GENERAL TREATMENT OF ALPHABET-MESSAGE SPACE AND INTEGRAL REPRESENTATION OF ENTROPY

By Hisaharu Umegaki

1. Introduction.

In this paper we shall clarify a topological structure of the alphabet-message space of the memory channel in information theory, and study the integral representation of entropy amount from a general view point of a certain generalized message space. In order to apply to the general theory of entropy, the present fashion will develop a message space into more general treatment, in which the basic space X will be assumed to be totally disconnected. As will be shown in § 2, the alphabet-message space A^I is a totally disconnected compact space, and in § 3, a kind of theorem relative to sufficiency for a σ-field generated by a partition and a homeoporphism (cf. Theorem 2) and the others (Theorems 3 and 4) are concerned with the semi-continuity of entropy amount which are general form of Breiman's Theorem [1]. Finally, in § 4, it will be discussed about the function $h(x)$ found by Parthasarathy [7] whose integral defines the corresponding amount of entropy (cf. Theorem 5). It is also shown that the results in [9] can be generalized (cf. the footnote 2) below). The function $h(x)$ may give useful and interesting tool for the general theory of entropy of measure preserving automorphism or flow over a probabiltiy space.

2. Structure of message space.

A Hausdorff space X is called *totally disconnected* if X has a base consisting of closed-open sets, *clopen* say. In such a space X, a measure μ is called *normal*, if μ is regular and the mass of every non-dense set is zero. The space X is called *hyper-Stonean*, if it is compact and the union of carriers of all finite normal measures is dense in X. Such a space X is characterized by the existence of a normal measure μ (not necessarily finite) on X such that $\mu(G)>0$ for every non-empty open set G (cf. Dixmier [2]). Whence, the Banach space $C(X)$ of real continuous functions on X, with sup-norm $||\cdot||$, is isometrical and lattice isomorphic to the conjugate space of L^1-space $L^1(X, \mu)$. It is known that, these concepts on X are closely related with the theories of Boolean algebras and especially of operator algebras (=von Neumann algebras, cf. Dixmier [3]).

Received September 4, 1963.

In this section, for the sake of functional analysistic necessity and interest, it will be investigated on the topological structure of the space A^I, defined below.

Let A be an alphabet, i.e., a set consisting of finite number of elements. Put $A_k=A$ ($k=0, \pm 1, \pm 2, \cdots$) and denote $A^I=\times_{k=-\infty}^{\infty} A_k$ the doubly infinite product set. A^I will be called *message space*. In the memory channel, the input space of alphabet information source is taken as the measurable space $(A^I, \mathfrak{A})$ where $\mathfrak{A}$ is the σ-field generated by all finite dimensional cylinder sets in A^I.

Since each coordinate space A_k is a finite set, they are compact metric spaces relative to each discrete topology, hence by Tychonoff's theorem, A^I is also compact with countable base relative to the weak product topology and is metrizable. For each point $\mathfrak{a} \in A^I=(\cdots, a_{-1}, a_0, a_1, a_2, \cdots)$, denote

$$(1) \qquad [a_m, \cdots, a_n] \qquad (m \leqq n)$$

the $(n-m+1)$-length message, say *finite message*, i.e., the set of all $\mathfrak{x} \in A^I$ whose *k-th* coordinate equals to a_k ($k=m, \cdots, n$). The messages are obviously clopen. Let U be any non-empty open set in A^I. Then there exists a finite set of integers $J \subset I$ such that the projections $\mathrm{pr}_k(U)$ of U into k-th coordinate spaces A_k, $k \in I-J$, are the whole spaces A_k, respectively. Consequently, putting $m=\min\{k; k \in J\}$ and $n=\max\{k; k \in J\}$, then for any fixed $\mathfrak{a} \in U$, $[a_m, \cdots, a_n]$ is contained in U. Thus we obtain the following

THEOREM 1. *The message space A^I is a compact metric and totally disconnected space relative to the product topology, in which the shift is a homeomorphism on A^I and the σ-field $\mathfrak{A}$ consists of all Borel sets. Especially every finite message is a clopen set and the family of all finite messages is base in A^I as its topology. Furthermore A^I is not hyper-Stonean.*

Since the shift is obviously continuous and one to one on A^I onto itsef, it is homeomorphism on A^I. Hence the proof is remained only in the last part. If A^I is hyper-Stonean, then $C(A^I)$ is identified with the conjugate space of $L^1(A^I, \mu)$ for certain normal measure μ on A^I. Let $\{E_n\}$ be an infinite sequence of mutually disjoint and non-empty clopen sets in A^I, and let M be the weakly* clopen subspace of $C(A^I)$ generated by the sequence of the characteristic functions $\{C_{E_n}\}$, where the closure is concerning the weak topology as conjugate space. While M is separable relative to the norm-topology, because so is $C(A^I)$. Therefore M must be finite dimensional. This is a contradiction. Thus A^I is not hyper-Stonean.

3. Properties of entropy functional.

In order to devolope the theory of entropy over the message space A^I from a general point of view, we shall take as an information basic space a *totally disconnected compact Hausdorff space X* with a fixed *homeomorphism S*, which contains the case of A^I as a special case. Here, *every colopen set in X and the homeomor-*

phism S are corresponding to the finite message and the shift in A^I, respectively.

Denote $\mathfrak{X}$ the σ-field of all Borel sets in X. Let $\boldsymbol{P}_S$ be the set of all S-invariant regular probability measures $p, q, \cdots$ on X. Let $\mathfrak{G}$ be a fixed covering of X consisting of clopen sets such as any pair $U, U'\in\mathfrak{G}$ being disjoint. Such an $\mathfrak{G}$ is always finite by the compactness of X and it will be called *clopen partition* of X. As in §2 of [9], putting $\mathfrak{G}_n=\vee_{k=1}^n S^{-k}\mathfrak{G}$ or $\mathfrak{G}_\infty=\vee_{n=1}^\infty \mathfrak{G}_n$ the σ-field generated by $\{S^{-k}\mathfrak{G}\}_{k=1}^n$ or $\{\mathfrak{G}_n\}_{n=1}^\infty$, respectively. Then the entropy $H(p)=H(p, \mathfrak{G}, S)$ of each $p\in\boldsymbol{P}_S$ is defined by the limit

$$H(p)=-\lim\frac{1}{n}\sum_U p(U)\log p(U) \qquad (n\to\infty)$$

where $\sum_U$ means the summation over U of the atomic sets in $\mathfrak{G}\vee\mathfrak{G}_{n-1}$.

For any $p\in\boldsymbol{P}_S$, denote $P_p(U\mid\mathfrak{G}_n)$ and $P_p(U\mid\mathfrak{G}_\infty)$ the conditional probability functions of $U\in\mathfrak{G}$ conditioned by $\mathfrak{G}_n$ and $\mathfrak{G}_\infty$ in the measure space $(X, \mathfrak{X}, p)$, respectively, where $\mathfrak{X}$ is the σ-field of all Borel sets, and for a pair $p, q\in\boldsymbol{P}_S$, denote $q\ll p$, when q is absolutely continuous with respect to p. Then we prove

THEOREM 2. *For any pair $p, q\in\boldsymbol{P}_S$ with $q\ll p$, it holds that*

(2) $$P_p(U\mid\mathfrak{G}_\infty)(x)=P_q(U\mid\mathfrak{G}_\infty)(x) \quad q\text{-a.e. } x\in X \text{ and for every } U\in\mathfrak{G}.$$

Proof. Putting $\mathfrak{B}$ the σ-subfield generated by $\mathfrak{G}$ and $S^{-k}\mathfrak{G}$, $(k=\pm1, \pm2, \cdots)$, then, since $(X, \mathfrak{B}, p)$ is separable for each fixed $p\in\boldsymbol{P}_S$, from the generic property of $\mathfrak{B}$, it follows that every S-invariant set $B\in\mathfrak{B}$ belongs to $\mathfrak{G}_\infty$ (mod p). Besides for every $p\in\boldsymbol{P}_S$, putting $p'=(p\mid\mathfrak{B})$, the restriction of p over $\mathfrak{B}$, then p' is S-invariant probability measure over $(X, \mathfrak{B})$ and $P_p(U\mid\mathfrak{G}_\infty)=P_{p'}(U\mid\mathfrak{G}_\infty)$ for every $U\in\mathfrak{G}$ on C_p the carrier of p over $\mathfrak{B}$. Since C_p is S-invariant, it is $\mathfrak{G}_\infty$-measurable. Taking $p, q\in\boldsymbol{P}_S$, $q\ll p$, then $q'\ll p'$ and the Radon-Nikodym derivative dq'/dp' is S-invariant and $\mathfrak{B}$-measurable, and hence $\mathfrak{G}_\infty$-measurable (mod p). Therefore for every $V\in\mathfrak{G}_\infty$ and every $U\in\mathfrak{G}$

$$\begin{aligned}\int_V P_q(U\mid\mathfrak{G}_\infty)(x)\,dq(x)&=\int_V C_U(x)\,dq(x)=\int_V C_U(x)\frac{dq'}{dp'}(x)\,dp(x)\\&=\int_V P_p(U\mid\mathfrak{G}_\infty)(x)\,\frac{dq}{dp}(x)\,dp(x)=\int_V P_p(U\mid\mathfrak{G}_\infty)(x)\,dq(x)\end{aligned}$$

and (2) holds, where $C_U(x)$ is the characteristic function of U.

This theorem implies that $\mathfrak{G}_\infty$ is a sufficient subfield for the set $\{p';\, p\in\boldsymbol{P}_S\}$ of measures on $(X, \mathfrak{B})$ (in the sense of Halmos-Savage).

THEOREM 3. *For each $p\in\boldsymbol{P}_S$, there exists, uniquely within p-a.e., a bounded, upper-semicontinuous and $\mathfrak{G}_\infty$-measurable function $h_p(x)$ on X such that*

(3) $$h_p(x)=-\sum_{U\in\mathfrak{G}}P_p(U\mid\mathfrak{G}_\infty)\log P_p(U\mid\mathfrak{G}_\infty)(x) \qquad p\text{-a.e. } x\in X.$$

Proof. Since each $V\in\mathfrak{G}_n$ is clopen, $P_p(U\mid\mathfrak{G}_n)\in C(X)$. Putting

(4) $$h_{p,n}(x)=-\sum_{U\in\mathfrak{G}}P_p(U\mid\mathfrak{G}_n)(x)\log P_p(U\mid\mathfrak{G}_n)(x),$$

then $h_{p,n} \in C(X)$ and the sequence $\{h_{p,n}\}_{n=1}^{\infty}$ is monotone decreasing, by the Jensen's inequality, say

(5) $$h_p(x) = \lim_{n\to\infty} h_{p,n}(x),$$

and $h_p(x)$ is upper-semicontinuous on X. Furthermore, since each $h_{p,n}(x)$ is $\mathfrak{F}_n$-measurable, $h_p(x)$ is $\mathfrak{F}_\infty$-measurable. Besides, $\{h_{p,n}\}$ is semi-martingale over the probability space $(X, \mathfrak{X}, p)$, and hence by the semi-martingale convergence it satisfies (3).

By Theorem 3 and by the well known theorem of McMillan, it holds that

(6) $$H(p) = \int_X h_p(x)\, dp(x)$$

and $h_p(x)$ *is S-invariant in the a.e. sense.*

THEOREM 4. *The functional $H(p)$ over $\boldsymbol{P}_S$ is weakly* upper-semicontinuous, where the continuity is one with respect to the weak topology as functional over the Banach space $C(X)$.*

Proof. By the proof of Theorem 3, for every $p \in \boldsymbol{P}_S$

$$\int_X h_{p,n}(x)\, dp(x)\ (= H_n(p),\ say) \downarrow \int_X h_p(x)\, dp(x) = H(p) \qquad (n\to\infty)$$

and hence it is sufficient to prove the weak* continuity of $H_n(p)$. But this follows immediately from that

$$H_n(p) = -\sum_{U\in\mathfrak{F}} \int_X C_U(x) \log P_p(U \mid \mathfrak{F}_n)(x)\, dp(x)$$
$$= \sum_{U\in\mathfrak{F}} \sum_{V\in\mathfrak{F}_n} \Big[p(U\cap V) \log p(V) - p(U\cap V) \log p(U\cap V) \Big]$$

and that every $U \in \mathfrak{F}$ and $V \in \mathfrak{F}_n$ are clopen where $\sum_{V\in\mathfrak{F}_n}$ means the summation over atomic V in $\mathfrak{F}_n$.

4. Integral representation of amount of entropy by a universal function.

We shall show the theorem of Parthasarathy [7] for the present case. Assume the notations given in §3.

THEOREM 5. *For any clopen partition $\mathfrak{F}$, there exists universally a Borel measurable functions $h(x) = h(x, \mathfrak{F}, S)$ on X such that it is bounded, non-negative, S-invariant and satisfies*

(7) $$H(p) = \int_X h(x)\, dp(x) \qquad \textit{for every } p \in \boldsymbol{P}_S$$

(8) $$h(x) = h_p(x) \qquad p\textit{-a.e. } x \in X \textit{ and for every } p \in \boldsymbol{P}_S.$$

This function $h(x)$ was introduced by Parthasarathy in case of $X=A^I$. The proof will be also done along his construction, combined with Theorems 2 and 3, in which use is made of the well-known theorem of Kryloff-Bogoliouboff relative to the ergodic decomposition of invariant measure, cf. Oxtoby [6]. Before the proof, we shall give a preliminary and several lemmas.

Let $\mathfrak{U}$ be the field, of clopen sets, generated by $\{S^nU; U\in\mathfrak{F}, n=0, \pm1, \pm2, \cdots\}$, and $\mathfrak{B}$ the σ-field of Borel sets generated by $\mathfrak{U}$. Putting $\boldsymbol{C}_{\mathfrak{F}}$ the uniformly closed linear subspace of $C(X)$ generated by $\{C_U; U\in\mathfrak{U}\}$, then $\boldsymbol{C}_{\mathfrak{F}}$ is uniformly separable and has a countable dense subset $\{f_n\}_{k=1}^{\infty}\subset\boldsymbol{C}_{\mathfrak{F}}$. Putting

$$d(x, y) = \sum_{n=1}^{\infty}\frac{|f_n(x)-f_n(y)|}{2^n}||f_n||, \qquad x, y\in X,$$

$d(\cdot, \cdot)$ is a quasi-metric on X, and denote $x\sim y$ $(x, y\in X)$ if and only if $d(x, y)=0$. Moreover, put $\tilde{X}=X/\sim$ the quotient space of X with respect to the equivalence relation $\sim$ and put

$$\tilde{d}(\tilde{x}, \tilde{y}) = d(x, y) \qquad \text{for each pair} \quad x, y\in X,$$

where $\tilde{x}$ $(x\in X)$ is the class containing $x\in X$. Then $(\tilde{X}, \tilde{d})$ is a compact metric space, the canonical mapping $x\to\tilde{x}$ from X onto $\tilde{X}$ is continuous, and $\boldsymbol{C}_{\mathfrak{F}}$ is isomorphic with $C(\tilde{X})$ under the isomorphism $f\in\boldsymbol{C}_{\mathfrak{F}}\to\tilde{f}\in C(\tilde{X})$ defined by that, for each $x\in X$

$$f(y) = \tilde{f}(\tilde{x}) \qquad \text{for every} \quad y\in\tilde{x}. \tag{9}$$

Furthermore, $\tilde{\mathfrak{B}}=\{\tilde{B}; B\in\mathfrak{B}\}$ $(\tilde{B}=\{\tilde{x}\in\tilde{X}; x\in B\})$ is the σ-field of all Borel sets in $\tilde{X}$, and hence the function $\tilde{f}(\tilde{x})$ defined by (9) is Borel measurable on $\tilde{X}$ if and only if $f(x)$ is $\mathfrak{B}$-measurable on X. Putting $\tilde{S}$: $\tilde{x}\to(Sx)^{\sim}$, which is well defined mapping on X onto $\tilde{X}$ because $d(x, y)=0$ if and only if $d(Sx, Sy)=0$, then $\tilde{S}$ is a homeomorphism on $\tilde{X}$.

LEMMA 1. *For any non-negative linear functional ρ on $\boldsymbol{C}_{\mathfrak{F}}$ with norm one, there corresponds a probability measure μ_ρ over the measurable space $(X, \mathfrak{B})$ such that*

$$\rho(f)=\int_X f(x)\,d\mu_\rho(x) \qquad \textit{for every} \quad f\in\boldsymbol{C}_{\mathfrak{F}}. \tag{10}$$

Proof. This follows from the Riesz theorem. Putting $\tilde{\rho}(\tilde{f})=\rho(f)$, $f\in\boldsymbol{C}_{\mathfrak{F}}$, then $\tilde{\rho}$ is a non-negative linear functional on $C(\tilde{X})$ with norm one and hence there exists a regular probabiltiy measure $\mu_{\tilde{\rho}}$ on $\tilde{X}$ such that

$$\tilde{\rho}(\tilde{f})=\int \tilde{f}(\tilde{x})\,d\mu_{\tilde{\rho}}(\tilde{x}) \qquad \text{for every} \quad \tilde{f}\in C(\tilde{X}). \tag{11}$$

Put $\mu_\rho(B)=\mu_{\tilde{\rho}}(\tilde{B})$ for every $B\in\mathfrak{B}$, then $\mu_\rho(\cdot)$ is a regular probability measure over $(X, \mathfrak{B})$ and (10) follows from (11).

LEMMA 2. *If the functional ρ given in Lemma 1 is S-stationary, i.e. $\rho(Sf)=\rho(f)$ $((Sf)(x)=f(Sx))$, then the corresponding measure μ_ρ is S-invariant.*

Indeed, this follows immediately from that

$$\mu_\rho(S^{-1}V)=\rho(SC_V)=\rho(C_V)=\mu_\rho(V) \qquad \text{for every } V\in\mathfrak{U}.$$

Now, we refer, as Parthasarathy [7], the notion of Kyloff-Bogoliouboff's (K-B, say) theorem; cf. Oxtoby [6]. Denote

$$M_{x,n}(f)=\frac{1}{n}\sum_{k=1}^{n} f(S^k x) \qquad \text{for each } x\in X \text{ and } f\in \boldsymbol{C}_{\mathfrak{F}},$$

$n=1, 2, \cdots$. If the limit of $M_{x,n}(f)$ $(n\to\infty)$ exists for every $f\in \boldsymbol{C}_{\mathfrak{F}}$, $=M_x(f)$ say, then it is non-negative and S-stationary linear functional, with norm one, on $\boldsymbol{C}_{\mathfrak{F}}$. Such an $x\in X$ is called a *quasi-regular point* in X relative to $\boldsymbol{C}_{\mathfrak{F}}$. *Denote Q the set of all such points $x\in X$.* Then, by Lemma 2, for each $x\in Q$, there corresponds an S-invariant measure $m_x=\mu_{M_x}$ over $(X, \mathfrak{B})$ such that

$$M_x(f)=\int_X f(y)\, dm_x(y) \qquad \text{for every } f\in \boldsymbol{C}_{\mathfrak{F}}, \tag{12}$$

and m_x satisfies

$$m_x(B)=m_{Sx}(B) \qquad \text{for every } B\in\mathfrak{B} \text{ and for every } x\in Q. \tag{13}$$

If m_x is ergodic with respect to S over $(X, \mathfrak{B})$, then x is called a *regular point* in X relative to $\boldsymbol{C}_{\mathfrak{F}}$. *Denote R the set of all regular points in X.* Then K-B theorem implies that

$$Q\in\mathfrak{B},\ R\in\mathfrak{B} \text{ and } p(Q)=p(R)=1 \qquad \text{for every } p\in \boldsymbol{P}_S. \tag{14}$$

Indeed, since $\tilde{X}$ is compact metric space with homeomorphism $\tilde{S}$, and since for each $x\in X$

$$M_{x,n}(f)=\frac{1}{n}\sum_{k=1}^{n} \tilde{f}(\tilde{S}^k\tilde{x}) \qquad \text{for every } \tilde{f}\in C(\tilde{X}),$$

and $\tilde{Q}$ or $\tilde{R}$ are the sets of all quasi-regular or regular point in $\tilde{X}$ relative to $C(\tilde{X})$, respectively, and both $\tilde{Q}, \tilde{R}$ are Borel sets in $\tilde{X}$ and invariant measure one, i.e., $\tilde{p}(\tilde{Q})=\tilde{p}(\tilde{R})=1$ for every $p\in \boldsymbol{P}_S$ (cf. Oxtoby [6], (2. 4)), where $\tilde{p}$ is the $\tilde{S}$-invariant and regular probability measure over $\tilde{X}$ defined by $\tilde{p}(\tilde{B})=p(B)$, $B\in\mathfrak{B}$, that is, (14) holds.

LEMMA 3. *For each bounded $\mathfrak{B}$-measurable functions f on X,*[1]

$$\int_X f(x)\, dm_r(x), =f^{\natural}(r) \quad \textit{say}, \tag{15}$$

---

1) The mapping $f\to f^{\natural}$, defined over the Banach space of all bounded $\mathfrak{B}$-measuable functions $B(X)$ (with sup-norm) into itself, coincides with the concept of the expectation in the sense of Nakamura-Turumaru [5] and also the conditional expectation in the sense of Umegaki [8].

($f^{\natural}(r)=M_r(f)$ if $f\in C_{\mathfrak{F}}$), is a bounded, $\mathfrak{B}$-measurable and S-invariant function over R, and it satisfies

$$\int_X f(x)\,dp(x)=\int_R f^{\natural}(r)\,dp(r)=\int_R\left(\int_X f(x)\,dm_r(x)\right)dp(r). \tag{16}$$

This follows from K-B Theorem (cf. [6], (2. 6)) that, for $r\in R$

$$f^{\natural}(r)=\int_{\tilde{X}}\tilde{f}(\tilde{x})\,dm_{\tilde{r}}(\tilde{x}),\quad =\tilde{g}(\tilde{r})\quad \text{say,}$$

and $\tilde{g}$ is Borel measurable over $\tilde{R}$, and that

$$\int_{\tilde{R}}\tilde{g}(\tilde{r})\,d\tilde{p}(\tilde{r})=\int_{\tilde{X}}\tilde{f}(\tilde{x})\,d\tilde{p}(\tilde{x}).$$

The S-invariance of $f^{\natural}(x)$ follows from (13).

Proof of Theorem 5. Put $\boldsymbol{P}'_S$ be the set of all S-invariant probability measures over the measurable space $(X,\mathfrak{B})$. Then the theorems and their proofs in § 3 hold for $p, q, \cdots$ in $\boldsymbol{P}'_S$ without chainging their statements. Since $m_r\in\boldsymbol{P}'_S$ for each $r\in R$, the function $h_{m_r}(x)$ over X can be defined by (5). Putting

$$h(r)=\begin{cases}\displaystyle\int_X h_{m_r}(x)\,dm_r(x) & \text{for}\quad r\in R,\\ 0 & \text{for}\quad r\notin R,\end{cases} \tag{17}$$

(i.e., $h(r)=H(m_r)$ for $r\in R$), then $h(\cdot)$ is bounded, S-invariant and $\mathfrak{B}$-measurable. Indeed, put $h_n(r,x)=h_{m_r,n}(x)$ $(r\in R)$ and put

$$g_n(r)=\int_X h_n(r,x)\,dm_r(x)=-\sum_{U\in\mathfrak{F}}\sum_{V\in\mathfrak{F}_n} m_r(U\frown V)[\log m_r(U\frown V)-\log m_r(V)].$$

Whence, since $m_r(W)$ $(=M_r(C_W),\ W\in\mathfrak{U})$ is $\mathfrak{B}$-measurable on R, so is $g_n(r)$ on R. Furthermore, as in the proof of Theorem 3, since $h_n(r,x)\downarrow h_{m_r}(x)$, $=h(r,x)$ say,

$$h(r)=\int_X h(r,x)\,dm_r(x)=\lim_{n\to\infty} g_n(r)\qquad (r\in R)$$

and $h(r)$ is $\mathfrak{B}$-measurable on R and hence $h(x)$ is so on X. The boundedness of $h(x)$ follows from (5) for $p=m_r$ and the definition (17) of $h(x)$, and S-invariance follows from $m_r=m_{Sr}$.

As Parthasarathy ([7], Theorem 2. 6), for each $U\in\mathfrak{F}$, $V\in\mathfrak{F}_\infty$ and $p\in\boldsymbol{P}_S$, and for every fixed $q\in\boldsymbol{P}_S$, $q\ll p$,

$$q(U\frown V)=\int_V P_q(U\mid\mathfrak{F}_\infty)(x)\,dq(x)=\int_V P_p(U\mid\mathfrak{F}_\infty)(x)\,dq(x)\qquad(\text{by Theorem 2})$$

$$=\int_R\left(\int_V P_p(U\mid\mathfrak{F}_\infty)(x)\,dm_r(x)\right)dq(r) \qquad \text{(by Lemma 3)}$$

and

$$q(U\frown V)=\int_X C_{U\frown V}(x)\,dq(x)=\int_R (C_{U\frown V})^{\natural}(r)\,dq(r)=\int_R m_r(U\frown V)\,dq(r)$$

$$=\int_R\left(\int_V P_{m_r}(U\mid\mathfrak{F}_\infty)(x)\,dm_r(x)\right)dq(r) \qquad \text{(by Lemma 3).}$$

This and (13) imply that $P_p(U\mid\mathfrak{F}_\infty)(x)=P_{m_r}(U\mid\mathfrak{F}_\infty)(x)$ and

(18) $h_p(x)=h_{m_r}(x)$ for m_r-*a.e.* $x\in X$ and for each fixed $r\in R$

within p-*a.e.* r in R. Thus we obtain the required equality (7):

$$H(p)=\int_X h_p(x)\,dp(x)=\int_R h_p^{\natural}(r)\,dp(r)=\int_R\left(\int_X h_p(x)\,dm_r(x)\right)dp(r)$$

$$=\int_R\left(\int_X h_{m_r}(x)\,dm_r(x)\right)dp(r)=\int_R h(r)\,dp(r)=\int_X h(x)\,dp(x).$$

Besides, by Theorems 2, 3 and (7), whenever $q\ll p$ ($p, q\in \boldsymbol{P}_S$),

$$H(q)=\int_X h_q(x)\,dq(x)=\int_X h_p(x)\,dq(x)=\int_X h(x)\,dq(x),$$

hence

$$\int_X h(x)f(x)\,dp(x)=\int_X h_p(x)f(x)\,dp(x)$$

for every $f\in L^1(X, p)$, S-invariant. Since both h and h_p are S-invariant, (8) is obtained.

The function $h(x)$ will be called *universal entropy function* associated with the clopen partition $\mathfrak{F}$ and the homeomorphism S, and sometimes precisely denote $h(x)=h(x, \mathfrak{F}, S)$. Let $\boldsymbol{L}$ be the Banach space of all bounded signed regular measures over X with the norm of total variation. Then putting

$$H(\xi)=\int_X H(x)\,d\xi(x) \qquad \text{for every} \quad \xi\in\boldsymbol{L},$$

$H(\cdot)$ is a bounded non-negative definite linear functional over $\boldsymbol{L}$, and is S-stationary, i.e., $H(S\xi)=H(\xi)$ for every $\xi\in\boldsymbol{L}$, where $S\xi\in\boldsymbol{L}$ is defined by $(S\xi)(V)=\xi(S^{-1}V)$ for every Borel set V in X. The functional $H(\xi)$ over $\boldsymbol{L}$ coincides with the functional

$H(\xi, \mathfrak{F}, S)$ in the paper [9],[2] and again it is called the *entropy functional* over $\boldsymbol{L}$ associated with a clopen patition $\mathfrak{F}$ and a homeomorphism S.

It is the author's pleasure to acknowledge that he gets valuable discussions from Prof. M. Nakamura.

References

[1] Breiman, L., On achieving channel capacity in finite-memory channels. Ill. J. Math. **4** (1960), 246–252.

[2] Dixmier, J., Sur certains especes considérés par M. H. Stone. Sum. Bras. Math. **11** (1951), 1–32.

[3] Dixmier, J., Les algèbres d'opérateurs dans l'espace Hibertien. Paris (1957).

[4] Halmos, P. R., Entropy in ergodic theory, Univerity of Chicago (1959).

[5] Nakamura, M., and T. Turumaru, Expectations in an operator algebra. Tôhoku Math. J. **6** (1954), 189–204.

[6] Oxtoby, J. C., Ergodic sets. Bull. Amer. Math. Soc. **58** (1952), 116–136.

[7] Parthasarathy. K. R., On the integral representation of the rate of transmission of a stationary channel. Ill. J. Math. **5** (1961), 299–305.

[8] Umegaki, H., Conditional expectation in an operator algebra. Tôhoku Math. J. **6** (1954), 177–181.

[9] Umegaki, H., A functional method on amount of entropy. Kōdai Math. Sem. Rep. **15** (1963), 162–175.

Department of Mathematics,
Tokyo Institute of Technology.

---

2) In the preceding paper [9], §4, it was assumed that if the measurable space $(X, \mathfrak{X})$ with measurable transformation S has denombrable generator, then it has maximal S-invariant probabitity measure relative to the ordering $\ll$ of absolute continuity. However, in general, this does not hold, for example, when S is identity mapping from X onto X, $X=$the interval $[0, 1]$ and $\mathfrak{X}$ is σ-field of Borel subsets. Therefore, it should be corrected such as

P. 168, line 3~line 4 "Then $P(X, S)$ is necessarily...dominates all $p \epsilon P(X, S)$" reads such as

"Assume that there exists an S-invariant probability measure μ, and denote $P(X, S)$ (resp., $L(X, S)$) the sets of all S-invariant 'probability' (resp., 'bounded signed') measures $p, \cdots$ (resp., $\xi, \cdots$) which are absolutely continuous with respect to μ."

Hence in the parts below in [9] the dominatedness for the sets $P(X, S)$ and $L(X, S)$ of measures should be assumed.

However, the measurable space $(X, \mathfrak{X})$ given in the paper [9], can be represented by a totally disconnected compact space in preserving the measurability structure and where the measurable transformation S is mapped to a homeomorphism. Therefore, by Theorem 5 and the discussions in §4 of the present paper, the theorems in [9] hold without assuming the denombrability of the measurable space.