left(P_{\sigma}(\alpha, 1]\right)+m\left(P_{\ell}(1, \infty)\right)+m\left(P_{\sigma}(1, \infty)\right)-$ $-m\left(P_{\psi}(1, \infty)\right)$. Let $c=m\left(P_{\rho}(1, \infty)\right)+m\left(P_{\sigma}(1, \infty)\right)-m\left(P_{\psi}(1, \infty)\right)$. Then $0 \leqq c<\infty$, and $m\left(P_{\psi}(\alpha, 1]\right) \leqq m\left(P_{\mathrm{e}}(\alpha, 1]\right)+m\left(P_{\sigma}(\alpha, 1]\right)+c$. Let $M$ be the unique Borel measure on $(0,1]$ such that $M(\alpha, 1]=m\left(P_{e}(\alpha, 1]\right)+m\left(P_{\sigma}(\alpha, 1]\right)+c$. Then $\int_{(0,1]} F(\alpha) d m\left(P_{\psi}(\alpha)\right) \leqq$ $\leqq \int_{(0,1]} F(\alpha) d M(\alpha)$, since $F$ is non-negative and non-decreasing on $(0,1]$, and $\int_{(0,1]}^{(0,1]} F(\alpha) d M(\alpha)<\infty$ since $\varrho$ and $\sigma$ have finite entropy.

We now prove part C. Assume $e(\varrho)=\infty$. Then $\int_{0}^{1} F(\alpha) d m\left(P_{e}(\alpha)\right)=\infty$, and by lemma $1, \int_{c}^{\infty} F\left(\Lambda_{\ell}(\alpha)\right) d \alpha=\infty$ for some $c$ such that $\Lambda_{\ell}(c) \leqq 1$ and $\Lambda_{\psi}(c) \leqq 1$. Note that $F$ is non-negative and non-decreasing on $[0,1]$. Since $\psi=\lambda \varrho+(1-\lambda) \sigma$,
$D_{\phi} \geqq \lambda D_{Q}$, so by [2] Corollary 1, $\Lambda_{\psi}(\alpha) \geqq \Lambda_{\lambda D_{Q}}(\alpha)=\lambda \Lambda_{Q}(\alpha)$. Then for $\alpha \geqq c$, $F\left(\Lambda_{\phi}(\alpha)\right) \geqq F\left(\lambda \Lambda_{e}(\alpha)\right) \geqq \lambda F\left(\Lambda_{e}(\alpha)\right)$ by convexity. Then

$$
\int_{c}^{\infty} F\left(\Lambda_{\phi}(\alpha)\right) d \alpha \geqq \lambda \int_{c}^{\infty} F\left(\Lambda_{\mathbf{e}}(\alpha)\right) d \alpha=\infty .
$$

We now prove part D. Assume $e(\varrho)=-\infty$, so that $\int_{0}^{\infty} F(\alpha) d m\left(P_{e}(\alpha)\right)=-\infty$. Choose $\varepsilon>0$ and $q>0$ so that $F(x) / F(x / \lambda) \geqq \varepsilon$ for $x \geqq \lambda q$. Since $D_{\psi} \geqq \lambda D_{e}$, $m\left(P_{\psi}[\alpha, \infty)\right) \geqq m\left(P_{\lambda D_{e}}[\alpha, \infty)\right)=m\left(P_{e}[\alpha / \lambda, \infty)\right)$. Then $\quad \int_{0}^{\infty} F(\alpha) d m\left(P_{e}(\alpha)\right)=-\infty$ implies $\int_{q}^{\infty} F(\alpha) d m\left(P_{e}(\alpha)\right)=-\infty$, so that $\int_{\lambda q}^{\infty} F(\alpha / \lambda) d m\left(P_{e}(\alpha / \lambda)\right)=-\infty$. Then $\int_{\lambda q}^{\infty} F(\alpha / \lambda) d m\left(P_{\psi}(\alpha)\right)=-\infty$ so $\int_{\lambda q}^{\infty} F(\alpha) d m\left(P_{\psi}(\alpha)\right)=-\infty$ and $e(\psi)=-\infty$.

Lemma 2. Let $R$ and $S$ be either finite sequences with the same number of members or countable sequences. Assume $r_{k} \geqq r_{k+1} \geqq 0, s_{k} \geqq s_{k+1} \geqq 0, \sum_{k} r_{k}=\sum_{k} s_{k}$, and $\sum_{k=1}^{j} r_{k} \geqq \sum_{k=1}^{j} s_{k}$ for $j \geqq 1$. Then there is a doubly stochastic matrix ${ }^{k} M$ with $s_{j}=\sum_{k} m_{j k} r_{k}$ for $j \geqq 1$.

Proof. If $R$ and $S$ are finite sequences the result is well known; our proof will contain this case if $R$ and $S$ are extended to countable sequences by adding a string of zeroes at the end. Let $R$ and $S$ be countable sequences and assume $r_{k} \neq 0$ for all $k$. $M$ will be constructed one row at a time; each row of $M$ will have finitely many non-zero entries. Let $w(1)$ be the smallest integer such that $s_{1} \geqq r_{w(1)}$. Express $s_{1}$ as a convex combination of $\left\{r_{i}: 1 \leqq i \leqq w(1)\right\}$ to obtain the first row of $M$.

Assume $k-1$ rows of $M$ have been obtained. If $s_{k} \geqq r_{w(k-1)}$, let $w(k)=1+$ $+w(k-1)$; otherwise, let $w(k)$ be the smallest integer such that $s_{k} \geqq r_{w(k)}$. We will show that $s_{k}$ can be expressed as a convex combination of $\left\{r_{i}: 1 \leqq i \leqq w(k)\right\}$ such that $\sum_{i=1}^{k} m_{i j} \leqq 1$ for $1 \leqq j \leqq w(k)$ by showing that there is such a convex combination which is $\geqq s_{k}$ and that there is such a convex combination (namely, $\left.\sum_{i=1}^{w(k)-1} 0 r_{i}+1 r_{w(k)}\right)$ which is $\leqq s_{k}$.

When $\sum_{i} m_{i j}=1$, we will say $r_{j}$ is "used up". Let the number $c=\sum_{i=1}^{k} c_{i} \dot{r}_{i}$ be formed as follows: $c_{1}$ is chosen so that $r_{1}$ is used up; i.e., $c_{1}=1-\sum_{i=1}^{k-1} m_{i 1}$. Choose $c_{2}$ so that $c_{1}+c_{2} \leqq 1$ and $r_{2}$ is used up if possible; $c_{2}=\min \left(1-c_{1}, 1-\sum_{i=1}^{k-1} m_{i 2}\right)$. Continue this process until $c_{k}$ is chosen. Then $c \geqq s_{k}$ follows from the hypothesis that $\sum_{i=1}^{k} r_{i} \geqq \sum_{i=1}^{k} s_{i}$.

This completes the construction of the matrix $M$. Clearly $s_{j}=\sum_{k} m_{j k} r_{k}$ for all $j, m_{j k} \geqq 0$ for all $j, k$, and the sum of the elements of any row of $M$ is 1 . It remains to show that the sum of the elements of any column of $M$ is $1.1=\sum_{i} s_{i}=\sum_{i} \sum_{j} m_{i j} r_{j}=$ $=\sum_{j} \sum_{i} m_{i j} r_{j}$; since all terms are non-negative the interchange of order of summation is valid. Since $1=\sum_{j} r_{j}, 0=\sum_{j}\left(1-\sum_{i} m_{i j}\right) r_{j}$. By the construction of $M$, $\left(1-\sum_{i} m_{i j}\right) \geqq 0$ for each $j$. Since $r_{j} \neq 0$ for all $j, 1=\sum_{i} m_{i j}$.

If $r_{j}=0$ for some $j$, then $r_{k}=0$ for all $k \geqq j$. The construction of $M$ must then be modified so that, for $k \geqq j, r_{k}$ is used up before one begins to use $r_{k+1}$.

Lemma 3. Let $(H, A, m)$ be a gage space, let $T \in L^{1}(H, A, m)$ with $T \geqq 0$, let $\gamma>0$, and let $q=m\left(P_{T}(\gamma, \infty)\right)$. Let $P$ be any projection in $A$ with $m(P)=q$. Then $m(P T) \leqq m\left(P_{T}(\gamma, \infty) T\right)$.

Proof. By lemma $1, \quad m(P T)=m(P T P)=\int_{0}^{m(I)} \Lambda_{P T P}(x) d x=\int_{0}^{q} \Lambda_{P T P}(x) d x$. By [2] Theorem 4, $\Lambda_{P T P}(x) \leqq \Lambda_{T}(x)$ for $0<x \leqq m(I)$. Note that $\Lambda_{P_{T}(y, \infty) T}(x)=$ $=\Lambda_{T}(x)$ for $0<x \leqq q$ so that $\Lambda_{P T P}(x) \leqq \Lambda_{P_{T}(\gamma, \infty) T}(x)$ for $0<x \leqq q$. Then

$$
\int_{0}^{q} \Lambda_{P T P}(x) d x \leqq \int_{0}^{q} \Lambda_{P_{T}(\gamma, \infty) T}(x) d x=\int_{0}^{m(I)} \Lambda_{P_{T}(\gamma, \infty) T}(x) d x=m\left(P_{T}(\gamma, \infty) T\right) .
$$

Proof of Theorem 2. Assume first that $A=B(H)$. Let $\varrho_{i}$ be the $i^{\text {th }}$ eigenvalue of $D_{e}$, where the eigenvalues of $D_{e}$ are arranged in decreasing order and are counted according to multiplicity.

Define a sequence $A$ by $a_{i}=\lambda \varrho_{i}+(1-\lambda) \sigma_{i}$ and a sequence $B$ by $b_{i}=\psi_{i}$. The first three hypotheses of lemma 2 are clearly satisfied. The last hypothesis of lemma 2 follows from lemma 3; a trivial modification of lemma 3 is needed if $D_{\boldsymbol{e}}$ or $D_{\sigma}$ has a repeated eigenvalue. By lemma 2, there is a doubly stochastic matrix $M$ with $\psi_{i}=\sum_{j} m_{i j}\left(\lambda \varrho_{j}+(1-\lambda) \sigma_{j}\right)$. Then $F\left(\psi_{i}\right) \geqq \sum_{j} m_{i j}\left(\lambda F\left(\varrho_{j}\right)+(1-\lambda) F\left(\sigma_{j}\right)\right)$. Summingthis relation yields $m\left(F\left(D_{\psi}\right)\right)=\sum_{i} F\left(\psi_{i}\right) \geqq \sum_{i} \sum_{j}^{J} m_{i j}\left(\lambda F\left(\varrho_{i}\right)+(1-\lambda) F\left(\sigma_{j}\right)\right)=$ $=\sum_{j} \sum_{i} m_{i j}\left(\lambda F\left(\varrho_{j}\right)+(1-\lambda) F\left(\sigma_{j}\right)\right)=\sum_{j}\left(\lambda F\left(\varrho_{j}\right)+(1-\lambda) F\left(\sigma_{j}\right)\right)=\lambda m\left(F\left(D_{e}\right)\right)+(1-\lambda) \cdot$ - $m\left(F\left(D_{\sigma}\right)\right)$; the interchange of the order of summation is valid since $\varrho$ and $\sigma$ e ach have finite entropy by hypothesis. If $\varrho \neq \sigma$, then $\psi_{i_{0}} \neq \lambda \varrho_{i_{0}}+(1-\lambda) \sigma_{i_{0}}$ for some $i_{0}$, so that $M$ is not the identity matrix. If $F$ is then strictly convex, then $F\left(\psi_{i_{0}}\right)>$ $>\sum_{j} m_{i_{0} j}\left(\lambda F\left(\varrho_{j}\right)+(1-\lambda) F\left(\sigma_{j}\right)\right)$.

We now prove the general case when $m(I)=\infty$; the proof when $m(I)<\infty$ is virtually identical. Let $\varepsilon$ be an arbitrary positive number. For $n$ a natural number, let

$$
\psi_{n}=\frac{1}{\varepsilon} \int_{(n-1) \varepsilon}^{n \varepsilon} \Lambda_{\psi}(x) d x
$$

define sequences $\varrho_{n}$ and $\sigma_{n}$ similarly. Assume that $D_{\psi}, D_{\varrho}, D_{\sigma}$ have no point masses at $\Lambda_{\psi}(k \varepsilon), \Lambda_{e}(k \varepsilon), \Lambda_{\sigma}(k \varepsilon)$ respectively, for all natural numbers $k$; arbitrarily small $\varepsilon$ can always be found so that this holds. By lemma 1 and lemma 3,

$$
\begin{gathered}
\varepsilon \sum_{n=1}^{k} \psi_{n}=\int_{0}^{k \varepsilon} \Lambda_{\psi}(x) d x=\int_{\Lambda_{\psi}(k \varepsilon)}^{\infty} \alpha d m\left(P_{\psi}(\alpha)\right)=m\left(D_{\psi} P_{\psi}\left(\Lambda_{\psi}(k \varepsilon, \infty)\right)\right)= \\
=\lambda m\left(D_{Q} P_{\psi}\left(\Lambda_{\psi}(k \varepsilon, \infty)\right)\right)+(1-\lambda) m\left(D_{\sigma} P_{\psi}\left(\Lambda_{\psi}(k \varepsilon, \infty)\right)\right) \leqq \\
\leqq \lambda m\left(D_{Q} P_{Q}\left(\Lambda_{\ell}(k \varepsilon, \infty)\right)\right)+(1-\lambda) m\left(D_{\sigma} P_{\sigma}\left(\Lambda_{\sigma}(k \varepsilon, \infty)\right)\right)=\varepsilon \lambda \sum_{n=1}^{k} \varrho_{n}+\varepsilon(1-\lambda) \sum_{n=1}^{k} \sigma_{n}
\end{gathered}
$$

By the first part of the proof of this theorem,

$$
\sum_{n} F\left(\psi_{n}\right) \geqq \lambda \sum_{n=1}^{\infty} F\left(\varrho_{n}\right)+(1-\lambda) \sum_{n=1}^{\infty} F\left(\sigma_{n}\right) .
$$

To complete the proof, it suffices to show that $\varepsilon \sum_{i=1}^{n} F\left(\psi_{n}\right)$ approximates $\int_{0}^{\infty} F\left(\Lambda_{\psi}(x)\right) d x$ for $\varepsilon$ small. This is immediate since $\Lambda_{\psi}$ is a non-increasing function implies $\Lambda_{\psi}((n-1) \varepsilon) \geqq \psi_{n} \geqq \Lambda_{\psi}(n \varepsilon)$ and $e(\psi)$ is finite by the hypotheses and Corollary 1.

Acknowledgement. This research was started while the author was a guest at M. I. T. and was completed at The Cleveland State University. The author wishes to thank M. I. T. for providing hospitality. This paper answers some questions proposed by Segal in [5]. The author wishes to thank Professor Segal for suggesting this topic to him and for many useful conversations while the research was in progress.

## Bibliography

[1] J. Bendat and S. Sherman, Monotone and convex operator functions, Trans. Amer. Math. Soc., 79 (1955), 58-71.
[2] A. Lieberman, Spectral distribution of the sum of self-adjoint operators, Pacific J. Math. 53 (1974), 211-216.
[3] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press (Princeton, 1955).
[4] I. E. Segal, A non-commutative extension of abstract integration, Ann. Math., 57 (1953), 401457.
[5] I. E. Segal, A note on the concept of entropy, J. Math. Mech., 9 (1960), 623-630.
[6] H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, Math. Ann., 71 (1912), 441-479.


[^0]:    Received January 25, 1977.

