AROUND BISTOCHASTIC MATRICES*

Marie Choda

Osaka Kyoiku University, Asahigaoka, Kashiwara 582-8582, Japan

(Received January 27, 2015)

We review some results about roles of bistochastic matrices from a point of view of entropy and the notion of the relative position between subalgebras of matrix algebras.

DOI:10.5506/APhysPolB.46.1883
PACS numbers: 47.27.eb, 47.27.ed

1. Introduction

In this paper, we review the papers [1–3], where bistochastic matrices play some important roles from the point of view of characterizations of positive maps related to the notion of "entropy" and relations between subalgebras of matrix algebras. So proofs are often omitted and I refer the interested reader to the introductory literature on the subject [1–3].

In order to simplify our discussion, throughout this paper, we restrict our target to the algebra $M_n(\mathbb{C})$ of $n \times n$ complex matrices. We denote the (i,j)-component of $x \in M_n(\mathbb{C})$ by x_{ij}, and the identity matrix by 1 simply. By Tr_n (often simply by Tr), we mean the standard trace of $M_n(\mathbb{C})$ such that $\mathrm{Tr}(p) = 1$ for every minimal projection p. The canonical tracial state Tr/n is denoted as τ simply.

In Section 2, we review the paper [3] *Around Shannon's Interpretation for Entropy-preserving Stochastic Averages*. The main topic in [3] is characterizations for entropy preserving positive unital trace-preserving maps via bistochastic matrices. A motivation in [3] is to give a generalized version of Shannon's interpretation for entropy-preserving stochastic averages of probability vectors in the framework of von Neumann entropy for states on $M_n(\mathbb{C})$. Shannon states in [4, p. 395] the following: If we perform any "averaging" operation on the $p = \{p_i\}_{i=1,\dots,n}$ of the form $p'_i = \sum_j a_{ij}p_j$, (where $a_{ij} \geq 0$, $\sum_i a_{ij} = \sum_j a_{ij} = 1$), the entropy H increases (except in

---

* Presented at the Conference "Random Matrix Theory: Foundations and Applications", Kraków, Poland, July 1–6, 2014.

the special case where this transformation amounts to no more than a permutation of the p_i with H, of course, remaining the same). By replacing a probability vector (resp. a bistochastic matrix) to a state ρ of $M_n(\mathbb{C})$ (resp. a unital positive Tr-preserving map Φ on $M_n(\mathbb{C})$), we showed, among the others, that the action of Φ on ρ preserves the von Neumann entropy if and only if Φ behaves just like an automorphism for the state ρ. Furthermore, we discuss various kinds of entropy in connection to this interpretation.

In Section 3, we describe some results of paper [1]: *Relative Entropy for Maximal Abelian Subalgebras of Matrices and the Entropy of Unistochastic Matrices*. The aim in [1] is to discuss the notion of the entropy for unistochastic matrices from the operator algebraic point of view. There are several notions on some relative position between two subalgebras of operator algebras. As one of such notions, Popa [5] introduced the notion of *mutually orthogonal subalgebras*. By the terminology *complementarity*, the same notion is investigated in the theory of quantum systems (see [6] for example).

In Section 4, we review some results about mutually orthogonal subalgebras. The most primary interest would be the case of two subalgebras of some full matrix algebra, both of which are either maximal Abelian or isomorphic to also some full matrix algebra. In such the cases, two subalgebras are connected by a unitary. We study the case when the subalgebras A and B in question are maximal Abelian subalgebras and in the next section, we study the case when the subalgebras A and B are isomorphic to some $M_n(\mathbb{C})$, *i.e.*, subfactors. First, we denote the results on mutually orthogonal maximal Abelian subalgebras in papers [1] and [2] *Von Neumann Entropy and Relative Position Between Subalgebras*, and then the results on mutually orthogonal subfactors in paper [2] by introducing some density matrix arising from the pair $\{A, B\}$. We show that the von Neumann entropy of the density matrix gives a characterization of the mutual orthogonality by using the notion of operational partition of unity.

Here, we summarize notations, terminologies and basic facts.

Entropy function $\boldsymbol{\eta}$. The *entropy function* η is defined on $[0,1]$ by

$$\eta(t) = -t\log t\,, \qquad 0 < t \leq 1 \qquad \text{and} \qquad \eta(0) = 0\,. \tag{1.1}$$

The η is *strictly operator-concave*, *i.e.* for a k-tuple of real numbers $\{t_i\}_{i=1}^k$ such that $t_i > 0$, $\sum_{i=1}^k t_i = 1$ and matrices $\{x_i\}_{i=1}^k$ with eigenvalues in [0,1], it holds in general that

$$\sum_{i=1}^k t_i\eta(x_i) \leq \eta\left(\sum_{i=1}^k t_i x_i\right) \tag{1.2}$$

and equality implies that $x_i = x_j$ for all i, j. (See, for example, [7, B], [6, 8].)

Shannon entropy. Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a probability vector in $\mathbb{R}^n$. The *Shannon entropy* $H(\lambda)$ for λ is given as

$$H(\lambda) = \eta(\lambda_1) + \cdots + \eta(\lambda_n)\,. \tag{1.3}$$

It holds always that $H(\lambda) \leq \log n$ and $H(\lambda) \leq H(\lambda b)$ ([4], *cf.* [6]) for a *bistochastic* matrix $b = [b_{ij}]$ (*i.e.* $b_{ij} \geq 0$, $\sum_i b_{ij} = \sum_j b_{ij} = 1$ for all $i, j = 1, \ldots, n$). A bistochastic matrix is also called a *doubly stochastic* matrix (see, for example, [6]).

Density and von Neumann entropy. Every positive linear functional ϕ on $M_n(\mathbb{C})$ is of the form

$$\phi(x) = \mathrm{Tr}(D_\phi x)\,, \qquad x \in M_n(\mathbb{C}) \tag{1.4}$$

for a unique positive element D_ϕ in $M_n(\mathbb{C})$. D_ϕ is called the *density matrix*. If ρ is a state, then $\mathrm{Tr}(D_\rho) = 1$.

Using the eigenvalue list $\{\lambda_1, \ldots, \lambda_n\}$ of D_ϕ, the *von Neumann entropy* $S(\phi)$ for a positive linear functional ϕ and the *von Neumann entropy* $S(D_\phi)$ for D_ϕ are defined by

$$S(\phi) = S(D_\phi) = \sum_{i=1}^{n} \eta(\lambda_i)\,. \tag{1.5}$$

Hence, if ρ is a state, then the vector $\lambda = (\lambda_1, \ldots, \lambda_n)$ of the eigenvalues of D_ρ is a probability vector and $S(\rho) = S(D_\rho) = H(\lambda)$.

The pair $\{\boldsymbol{\rho}, \boldsymbol{\Phi}\}$ of a state and a positive map. Let ρ be a state of $M_n(\mathbb{C})$, and let D_ρ be the density matrix of ρ. Let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a positive unital Tr-preserving map. Then, $\Phi(D_\rho)$ is a positive matrix and $\mathrm{Tr}(\Phi(D_\rho)) = 1$. The most basic examples of positive unital Tr-preserving maps are automorphisms and conditional expectations, that is, there exists always a unique positive linear map E_A of $M_n(\mathbb{C})$ onto a subalgebra A of $M_n(\mathbb{C})$ such that $\tau(xa) = \tau(E_A(x)a)$ for all $x \in M_n(\mathbb{C}), a \in A$. It is called the *conditional expectation* and satisfies that $E_A(axb) = aE_A(x)b$, for all $x \in M,\ a, b \in A$.

In order to see the state whose density matrix is $\Phi(D_\rho)$, we need the Hilbert–Schmidt inner product of $M_n(\mathbb{C})$. The inner product is given by

$$\langle x, y\rangle = \mathrm{Tr}\left(y^* x\right)\,, \qquad x, y \in M_n(\mathbb{C}) \tag{1.6}$$

and the adjoint map $\Phi^* : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ of Φ is given by

$$\mathrm{Tr}\left(y\Phi^*(x)\right) = \mathrm{Tr}\left(\Phi(y)x\right)\,, \qquad x, y \in M_n(\mathbb{C})\,. \tag{1.7}$$

Since Φ is positive and Tr-preserving, it implies that Φ^* is positive and unital so that $\rho \circ \Phi^*$ is a state, whose density matrix is $\Phi(D_\rho)$

$$\rho \circ \Phi^*(x) = \mathrm{Tr}\,(D_\rho \Phi^*(x)) = \mathrm{Tr}(\Phi(D_\rho)x)\,, \qquad x \in M_n(\mathbb{C})\,. \tag{1.8}$$

Let

$$\lambda = (\lambda_1, \ldots, \lambda_n) \qquad \text{and} \qquad \mu = (\mu_1, \ldots, \mu_n)\,,$$

be the probability vectors of the eigenvalues of D_ρ and $\Phi(D_\rho)$ respectively. Here, we arrange them always in a decreasing order

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \qquad \text{and} \qquad \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n\,.$$

Let $\{e_1, \ldots, e_n\}$ (resp. $\{p_1, \ldots, p_n\}$) be the mutually orthogonal minimal projections inducing the following decomposition of D_ρ (resp. $\Phi(D_\rho)$)

$$D_\rho = \sum_{i=1}^n \lambda_i e_i \qquad \text{and} \qquad \Phi(D_\rho) = \sum_{j=1}^n \mu_j p_j\,. \tag{1.9}$$

We denote by $u_{(\rho,\Phi)}$ the unitary such that

$$u_{(\rho,\Phi)} e_i {u_{(\rho,\Phi)}}^* = p_i \qquad \text{for each} \qquad i = 1, \ldots, n\,. \tag{1.10}$$

Also we denote by A (resp. B) the maximal Abelian subalgebra of $M_n(\mathbb{C})$ generated by $\{e_1, \ldots, e_n\}$ (resp. $\{p_1, \ldots, p_n\}$).

2. Around Shannon's interpretation

2.1. The bistochastic matrix $b_\rho(\Phi)$ for the pair $\{\rho, \Phi\}$

For a state ρ of $M_n(\mathbb{C})$ and a unital positive Tr-preserving map Φ on $M_n(\mathbb{C})$, we define the matrix $b_\rho(\Phi)$ by

$$b_\rho(\Phi)_{ij} = \mathrm{Tr}(\Phi(e_i)p_j)\,, \qquad 1 \leq i \leq n\,, \qquad 1 \leq j \leq n\,. \tag{2.1}$$

Our discussions do not depend on this kind of matrix representations as follows, we denote them simply by $b_\rho(\Phi)$:

Assume that $\lambda_i > \lambda_{i+1}$ and $\mu_i > \mu_{i+1}$ for all $i = 1, \ldots, n-1$. Then, the value $b_\rho(\Phi)_{ij}$ is uniquely determined for all i, j because the spectral projections $\{e_i\}_i$ and $\{p_j\}_j$ are uniquely determined. In the other case, the value $b_\rho(\Phi)_{ij}$ is not always uniquely determined. For example, if it happened that $D = \sum_i \lambda_i f_i$ and $\Phi(D) = \sum_j \mu_j q_i$ for some projections $\{f_i\}_{i=1}^n$ and $\{q_j\}_{j=1}^n$ different from $\{e_i\}_{i=1}^n$ and $\{p_j\}_{j=1}^n$, then the matrix $b_\rho(\Phi)^{e,p}$ with the (i,j)-coefficient $\mathrm{Tr}(\Phi(e_i)p_j)$ may be different from the matrix $b_\rho(\Phi)^{f,q}$ with the (i,j)-coefficient $\mathrm{Tr}(\Phi(f_i)q_j)$. But the difference is covered by the permutation matrices $[\sigma]$ and $[\pi]$ via the permutations σ and π such that $f_i = e_{\sigma(i)}$ and $q_i = p_{\sigma(i)}$ for all i and j: $b_\rho(\Phi)^{f,q} = [\sigma] b_\rho(\Phi)^{e,p}[\pi]$. Then, $b_\rho(\Phi)$ is a bistochastic matrix, *i.e.*,

$$b_{ij} \geq 0 \qquad \text{and} \qquad \sum_i b_{ij} = \sum_j b_{ij} = 1 \quad \text{for all } i, j$$

and $b_\rho(\Phi)$ transposes the vector λ to the vector μ, *i.e.* $\lambda b_\rho(\Phi) = \mu$.

By using this bistochastic matrix $b_\rho(\Phi)$, in [3] we had the following:

Theorem 2.1. *Let ρ be a state on $M_n(\mathbb{C})$ and let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a unital positive Tr-preserving map. Then, the following are equivalent:*

(i) $S(\rho \circ \Phi^*) = S(\rho)$, *i.e.* $S(\Phi(D_\rho)) = S(D_\rho)$.

(ii) $\lambda = \mu b_\rho(\Phi)^T$, *where x^T denotes the transpose of the matrix x.*

(iii) $\lambda_i = \mu_i$ *for all* $i = 1, \ldots, n$.

(iv) $\Phi(D_\rho) = u D_\rho u*$ *for some unitary* $u \in M_n(\mathbb{C})$.

(v) $\Phi^* \Phi(D_\rho) = D_\rho$.

Remark 2.2. Under the assumption that Φ is 2-positive, the corresponding relation to *(i)* $\Leftrightarrow$ *(v)* is obtained for the discussion on relative entropy in [9, Theorem 7.1].

In our case, Φ is not necessarily 2-positive.

Example 2.3. Now, we consider the transpose mapping $\Phi : x \to x^T$ on $M_n(\mathbb{C})$. It is a typical example of a unital Tr-preserving positive but not 2-positive map, and Φ satisfies the all conditions in the theorem for every state ρ.

In fact, the Φ is a symmetry as follows:

$$\begin{aligned} \langle \Phi^*(x), y \rangle &= \langle x, \Phi(y) \rangle = \mathrm{Tr}\left(\Phi(y)^* x\right) = \mathrm{Tr}\left(\left(y^T\right)^* x\right) \\ &= \sum_{i,j=1}^n \overline{y_{i,j}} x_{j,i} = \sum_{i,j=1}^n \overline{y_{j,i}} x_{i,j} = \mathrm{Tr}\left(y^* x^T\right) = \mathrm{Tr}(y^* \Phi(x)) \\ &= \langle \Phi(x), y \rangle \qquad \text{for all} \qquad x = (x_{ij}), \qquad y = (y_{ij}). \end{aligned}$$

Hence, $\Phi^* \Phi$ is the identity map on $M_n(\mathbb{C})$.

Remark 2.4. If the state ρ is the normalized trace Tr/n, then $D_\rho = \mathrm{I}_n/n$ so that the statements *(i)–(v)* are all trivial for every Φ.

2.2. Relations among various entropies

Definition 2.5. We modify the notion of weighted entropy for a bistochastic matrix defined in [10] to our bistochastic matrix $b_\rho(\Phi)$. We let

$$H^\lambda(b_\rho(\Phi))=\sum_{i=1}^n \lambda_i \sum_{j=1}^n \eta\left(b_\rho(\Phi)_{ij}\right) \quad \text{and} \quad H_\mu(b_\rho(\Phi))=\sum_{j=1}^n \mu_j \sum_{i=1}^n \eta\left(b_\rho(\Phi)_{ij}\right) .$$

This is well-defined, *i.e.* the values $H^\lambda(b_\rho(\Phi))$ and $H_\mu(b_\rho(\Phi))$ depend only on the pair $\{\rho, \Phi\}$. In fact, assume that $D_\rho = \sum_i \lambda_i f_i$ and $\Phi(D_\rho) = \sum_j \mu_j q_i$ for minimal projections $\{f_i\}_{i=1}^n$ and $\{q_j\}_{j=1}^n$ which are not always the same as $\{e_i\}_{i=1}^n$ and $\{p_j\}_{j=1}^n$. Then, there are permutations σ and π of $\{1, \ldots, n\}$ such that $f_i = e_{\sigma(i)}$ and $q_i = p_{\pi(i)}$ for all i. If $\lambda_i = \lambda_{\sigma(i)}$ for all i, then

$$\begin{aligned}
\sum_{i=1}^n \lambda_i \sum_{j=1}^n \eta\left(\operatorname{Tr}\left(\Phi(f_i) q_j\right)\right) &= \sum_{i=1}^n \lambda_i \sum_{j=1}^n \eta\left(\operatorname{Tr}\left(\Phi\left(e_{\sigma(i)}\right) p_{\pi(j)}\right)\right) \\
&= \sum_{i=1}^n \lambda_{\sigma(i)} \sum_{j=1}^n \eta\left(\operatorname{Tr}\left(\Phi\left(e_{\sigma(i)}\right) p_{\pi(j)}\right)\right) \\
&= \sum_{i=1}^n \lambda_i \sum_{j=1}^n \eta\left(\operatorname{Tr}\left(\Phi(e_i) p_j\right)\right) .
\end{aligned}$$

Hence, the value $H^\lambda(b_\rho(\Phi))$ does not depend on the choice of minimal projections. Similarly, it holds for $H_\mu(b_\rho(\Phi))$ since $\mu_j = \mu_{\pi(j)}$ for all j.

Definition 2.6. Since Φ is a positive unital Tr-preserving map, $\Phi(e_i)$ and $\Phi^*(p_j)$ are density matrices for all i, j. We put $S_\rho(\Phi)$ and $S^\rho(\Phi^*)$ as the following:

$$S_\rho(\Phi) = \sum_{i=1}^n \lambda_i S(\Phi(e_i)) \qquad \text{and} \qquad S^\rho(\Phi^*) = \sum_{j=1}^n \mu_j S(\Phi^*(p_j)) .$$

Similarly to the case of $H^\lambda(b_\rho(\Phi))$ and $H_\mu(b_\rho(\Phi))$, the values $S_\rho(\Phi)$ and $S^\rho(\Phi^*)$ are uniquely determined by the pair $\{\rho, \Phi\}$.

Proposition 2.7. *For the conditional expectation E_B (resp. E_A) onto B (resp. A), the following holds:*

1. $E_B(\Phi(e_i)) = \sum_{j=1}^n b_\rho(\Phi)_{ij} p_j$, $E_A(\Phi^*(p_j)) = \sum_{i=1}^n b_\rho(\Phi)_{ij} e_i$,
2. $H^\lambda(b_\rho(\Phi)) = \sum_i \lambda_i S(E_B(\Phi(e_i)))$, $H_\mu(b_\rho(\Phi)) = \sum_j \mu_j S(E_A(\Phi^*(p_j)))$.

For a probability vector $\lambda = (\lambda_1, \ldots, \lambda_n)$, we set

$$J_\lambda = \{k; \lambda_k \neq 0\}\,. \tag{2.2}$$

Theorem 2.8. *Let ρ be a state of $M_n(\mathbb{C})$, and let Φ be a unital positive Tr-preserving map on $M_n(\mathbb{C})$. Then, the following statements hold:*

1. $$\begin{array}{c} S_\rho(\Phi) \leq H^\lambda\left(b_\rho(\Phi)\right) \leq S\left(\rho \circ \Phi^*\right) = S\left(\Phi\left(D_\rho\right)\right) \leq S(\rho) + S_\rho(\Phi) \\ \vee \ \| \\ S(\rho) = S\left(D_\rho\right). \end{array}$$

2. $S_\rho(\Phi) = H^\lambda(b_\rho(\Phi))$ *if and only if* $\Phi(e_i) \in B$ *for all* $i \in J_\lambda$*:*

$$\Phi(e_i) = \sum_{j=1}^{n} b_\rho(\Phi)_{ij} p_j\,, \quad \text{for all} \quad i \in J_\lambda\,.$$

3. $H^\lambda(b_\rho(\Phi)) = S(\rho \circ \Phi^*)$ *if and only if* $b_\rho(\Phi)_{ij} = \mu_j$ *for all* $i \in J_\lambda$ *and* j*:*

$$\Phi(D_\rho) = \sum_{j=1}^{n} b_\rho(\Phi)_{ij} p_j\,.$$

4. $S_\rho(\Phi) = S(\rho \circ \Phi^*)$ *if and only if* $\Phi(D_\rho) = \Phi(e_i)$ *for every* $i \in J_\lambda$*.*
5. $S(\rho \circ \Phi^*) = S(\rho) + S_\rho(\Phi)$ *if and only if the* ρ *is a pure state.*

Corollary 2.9. *Assume that all eigenvalues of D_ρ are non-zero and that $H^\lambda(b_\rho(\Phi)) = S(\rho \circ \Phi^*)$. Then the state $\rho \circ \Phi^*$ is the canonical tracial state, and so that $H^\lambda(b_\rho(\Phi)) = S(\rho \circ \Phi^*) = log\ n$.*

Remark 2.10. (**Unistochastic matrix and Hadamard matrix.**) A bistochastic matrix b is said to be *unistochastic* if it is induced from a unitary matrix u by that $b_{ij} = |u_{ij}|^2$ for all i, j. A unitary matrix u is called a *Hadamard matrix* if $|u_{ij}| = 1/\sqrt{n}$ for all i, j. The above shows that if all eigenvalues of D_ρ are non-zero and if $H^\lambda(b_\rho(\Phi)) = S(\Phi(D_\rho))$ then $b_\rho(\Phi)$ is a unistochastic matrix induced from a Hadamard matrix.

Example 2.11.

(1) Assume that ρ is a pure state. Then, it is clear that

$$S(\rho) = 0 \qquad \text{and} \qquad S_\rho(\Phi) = H^\lambda(b_\rho(\Phi)) = S(\rho \circ \Phi^*)$$

for every positive unital Tr-preserving map Φ.

Furthermore, for any given value s with $0 \leq s \leq \log n$, there exists a positive unital Tr-preserving map Φ such that $S_\rho(\Phi) = s$.

(2) If Φ is a *-isomorphism, then for each state ρ,

$$S_\rho(\Phi) = H^\lambda(b_\rho(\Phi)) = 0 \quad \text{and} \quad S(\rho \circ \Phi^*) = S(\rho)\,.$$

(3) If Φ is a map $M_n(\mathbb{C}) \to \mathbb{C}1_M$, then for every state ρ

$$S_\rho(\Phi) = H^\lambda(b_\rho(\Phi)) = S(\rho \circ \Phi^*) = \log n\,.$$

(4) A typical counter example of Φ for the statement (4) in the above theorem is the transpose mapping $\Phi(x) = x^T$, where ρ is not a pure state.

More general examples are given as follows: Let $D = \sum_i \lambda_i e_i$ be a given density matrix. Let $\{p_j\}_{j=1}^n$ be mutually orthogonal minimal projections. Then, we have a family of partially isometries $\{v_{ij}\}_{ij}$ such that $v_{ij}^* v_{ij} = e_j$ and $v_{ij} v_{ij}^* = p_j$. Let $a = [a_{ij}]$ be a bistochastic matrix and let

$$\Phi(x) = \sum_{i,j} a_{ij} v_{ij} x v_{ij}^*\,, \qquad x \in M_n(\mathbb{C})\,.$$

Then, Φ is a unital positive Tr-preserving map and

$$\Phi(D) = \sum_i (\sum_j a_{ij}\lambda_j) p_i \qquad \text{and} \qquad \Phi(e_i) = \sum_j a_{ji} p_j \quad \text{for all} \quad i\,.$$

Hence, $\Phi(e_i) \in B$ for all i, that is the condition (2).

Also we can choose bistochastic matrices $a = [a_{ij}]$, one of which induces Φ satisfying the condition (4) and the other of which induces Φ not satisfying the condition (4).

The following show us that the bistochastic matrix $b_\rho(\Phi)$ plays an important role:

Theorem 2.11. *Let ρ be a state of $M_n(\mathbb{C})$, and let Φ be a unital positive Tr-preserving map on $M_n(\mathbb{C})$. Then, the following conditions are equivalent:*

(0) $H^\lambda(b_\rho(\Phi)) = 0$,

(1) for each $i \in J_\lambda$, there exists a unique $j(i)$ such that

$$\lambda_i = \mu_{j(i)} \qquad \textit{and} \qquad \Phi(e_i) = p_{j(i)}\,,$$

(2) $S(\rho) = S(\rho \circ \Phi^*) = S(\Phi(D_\rho))$,

(3) there exists a unitary u such that $\Phi(D_\rho) = u D_\rho u^*$,

(4) $\Phi^*\Phi(D_\rho) = D_\rho$.

3. Relative entropy and the entropy of unistochastic matrices

In this section, we restrict our attention to the maximal Abelian subalgebras (abbreviated as MASA's) of $M_n(\mathbb{C})$. The most typical example of a MASA is the algebra $D_n(\mathbb{C})$ of all diagonal matrices.

If A and B are two MASAs of $M_n(\mathbb{C})$, then there exists a unitary matrix u with $B = uAu^*$, which we denote by $u(A, B)$. Each unitary matrix u induces a unistochastic matrix $b(u)$, which is a typical example of a bistochastic matrix.

The aim of this section is to discuss the notion of the entropy for unistochastic matrices from the operator algebraic point of view.

3.1. Relative entropy of Connes–Størmer

First, we review about the formula of the relative entropy defined by Connes and Størmer in [2] (*cf.* [11]).

Let S be the set of all finite families (x_i) of positive elements in $M_n(\mathbb{C})$ with $1 = \sum_i x_i$. Let A and B be two subalgebras of $M_n(\mathbb{C})$. The relative entropy $H(A|B)$ is

$$H(A \mid B) = \sup_{(x_i)\in S} \sum_i \left(\tau\eta E_B(x_i) - \tau\eta E_A(x_i)\right) .$$

Let ϕ be a state on $M_n(\mathbb{C})$ and let Φ be the all finite families (ϕ_i) of positive linear functionals on $M_n(\mathbb{C})$ with $\phi = \sum_i \phi_i$. The relative entropy $H_\phi(A|B)$ of A and B related to ϕ is defined by

$$H_\phi(A|B) = \sup_{(\phi_i)\in\Phi} \sum_i (S(\phi_i\mid_A, \phi\mid_A) - S(\phi_i\mid_B, \phi\mid_B)) .$$

Here, $S(\psi|\varphi)$ is the relative entropy for two positive linear functionals ψ and φ such that $\psi \le \lambda\varphi$ for some $\lambda > 0$ given as

$$S(\psi, \varphi) = \mathrm{Tr}(D_\psi(\log D_\psi - \log D_\varphi)) .$$

3.2. Conditional relative entropy

As a replacement of $H(A|B)$ (resp. $H_\phi(A|B)$), we defined $h(A|B)$ (resp. $h_\phi(A|B)$) in [1] as follows:

The *conditional relative entropy* $h(A \mid B)$ conditioned by A is

$$h(A \mid B) = \sup_{(x_i)\in S} \sum_i (\tau\eta E_B(E_A(x_i)) - \tau\eta E_A(x_i)) .$$

Let $S(A) \subset S$ be the set of all finite families (x_i) of positive elements in A with $1 = \sum_i x_i$. Then, it is clear that

$$h(A \mid B) = \sup_{(x_i)\in S(A)} \sum_i (\tau\eta E_B(x_i) - \tau\eta(x_i))\,.$$

Let $S'(A) \subset S(A)$ be the set of all finite families (x_i) with each x_i a scalar multiple of a projection in A. Then, by the same proof as in [12], we have

$$h(A \mid B) = \sup_{(x_i)\in S'(A)} \sum_i (\tau\eta E_B(x_i) - \tau\eta(x_i))\,.$$

Hence, we only need to consider the families consisting of scalar multiples of orthogonal minimal projections.

The *conditional relative entropy* of A and B with respect to a state ϕ conditioned by A is defined in [1] as

$$h_\phi(A|B) = \sup_{(\phi_i)\in\Phi} \sum_i (S(\phi_i\mid_A, \phi\mid_A) - S((\phi_i\circ E_A)\mid_B, (\phi\circ E_A)\mid_B))\,.$$

Assume that $D_\phi \in A$. Let $\Phi(A) \subset \Phi$ be the set of all finite families $(\phi_i)_i$ such that $(D_{\phi_i})_i \subset A$. Then, by the fact that $D_{\phi_i\circ E_A} = E_A(Q_i)$, we have

$$h_\phi(A|B) = \sup_{(\phi_i)\in\Phi(A)} \sum_i (S(\phi_i\mid_A, \phi\mid_A) - S(\phi_i\mid_B, \phi\mid_B))\,.$$

Here, we just remark the following facts:

(1) If ϕ is the normalized trace τ, then by [11, Theorem 2.3.1(x)]

$$h_\tau(A|B) = h(A \mid B)\,.$$

(2) In general, $0 \le h_\phi(A|B) \le H_\phi(A|B)$. If A and B are contained in an Abelian algebra, then $h_\phi(A \mid B)$ coincides with the conditional entropy in the ergodic theory by a similar proof as in [11, p. 158].

3.2.1. Relation to Schur's Lemma

What is the meaning of the entropy H(b(u)) from the theory of operator algebras?

We would like to consider that it should be an invariant related to inner automorphisms. For example, we just remember Schur's Lemma. Let

$$\lambda = (\lambda_1, \ldots, \lambda_n)\,, \quad \lambda_i \in \mathbb{R} \qquad \text{and} \qquad d = (d_1, \ldots, d_n)\,, \quad d_i \in \mathbb{R}\,.$$

Then, Schur's Lemma says that if λ is the eigenvalue sequence of a self-adjoint matrix $a \in M_n(\mathbb{C})$, and if d is the diagonal sequence of a, then λ majorizes d, that is, for each k with $1 \le k \le n$,

$$\sum \text{ the } k \text{ largest } \lambda's \ \ge \sum \ \text{ the } k \text{ largest } d's\,,$$

and

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} d_i\,.$$

These were shown by using the fact that there exists a unitary matrix $u = (u(i,j))_{ij}$ such that

$$d_i = \sum_j |u(i,j)|^2 \lambda_j\,, \qquad j = 1, \ldots, n\,.$$

Thus, a unitary u and the unistochastic matrix $b(u)$ defined by u appeared in the step to get a diagonal matrix from d with $d = d^*$.

3.3. Relations between $h_\phi(A|uAu^*)$ and $H(b(u))$

Let ϕ be a positive linear functional on $M_n(\mathbb{C})$. We set the eigenvalues of the density D_ϕ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Let us decompose D_ϕ into the form of $D_\phi = \sum_{i=1}^{n} \lambda_i e_i$, where $\{e_1, \ldots, e_n\}$ are mutually orthogonal minimal projections in $M_n(\mathbb{C})$, which we fix. By keeping in mind that each MASA is isomorphic to the algebra $D_n(\mathbb{C})$ of diagonal matrices, we describe our characterizations in [1] by the following simple form:

Theorem 3.1. *Let ϕ be a state of $M_n(\mathbb{C})$, and let $u \in M_n(\mathbb{C})$ be a unitary. Then, we have that*

$$h_\phi(D_n(\mathbb{C}) \mid uD_n(\mathbb{C})u^*) = H_\lambda(b(u)^*) + S(\phi|_{D_n(\mathbb{C})}) - S(\phi|_{uD_n(\mathbb{C})u^*})\,. \quad (3.1)$$

In the special case, where $\phi = \tau$,

$$h(D_n(\mathbb{C}) \mid uD_n(\mathbb{C})u^*) = H(b(u)) = max_\phi h_{\phi_v}(D_n(\mathbb{C}) \mid uD_n(\mathbb{C})u^*)\,, \quad (3.2)$$

where the maximum is taken over all states ϕ of $M_n(\mathbb{C})$ and ϕ_v is the state given by the inner perturbation of ϕ by v : $\phi_v(x) = \phi(vxv^*)$.

4. Mutually orthogonal subalgebras

Popa [5] defined that two subalgebras A and B of $M_n(\mathbb{C})$ are *mutually orthogonal* (sometimes denoted as $A \perp B$) if $\tau(ab) = 0$ for $a \in A, b \in B$ with $\tau(a) = \tau(b) = 0$.

4.1. Maximal Abelian subalgebras via crossed product decomposition

First, remark that $M_n(\mathbb{C})$ is represented as the crossed product of a maximal Abelian subalgebra A by the group $\mathbb{Z}_n$ with respect to α

$$M_n(\mathbb{C}) = A \times_\alpha \mathbb{Z}_n\,,$$

where α is the automorphism of A with $\alpha(e_i) = e_{i+1}$, (mod n) for a mutually orthogonal minimal projections $\{e_1, e_2, \ldots, e_n\}$ of A. Let $\{e_{ij}; i, j = 1, 2, \ldots, n\}$ be a system of a matrix units of $M_n(\mathbb{C})$ with $e_{ii} = e_i$ for all i. Then, the unitary $v = \sum_{i=1}^n e_{i\ i-1}$ satisfies that $\alpha(a) = vav^*$ for all $a \in A$. Each $x \in M_n(\mathbb{C})$ is uniquely written as

$$x = \sum_{i=0}^{n-1} x_i v^i\,, \qquad x_i \in A\,. \tag{4.1}$$

The conditional expectation E_A of M onto A is given by $E_A(x) = x_0$, and it holds that $x_i = E_A(x{v^i}^*)$ for all $i = 0, \ldots, n-1$ and $x \in M_n(\mathbb{C})$.

We defined in [2] the entropy $S(u)$ of a unitary $u \in M_n(\mathbb{C})$ by using the decomposition of $u = \sum_{j=0}^{n-1} u_j v^j$ as the following:

$$S(u) = \sum_{j=0}^{n-1} \tau_A \eta\left(u_j u_j^*\right) = \frac{1}{n}\sum_{j=0}^{n-1} \mathrm{Tr}\,\eta\left(u_j u_j^*\right) = \frac{1}{n}\sum_{j=0}^{n-1} S\left(u_j u_j^*\right)\,.$$

Remark that the family $\{u_i u_i^*; i = 0, \ldots, n-1\}$ is a finite partition of unity in A and so the entropy $S(u)$ is nothing else but the average of the von Neumann entropy $\{S(u_j u_j^*) : j = 0, 1, \ldots, n-1\}$.

Now, we can give characterization for the notion of mutually orthogonality by entropy as the following:

Theorem 4.1.

(I) *Let A be a maximal Abelian subalgebras of $M_n(\mathbb{C})$, and let u be a unitary. Then the following characterization in [2] holds:*

$$A \perp uAu^* \iff S(u) = log\ n = max\{S(w) \mid w \in M\,, unitary\}\,.$$

(II) *Another characterization in [1] is the following: Let $\{A_0, B_0\}$ be a pair of maximal Abelian subalgebras of $M_n(\mathbb{C})$. Then,*

$$A_0 \perp B_0 \iff h(A_0 \mid B_0) = log\ n = max\ h(A \mid B)\,,$$

where the maximum is taken over the set of pairs $\{A, B\}$ of maximal Abelian subalgebras of $M_n(\mathbb{C})$.

Remark 4.2. Jones and Sunder [13] characterized that a unitary u is complex Hadamard matrix if and only if $D_n(\mathbb{C}) \perp uD_n(\mathbb{C})u^*$.

Here, our characterization is that u is a complex Hadamard matrix if and only if $S(u) = \log n$.

4.2. Case of subfactors

In the case of mutual orthogonality for subfactors, we need to use another kind of entropy to get a characterization. Now, let $M = M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$, and let $\tau_M = \mathrm{Tr}_n/n \otimes \mathrm{Tr}_m/m$. Consider the subalgebra $N = M_n(\mathbb{C}) \otimes 1 \subset M$. Then, the conditional expectation E_N on N has the following form:

$$E_N(x \otimes y) = x \otimes \frac{\mathrm{Tr}_m}{m}(y)1_{M_m(\mathbb{C})}\,, \qquad x \in M_n(\mathbb{C})\,, \qquad y \in M_m(\mathbb{C})\,.$$

As an easy consequence, for each unitary $u \in M$, we have the following:

$$N \perp uNu^* \Longleftrightarrow E_N(u^*(a \otimes 1_L)u) = \frac{\mathrm{Tr}(a)}{n}1_M\,, \qquad \forall a \in M_n(\mathbb{C})\,.$$

Now, let $\{e_{ij}; i,j = 1,\ldots,n\}$ be a system of matrix units of $M_n(\mathbb{C})$. Each x in $M = M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ is written in the unique form

$$x = \sum_{i,j=1}^{n} e_{ij} \otimes x_{ij}\,, \qquad x_{ij} \in M_m(\mathbb{C})\,.$$

We apply the notion of a finite operational partition X of unity and the density matrix $\rho_\phi[X]$ introduced by Lindblad [12]:

A *finite operational partition of unity* is a subset $X = \{x_1,\ldots,x_k\}$ of $M_n(\mathbb{C})$ such that $\sum_i^k x_i^* x_i = 1$. The k is called the size of X. To a finite operational partition X of unity of size k, we associate a $k \times k$ density matrix $\rho[X]$ given by $\rho[X]_{ij} = \tau(x_j^* x_i) = \mathrm{Tr}/n(x_j^* x_i)$, $(i,j = 1,\ldots,k)$.

4.2.1. Finite operational partition induced by a unitary u

Let $M = M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ so that $\tau_M = \mathrm{Tr}_n/n \otimes \mathrm{Tr}_m/m$. Let $u = \sum_{i,j} e_{ij} \otimes u_{ij}$, $(u_{ij} \in M_m(\mathbb{C}))$ be the decomposition of a unitary $u \in M$, where $\{e_{ij}\}_{i,j}$ is a set of matrix units of $M_n(\mathbb{C})$. We set

$$U = \left\{ \frac{1}{\sqrt{n}} u_{ij} \;;\; i,j = 1,\ldots,n \right\}.$$

It is clear that U is a finite operational partition of unity of size n^2. We call this set U the *finite operational partition of unity induced by u*.

4.2.2. Mutually orthogonal subfactors

First, we remark that every subfactor of $M_k(\mathbb{C})$ is isomorphic to $M_n(\mathbb{C})$ for some n, and let $N \subset M_k(\mathbb{C})$ be isomorphic to $M_n(\mathbb{C})$. Then, $k = mn$ for some m. We can assume that $M_k(\mathbb{C}) = M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ and $N = M_n(\mathbb{C}) \otimes \mathbb{C}1$. Let $u \in M_k(\mathbb{C})$ be a unitary. Then, by considering the von Neumann entropy $S(\rho[U])$, we had the following characterization in [2]

$$N \perp uNu^* \Longleftrightarrow S(\rho[U]) = 2\log n = \log \dim N\,,$$

where U is the finite operational partition of unity induced by u.

REFERENCES

[1] M. Choda, *Int. J. Math.* **19**, 767 (2008).

[2] M. Choda, *Int. J. Math.* **24**, 1350066 (2013).

[3] M. Choda, *Int. J. Math.* **25**, 1450096 (2014).

[4] C.E. Shannon, *Bell Syst. Tech.* **27**, 379 (1948).

[5] S. Popa, *J. Operator Theory* **9**, 253 (1983).

[6] D. Petz, *Quantum Information Theory and Quantum Statistics*, Theoretical and Mathematical Physics, Springer, Berlin–Heidelberg 2008.

[7] S. Neshveyev, E. Størmer, *Dynamical Entropy in Operator Algebras*, Springer-Verlag, Berlin 2006.

[8] M. Ohya, D. Petz, *Quantum Entropy and Its Use*, Texts and Monographs in Physics, Springer-Verlag, Berlin 1993.

[9] F. Hiai, M. Mosonyi, D. Petz, C. Beny, *Rev. Math. Phys.* **23**, 691 (2011).

[10] K. Życzkowski, M. Kuś, W. Słomczyński, H.-J. Sommers, *J. Phys. A* **36**, 3425 (2003).

[11] M.D. Choi, *Ill. J. Math.* **18**, 565 (1974).

[12] G. Lindblad, *Commun. Math. Phys.* **65**, 281 (1979).

[13] V.F.R. Jones, V.S. Sunder, *Introduction to Subfactors*, Cambridge University Press, Cambridge 1997.