37. A Note on the Entropy for Operator Algebras

By Masahiro NAKAMURA*) and Hisaharu UMEGAKI**)

(Comm. by K. KUNUGI, M.J.A., March 13, 1961)

Recently, I. E. Segal [9] established the notion of the entropy of states of semi-finite von Neumann algebras. Segal's entropy contains the cases of the information theory, e.g. A. I. Khinchin [5], and the quantum statistical mechanics due to J. von Neumann [8]. The purpose of the present note is to discover the background of Segal's definition basing on a study of the so-called convex operator functions due to originally C. Loewner and extensively J. Bendat and S. Sherman [1].

1. A real-valued continuous function f defined on an interval I will be called *operator-convex* in the sense of Loewner-Bendat-Sherman provided that

$$f(\alpha a+\beta b)\leqq\alpha f(a)+\beta f(b), \tag{1}$$

for any hermitean operators a and b having their spectra in I, and for any non-negative real numbers α and β with $\alpha+\beta=1$. According to a theorem of Bendat-Sherman [1; Theorem 3.5], an analytic function,

$$f(\lambda)=\sum_{i=2}^{\infty}\gamma_i\lambda^i, \tag{2}$$

with the convergence radius R, is operator-convex for $|\lambda|<R$ if and only if

$$\sum_{i,k=0}^{n}\frac{f^{(i+k+2)}(0)}{(i+k+2)!}\alpha_i\alpha_k\geqq 0, \tag{3}$$

for any sequence of real numbers α_i and for all n.

LEMMA 1. *$\lambda\log(1+\lambda)$ is operator-convex for $|\lambda|<1$.*

Proof. Put $f(\lambda)=\lambda\log(1+\lambda)$. Clearly f satisfies (2) for $R=1$. Calculating, for $k=2,3,\cdots$,

$$f^{(k)}(\lambda)=(-1)^k[(k-2)!\,(1+\lambda)^{-(k-1)}+(k-1)!\,(1+\lambda)^{-k}].$$

Putting $\lambda=0$, one has $f^{(k)}(0)=(-1)^k(k-2)!\,k$ for $k=2,3,\cdots$. Applying (3), one has, for any real numbers α_i,

$$\sum_{i,k=0}^{n}\frac{f^{(i+k+2)}(0)}{(i+k+2)!}\alpha_i\alpha_k=\sum_{i,k=0}^{n}(-1)^{i+k}\frac{(i+k)!\,(i+k+2)}{(i+k+2)!}\alpha_i\alpha_k$$

$$=\sum_{i,k=0}^{n}(-1)^{i+k}\frac{\alpha_i\alpha_k}{i+k+1}.$$

Replacing $(-1)^i\alpha_i$ by α_i, it is non-negative, since the matrix,

*) Osaka Gakugei Daigaku.
**) Tokyo Institute of Technology.

$$C_n=\begin{pmatrix} \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ & \cdots\cdots & & \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix},$$

is positive definite, according to the well-known Hilbert's formula:

$$(4)\qquad \det C_n=\frac{[2!\,3!\cdots(n-1)!]^3}{n!\,(n+1)!\cdots(2n-1)!}>0,$$

for $n=1,2,\cdots$, which proves the lemma.

It is noteworthy that Lemma 1 is implied also by an another theorem of Bendat-Sherman [1; Theorem 3.4] since $\log(1+\lambda)$ is clearly operator-monotone.

LEMMA 2. *$(1+\lambda)\log(1+\lambda)$ is operator-convex for $|\lambda|<1$.*

It is clearly sufficient to show the lemma that the function,

$$(5)\qquad g(\lambda)=(1+\lambda)\log(1+\lambda)-\lambda,$$

is operator-convex for $|\lambda|<1$. By computation, (5) implies

$$g^{(k)}(\lambda)=(-1)^k(k-2)!\,(1+\lambda)^{-(k-1)},$$

for $k=2,3,\cdots$. Since g is analytic for $|\lambda|<1$ and satisfies (2), it is also sufficient to prove the lemma that

$$\sum_{i,k=0}^{n}\frac{g^{(i+k+2)}(0)}{(i+k+2)!}\alpha_i\alpha_k=\sum_{i,k=0}^{n}(-1)^{i+k}\frac{(i+k)!}{(i+k+2)!}\alpha_i\alpha_k\geqq 0,$$

for all n and for any real numbers α_i. Therefore, it is sufficient to show the lemma that the matrix,

$$D_{n+1}=\begin{pmatrix} \frac{1}{1\cdot 2} & \frac{1}{2\cdot 3} & \frac{1}{3\cdot 4} & \cdots & \frac{1}{(n+1)(n+2)} \\ \frac{1}{2\cdot 3} & \frac{1}{3\cdot 4} & \frac{1}{4\cdot 5} & \cdots & \frac{1}{(n+2)(n+3)} \\ & & \cdots\cdots\cdots\cdots & & \\ \frac{1}{(n+1)(n+2)} & & & \cdots & \frac{1}{(2n+1)(2n+2)} \end{pmatrix},$$

is non-negative definite, for $n=1,2,\cdots$. It will be shown in the next section that $\det D_n>0$ for $n=1,2,\cdots$.

2. In this section, Lemma 2 is proved by establishing a determinant formula in the following

LEMMA 3. *For $n=1,2,\cdots$,*

$$(6)\qquad \det D_n=\begin{vmatrix} \frac{1}{1\cdot 2} & \frac{1}{2\cdot 3} & \cdots & \frac{1}{n(n+1)} \\ \frac{1}{2\cdot 3} & \frac{1}{3\cdot 4} & \cdots & \frac{1}{(n+1)(n+2)} \\ & \cdots\cdots\cdots\cdots & & \\ \frac{1}{n(n+1)} & & \cdots & \frac{1}{(2n-1)2n} \end{vmatrix}=\frac{[(n-1)!\,(n-2)!\cdots 3!\,2!]^3 n!}{(n+1)!\,(n+2)!\cdots(2n-1)!\,2n!}.$$

Our proof is an imitation of a proof which gives Hilbert's formula (4) from Cauchy's formula:

$$\begin{vmatrix} \frac{1}{x_1-a_1} & \frac{1}{x_1-a_2} & \cdots & \frac{1}{x_1-a_n} \\ \frac{1}{x_2-a_1} & \frac{1}{x_2-a_2} & \cdots & \frac{1}{x_2-a_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{x_n-a_1} & \frac{1}{x_n-a_2} & \cdots & \frac{1}{x_n-a_n} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \frac{\prod_{i>k}(x_i-x_k)(a_i-a_k)}{\prod_{i=1}^{n} p(x_i)},$$

where $p(x)=\prod_{k=1}^{n}(x-a_k)$.

Proof. Put $p(x)=\prod_{k=1}^{n+1}(x-a_k)$,

$$D=\begin{vmatrix} \frac{1}{(x_1-a_1)(x_1-a_2)} & \frac{1}{(x_1-a_2)(x_1-a_3)} & \cdots & \frac{1}{(x_1-a_n)(x_1-a_{n+1})} \\ \frac{1}{(x_2-a_1)(x_2-a_2)} & \frac{1}{(x_2-a_2)(x_2-a_3)} & \cdots & \frac{1}{(x_2-a_n)(x_2-a_{n+1})} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{(x_n-a_1)(x_n-a_2)} & \frac{1}{(x_n-a_2)(x_n-a_3)} & \cdots & \frac{1}{(x_n-a_n)(x_n-a_{n+1})} \end{vmatrix},$$

and $C=D\cdot\prod_{i=1}^{n} p(x_i)$. Then we have

$$C=\begin{vmatrix} \prod_{k\neq 1,2}(x_1-a_k) & \prod_{k\neq 2,3}(x_1-a_k) & \cdots & \prod_{k\neq n,n+1}(x_1-a_k) \\ \prod_{k\neq 1,2}(x_2-a_k) & \prod_{k\neq 2,3}(x_2-a_k) & \cdots & \prod_{k\neq n,n+1}(x_2-a_k) \\ \cdots & \cdots & \cdots & \cdots \\ \prod_{k\neq 1,2}(x_n-a_k) & \prod_{k\neq 2,3}(x_n-a_k) & \cdots & \prod_{k\neq n,n+1}(x_n-a_k) \end{vmatrix}.$$

Since C is divisible by x_i-x_k for $i\neq k$, $i, k=1, 2, \cdots, n$, C is also divisible by $\prod(x_i-x_k)$ $(n\geqq i>k\geqq 1)$. Similarly, D is divisible by a_i-a_k for $i-k\neq 1$. Hence C is divisible by

$$\prod_{n+1\geqq i>k\geqq 1}(a_i-a_k)/\prod_{n\geqq k\geqq 1}(a_{k+1}-a_k),$$

that is,

$$C=c\cdot\frac{\prod_{n\geqq i>k\geqq 1}(x_i-x_k)\prod_{n+1\geqq i>k\geqq 1}(a_i-a_k)}{\prod_{k=1}^{n}(a_{k+1}-a_k)}.$$

Comparing with the order of the both sides of the above equality, c is known as a constant. Also comparing with the corresponding terms of the expansions, $c=(-1)^{n(n-1)}/(-1)^{n(n-1)/2}=(-1)^{n(n-1)/2}$. Therefore, one has

$$(7)\qquad D=(-1)^{\frac{n(n-1)}{2}}\frac{\prod_{n\geqq i>k\geqq 1}(x_i-x_k)\prod_{n+1\geqq i>k\geqq 1}(a_i-a_k)}{\prod_{k=1}^{n}(a_{k+1}-a_k)\prod_{i=1}^{n}p(x_i)}.$$

Finally, it is shown that (7) implies (6) assuming

$$(8)\qquad x_i - a_k = i + k - 1.$$

Since (8) implies at once $x_k - a_i = k + i - 1 = x_i - a_k$, one has

$$x_i - x_k = i - k,\quad a_i - a_k = k - i,\quad a_{k+1} - a_k = -1,$$

and $p(x_i) = \prod_{k=1}^{n+1}(x_i - a_k) = \prod_{k=1}^{n+1}(i+k-1)$. Replacing them in (7), one has

$$\det D_n = D = (-1)^{\frac{n(n-1)}{2}} \frac{\prod_{n\geqq i>k\geqq 1}(i-k) \prod_{n+1\geqq i>k\geqq 1}(k-i)}{\prod_{k=1}^{n}(-1)^k \prod_{i=1}^{n}\prod_{k=1}^{n+1}(i+k-1)}$$

$$= (-1)^{\frac{n(n-1)}{2}+\frac{n(n+1)}{2}-n} \frac{\left[\left(\prod_{k=1}^{n-1}(n-k)\right)\cdot\left(\prod_{k=1}^{n-2}(n-1-k)\right)\cdots 2\cdot 1\right]^2 \prod_{k=1}^{n}(n-k+1)}{\prod_{i=1}^{n}[i(i+1)\cdots(i+n)]}$$

$$= (-1)^{n(n-1)} \frac{[(n-1)!\,(n-2)!\cdots 3!\,2!]^2[n(n-1)\cdots 2\cdot 1]}{n!\,(n+1)!\,\frac{(n+2)!}{2!}\cdots\frac{(2n-1)!}{(n-1)!}\,\frac{2n!}{n!}}$$

$$= \frac{[(n-1)!\cdots 3!\,2!]^3 n!}{(n+1)!\,(n+2)!\cdots 2n!},$$

which proves the lemma.

3. In this section, the following theorem will be proved:

THEOREM 1. *$f(\lambda)=\lambda\log\lambda$ is operator-convex for $\lambda\geqq 0$, where $f(0)$ is defined by*

$$(9)\qquad f(0)=0.$$

The proof will be divided into two steps. At first, it will be shown that f is operator-convex in $[0,2\rangle$. If a and b are two hermitean operators having their spectra in $[0,2\rangle$, then there exist nets $\{a_\delta\}$ and $\{b_\delta\}$ which have their spectra in $\langle 0,2\rangle$ and converge strongly to a and b respectively. For each δ Lemma 2 implies $f(\alpha a_\delta + \beta b_\delta) \leqq \alpha f(a_\delta) + \beta f(b_\delta)$, where $\alpha, \beta \geqq 0$ with $\alpha+\beta=1$. By the continuity of $f(\lambda)$ for $\lambda\geqq 0$ and by a lemma of Kaplansky [4], the mapping $a\to f(a)$ is strongly continuous on $0\leqq a<2$, whence the above inequality implies (1) for $I=[0,2\rangle$.

Now let c and d are non-negative hermitean operators, and choose a constant $k>0$ such that $a=c/k$ and $b=d/k$ have their spectra in $[0,2\rangle$. By the above, a and b satisfy (1). Hence,

$$\frac{1}{k}(\alpha c+\beta d)[\log(\alpha c+\beta d)-\log k]$$

$$\leqq \frac{1}{k}[\alpha c\log c+\beta d\log d-(\alpha c+\beta d)\log k],$$

which implies that c and d satisfy (1) in place of a and b. This completes the proof of the theorem.

4. Suppose that A is a semi-finite von Neumann algebra in the sense of J. Dixmier [3] having a normal trace or a gage τ. If a is

a non-negative hermitean member of A, then $a \log a$ belongs to A too. Hence the following definition has a meaning:

DEFINITION 1. For a non-negative hermitean a of A, the *entropy* of a is defined by

$$H(a) = -\tau(a \log a). \tag{10}$$

If τ is semi-finite, a is called to have the *bounded entropy* provided that $H(a)$ is finite. If τ is finite, the entropy is always finite.

Since $f(a) = a \log a$ satisfies (1) by Theorem 1, and τ is monotone, the definition implies at once

THEOREM 2. *The entropy is concave on A^+, that is,*

$$H(\alpha a + \beta b) \geqq \alpha H(a) + \beta H(b),\text{[1]} \tag{11}$$

for non-negative hermiteans a and b of A, where $\alpha, \beta \geqq 0$ and $\alpha + \beta = 1$.

Moreover, the following theorem holds:

THEOREM 3. *The entropy does not decrease after an application of the conditional expectation ϵ conditioned by a von Neumann subalgebra B in the sense of* [10, I and II]:[2]

$$H(a^\epsilon) \geqq H(a). \tag{12}$$

Proof.[1] Since $\lambda \log \lambda$ is operator-convex by Theorem 1 and satisfies (9), theorems in [2] and [6] imply

$$a^\epsilon \log a^\epsilon \leqq [a \log a]^\epsilon. \tag{13}$$

By the monotonity of τ, one has

$$H(a^\epsilon) = -\tau(a^\epsilon \log a^\epsilon) \geqq -\tau([a \log a]^\epsilon) = -\tau(a \log a) = H(a),$$

which is desired.

Since Segal [9] defined the entropy of a state σ of A by the Radon-Nikodym derivative a of σ with respect to the trace τ, the above theorems imply the corresponding theorems of Segal.

To conclude the note, it may be observed with some interests, that Theorem 3 allows us to introduce the following

DEFINITION 2. If B is a von Neuman subalgebra of A, and if a is a non-negative hermitean element of A, then the *information of a with respect to B* is defined by

$$I(a; B) = H(a^\epsilon) - H(a), \tag{14}$$

where a^ϵ is the conditional expectation of a conditioned by B. By Theorem 3, $I(a; B)$ is non-negative.

---

1) By the same methods, we can prove the followings: For a state σ of a C*-algebra, we define $H_\sigma(a) = -\sigma(a \log a)$ $(a \geqq 0)$ and call it by σ-entropy. If the ϵ is an expectation ϵ_σ on a finite von Neumann algebra in the sense of [10, III], then the inequalities (11) and (12) for H_σ in places of H also hold, where σ is a normal state in the tracelet space defined by the von Neumann subalgebra. More generally, if the ϵ is an expectation on a C*-algebra A in the sense of [7], then the same facts also hold for any state σ, invariant by ϵ, that is, $\sigma(a) = \sigma(a^\epsilon)$ for all $a \in A$.

2) The subalgebra B is clearly defined such that the restriction of τ onto B, is also a gage. Therefore, $\tau(a) = \tau(a^\epsilon)$ for all $a \in A$ (cf. [10, II]).

References

[1] J. Bendat and S. Sherman: Monotone and convex operator functions, Trans. Amer. Math. Soc., **79**, 58–71 (1955).
[2] C. Davis: A Schwarz inequality for convex operator functions, Proc. Amer. Math. Soc., **8**, 42–44 (1957).
[3] J. Dixmier: Les Algèbres d'Opérateurs dans l'Espace Hilbertien, Gauthier-Villars, Paris (1957).
[4] I. Kaplansky: A theorem on rings of operators, Pacific J. Math., **1**, 227–232 (1951).
[5] A. I. Khinchin: Mathematical Foundations of Information Theory, Dover, New York (1957).
[6] M. Nakamura, M. Takesaki, and H. Umegaki: A remak on the expectations of operator algebras, Kōdai Math. Sem. Rep., **12**, 82–90 (1960).
[7] M. Nakamura and T. Turumaru: Expectations in an operator algebra, Tôhoku Math. J., **6**, 182–188 (1954).
[8] J. von Neumann: Mathematical Foundations of Quantum Mechanics, Princeton Univ. Press, Princeton (1955).
[9] I. E. Segal: A note on the concept of entropy, J. Math. Mech., **9**, 623–629 (1960).
[10] H. Umegaki: Conditional expectation in an operator algebra, I, Tôhoku Math. J., **6**, 177–181 (1954); II, **8**, 86–100 (1956); and III, Kōdai Math. Sem. Rep., **11**, 51–64 (1959).