A FUNCTIONAL METHOD FOR STATIONARY CHANNELS

By Hisaharu Umegaki

Dedicated to Professor K. Kunugi on his sixtieth birthday

1. Introduction.

The concept of the finite memory channel of Shannon has been formulated in the purely mathematical form by McMillan and Khinchin (Cf. [10]), and established in the present elegant style by Feinstein [6]. Around the finite memory channels as its focal point, there exist various theorems, in which one of the most important results is the theorem of the equality of C_s and C_e of stationary and ergodic capacities. The problem 'whether the equality holds' has been an open question since Khinchin's paper [10]. This equality has been recently proved by many authors: Tsaregradsky [13], Carleson [3], Feinstein [7], Breiman [2], Parthasarathy [11] and others. In this paper we shall describe it in an abstract form.

The purpose of this paper is to introduce an abstract characterization of finite or infinite memory channel in which the input space and the output space are compact (totally disconnected) Hausdorff spaces with a pair of fixed homeomorphisms, and in which the channel distribution has a continuous property. In particular, every memory channel has always these properties. The usual memory channels are based upon their message symbols with practical applications. However their symbols may sometimes produce certain troublesome complications for the developments of several mathematical computations of them. The message symbols, in the present construction of the channel, will not be presented, and they will be replaced by sets with the property of the closed-openness (clopen, say). The descriptions will be given only by topological and functional forms, that is, they will be described by topological and Banach spaces methods. The entropy functional $H(\cdot)$ (cf. Umegaki [15] and [16]) and the transmission functional $\mathfrak{R}(\cdot)$ are defined over the Banach space of bounded signed regular measures, and they depend upon a clopen partition, or upon a pair of such partitions in the input and output spaces.

In § 2, in order to clarify the abstract stationary channel (X, ν, Y) defined below, the definition of stationary finite memory channel (A^I, ν, B^I) will be first stated with respect to the conditions (m 1)$\sim$(m 5). These conditions will be replaced below by the conditions (C 1)$\sim$(C 5) in the channel (X, ν, Y), respectively. In § 3, several notations and preliminaries will be given, and in § 4 the stationary channel (X, ν, Y)

Received September 5, 1963.

will be defined by the conditions (C 1)$\sim$(C 3), and it will be described the linear transformations K' and K'' associated with the channel distribution, which was previously introduced, from a point of view of the theory of von Neumann algebra, by Echigo and Nakamura [5]. In order to derive a functional analysis for the theory of channel, it seems to us that the description of K' and K'' may become one of the starting points of it. In § 5, using a method in the author's recent papers [15] and [16] it will be defined a functional derived from the transmission rate, and under assumption of total disconnectedness for the input and output spaces, the integral representation theorems of Parthasarathy [11] for the memory channel will be proved for the channel (X, ν, Y), and also proved that the transmission rate is always represented by an integral of the universal entropy function (cf. § 4, p. 25 of [16]) with respect to a measure over the compound space. In § 6, it will be introduced the condition (C 4) of a concept of continuity of the channel (X, ν, Y), which is an abstractly generalized condition (m 4) of Khinchin. Under the condition (C 4), the weak continuities of the transformations K' and K'' are proved. Furthermore, along Breiman's construction, the existence of the stationary capacity C_s is proved, and combining Adler's condition (C 5), cf. [1], of asymptotic independence of the channel distribution ν, the equality $C_s = C_e$ of stationary and ergodic capacities will be proved, in (X, ν, Y).

The abstract of this paper was partly published in [14].

2. A concept of finite memory channel.

Let A be an alphabet, that is, A is a discrete set consisting of finite number of elements. Denote $A^I = \times_{k=-\infty}^{\infty} A_k$ $(A_k = A,\ k = 0, \pm 1, \pm 2, \cdots)$ the doubly infinite product of A, this is the set of all doubly infinite sequences

$$\mathfrak{a} = (\cdots, a_{-1}, a_0, a_1, \cdots),$$

$a_k \in A$, $k = 0, \pm 1, \pm 2, \cdots$. Let $\mathfrak{A}_0$ be the family of all finite dimensional cylinder sets and $\mathfrak{A}$ the Borel field generated by $\mathfrak{A}_0$. Then $(A^I, \mathfrak{A})$ is a measurable space with the denombrable (measurable) generator $\mathfrak{A}_0$. Let S be the shift transformation defined by

$$S(\cdots, a_{-1}, a_0, a_1, \cdots) = (\cdots, a'_{-1}, a'_0, a'_1, \cdots)$$

where $a'_k = a_{k+1}$, $k = 0, \pm 1, \pm 2, \cdots$. Then S is an invertible measurable transformation from A^I onto A^I. Denote the cylinder sets

$$\{\mathfrak{a} \in A^I;\ k\text{-th coordinate} = a_k,\ k = n, n-1, \cdots\}$$

or

$$\{\mathfrak{a} \in A^I;\ k\text{-th coordinate} = a_k,\ k = m, \cdots, n\} \qquad (m \leqq n)$$

by

$$[\cdots a_n] \qquad \text{or} \qquad [a_m \cdots a_n],$$

respectively. These sets are called *messages*, especially the set $[a_m \cdots a_n]$ is called *finite message* of length $n-m+1$, and A^I or $(A^I, \mathfrak{A}, S)$ is called *message space*.

Let $(A^I, \mathfrak{A}, S)$ and $(B^I, \mathfrak{B}, T)$ be a pair of the message spaces, with their shift transformations S and T, which are introduced by the corresponding pair A and B of alphabets, respectively.

A function $\nu(\mathfrak{a}, D)$ defined over the cartesian product $A\times\mathfrak{B}$ (or the triple (A^I, ν, B^I)) is called *stationary memory channel*, if

(m 1) For each $D\in\mathfrak{B}$, $\nu(\cdot, D)$ is a measurable function over the space $(A^I, \mathfrak{A})$,

(m 2) For each $\mathfrak{a}\in A^I$, $\nu(\mathfrak{a}, \cdot)$ is a probability measure over the space $(B^I, \mathfrak{B})$,

(m 3) the equality $\nu(S\mathfrak{a}, TD)=\nu(\mathfrak{a}, D)$ holds for every $\mathfrak{a}\in A^I$ and for every $D\in\mathfrak{B}$; and it is called *nonanticipating* if it satisfies that: for each $\mathfrak{a}^0=(\cdots, a_{-1}^0, a_0^0, a_1^0, \cdots)\in A^I$ and for any message $[\cdots b_{n-1}\, b_n]\subset B^I$

$$\nu(\mathfrak{a}^0, [\cdots b_{n-1}\, b_n])=\nu(\mathfrak{a}, [\cdots b_{n-1}\, b_n])$$

holds for every $\mathfrak{a}=(\cdots, a_{-1}, a_0, a_1, \cdots)\in A^I$ within $a_k=a_k^0$ $(k=n, n-1, \cdots)$.

Moreover the distribution $\nu(\cdot, \cdot)$ (or the triple (A^I, ν, B^I)) is called *finite memory*, if there exists an integer $l>0$ satisfying the following conditions:

(m 4) for each $\mathfrak{a}^0\in A^I$ and for any message $[b_m \cdots b_n]\subset B^I$ $(m\leqq n)$ the equality

$$\nu(\mathfrak{a}^0, [b_m \cdots b_n])=\nu(\mathfrak{a}, [b_m \cdots b_n])$$

holds for every $\mathfrak{a}=(\cdots, a_{-1}, a_0, a_1, \cdots)\in A^I$ within $a_k=a_k^0$ $(k=m-l, \cdots, n)$, and

(m 5) for any two finite messages $[b_i \cdots b_j]$ and $[b_m' \cdots b_n']\subset B^I$ such that $j+l<m$, the equality

$$\nu(\mathfrak{a}, [b_i \cdots b_j]\frown[b_m' \cdots b_n'])=\nu(\mathfrak{a}, [b_i \cdots b_j])\cdot\nu(\mathfrak{a}, [b_m' \cdots b_n'])$$

holds for all $\mathfrak{a}\in A^I$.

The smallest integer $l>0$ for which (m 4) and (m 5) hold, is called the *memory length* or the *memory of the channel.* If no such $l>0$ exists, the channel (A^I, ν, B^I) will be called *infinite memory.*

The *message space* A^I is a *totally disconnected compact* metric space by the weak product topology, the field $\mathfrak{A}_0$ of finite dimensional cylinder sets is the base of this topology and it consists of denombrable clopen sets, in which S is a *homeomorphism* (cf. [16], Theorem 1). And so are the *message space* B^I and the *compound message space* $A^I\times B^I$ with the denombrable bases $\mathfrak{B}_0$ and $\mathfrak{A}_0\otimes\mathfrak{B}_0$ of all finite dimensional cylinder sets, respectiveley.

3. Preliminary and notatinns.

Let X be a compact Hausdorff space with a (fixed) homeomorphism S. Denote $C(X)$ the Banach space of all (real) continuous functions on X with the sup-norm and $\boldsymbol{L}(X)$ the Banach space of all the bounded signed regular measures with norm $\|\ \|_1$ of total variation, then each measure $\xi\in\boldsymbol{L}(X)$ corresponds to a bounded linear functional F_ξ of $C(X)$ such as

$$F_\xi(f)=\int_X f(x)\, d\xi(x) \qquad \text{for every} \quad f\in C(X), \tag{1}$$

and conversely. The correspondence $\xi\rightarrow F_{\xi}$ defines a linear isometric mapping between $\boldsymbol{L}(X)$ and the conjugate space $C^{*}(X)$ of $C(X)$, and for the sake of notational convenience they will be identified: $\boldsymbol{L}(X)=C^{*}(X)$ by $\xi=F_{\xi}$, or put

$$(1') \qquad F_{\xi}(f)=<f, \xi>.$$

Denote $\boldsymbol{P}(X)$ the set of all probability regular measures on X. Then $\boldsymbol{P}(X)$ coincides with the space of functionals $\rho\in C^{*}(X)$ with norm one and $\rho(I)=1$ (I being the identity function on X), and hence $\boldsymbol{P}(X)$ is convex and weakly* compact set.[1] For any function f and measure ξ denote $(Sf)(x)=f(Sx)$ and $(S\xi)(U)=\xi(S^{-1}U)$. Denote $\boldsymbol{L}(X, S)\subset\boldsymbol{L}(X)$ the set of all measures with $S\xi=\xi$ (called S-invariant) and put $\boldsymbol{P}(X, S)=\boldsymbol{P}(X)\frown\boldsymbol{L}(X, S)$. By the fixed point theorem, $\boldsymbol{L}(X, S)$ and $\boldsymbol{P}(X, S)$ are non-empty, and hence $\boldsymbol{L}(X, S)$ is closed linear subspace of $\boldsymbol{L}(X)$ and $\boldsymbol{P}(X, S)$ is the weakly* closed convex hull of the set of extreme points in $\boldsymbol{P}(X, S)$. Donote $\boldsymbol{P}_{e}(X)$ the set of all ergodic (relative to S) measures $p\in\boldsymbol{P}(X, S)$. Then $\boldsymbol{P}_{e}(X)$ coincides with the set of all extreme points of $\boldsymbol{P}(X, S)$.

Let $\mathfrak{X}$ be the σ-field of all Borel subsets of X. A subfamily $\mathfrak{F}$ of $\mathfrak{X}$ is called a *partition*, if it covers X and any pair of different sets in $\mathfrak{F}$ is disjoint. Denote by $\mathfrak{F}_{n}=\vee_{k=1}^{n}S^{-k}\mathfrak{F}$ $(S^{-k}\mathfrak{F}=\{S^{-k}U;\ U\in\mathfrak{F}\})$ or $\mathfrak{F}_{\infty}=\vee_{k=1}^{\infty}S^{-k}\mathfrak{F}$ the σ-subfield of $\mathfrak{X}$ generated by $\{S^{-k}\mathfrak{F}\}_{k=1}^{n}$ or by $\{S^{-k}\mathfrak{F}\}_{k=1}^{\infty}$, respectively, and by $\mathfrak{F}_{n}^{\circ}$ the partition which generates $\mathfrak{F}_{n}$.

For any fixed measure $p\in\boldsymbol{P}(X)$ and for any fixed σ-subfield $\mathfrak{B}\subset\mathfrak{X}$, denote $P_{p}(U|\mathfrak{B})$ the conditional probability, in the probability measure space $(X, \mathfrak{X}, p)$, of a sets $U\in\mathfrak{X}$ conditioned by $\mathfrak{B}$. Let $\boldsymbol{L}^{+}(X)$ be the set of all non-negative measures in $\boldsymbol{L}(X)$, i.e., the positive cone of $\boldsymbol{L}(X)$. Then for any non-trivial $\xi\in\boldsymbol{L}^{+}(X)$, $\xi_{1}=\xi/||\xi||_{1}$ belongs to $\boldsymbol{P}(X)$ and put $P_{\xi}(U|\mathfrak{B})=P_{\xi_{1}}(U|\mathfrak{B})$. For a finite partition $\mathfrak{F}$, define a functional over $\boldsymbol{L}^{+}(X, S)$

$$(2) \qquad H(\xi, \mathfrak{F}, S)=-\sum_{U\in\mathfrak{F}}\int_{X}P_{\xi}(U|\mathfrak{F}_{\infty})\log P_{\xi}(U|\mathfrak{F}_{\infty})\,d\xi(x)$$

for every $\xi\in\boldsymbol{L}^{+}(X, S)$, where for $\xi=0$ put $H(\xi, \mathfrak{F}, S)=0$ and the base of log is 2. Then, as the previous paper [15], this satisfies the following (cf. Halmos [8])

$$(3) \qquad \begin{aligned} H(\xi, \mathfrak{F}, S)&=-\sum_{U\in\mathfrak{F}}\lim_{n\rightarrow\infty}\int_{X}P_{\xi}(U|\mathfrak{F}_{n})\log P_{\xi}(U|\mathfrak{F}_{n})\,d\xi(x)\\ &=-\lim\frac{1}{n}\sum_{U\in(\mathfrak{F}\vee\mathfrak{F}_{n-1})^{\circ}}\xi(U)\log\xi(U)\\ &\leqq-\sum_{U\in\mathfrak{F}}\xi(U)\log\xi(U)+\xi(X)\log\xi(X)\end{aligned}$$

---

1) When E is a Banach space and E^{*} is its conjugate space, the weak topology on E^{*} as functional over E is said to be weak* topology. That is, the generic (weak*) neighbourhoods are determined by $F_{0}\epsilon E^{*}$, $\varepsilon>0$ and a finite set, $f_{1}, \cdots, f_{n}$, of elements in E:

$N(F_{0}, f_{1}, \cdots, f_{n}, \varepsilon)=\{F\epsilon E^{*};\ |F_{0}(f_{i})-F(f_{i})|<\varepsilon, i=1, \cdots, n\}$. This topology is equivalent to that defined by the weak* convergence of nets, $F_{\alpha}\rightarrow F$ weak* if and only if $F_{\alpha}(f)\rightarrow F(f)$ for every $f\epsilon E$.

Whence, $H(\xi, \mathfrak{F}, S)$ is uniquely extended to non-negative bounded linear functional over the space $\boldsymbol{L}(X, S)$ and is called the *entropy functional,* cf. Umegaki [15] and [16].

4. Stationary channel.

As described for X, let Y be a compact Hausforff space with the fixed homeomorphism T and with the σ-field $\mathfrak{Y}$ of the Borel subsets in Y. Denote $Z=X\times Y$ the cartesian product X and Y, $S\otimes T$ the product homeomorphism on Z of S and T, and $\mathfrak{Z}$ the σ-field $\mathfrak{X}\otimes\mathfrak{Y}$. The spaces $C(Y)$, $\boldsymbol{L}(Y)$, $\boldsymbol{L}(Y, T)$, $\boldsymbol{P}(Y, T)$ and $C(Z)$, $\boldsymbol{L}(Z)$, $\boldsymbol{L}(Z, S\otimes T)$, $\boldsymbol{P}(Z, S\otimes T)$, etc. associated with Y and Z are defined as in § 3 for X.

DEFINITION 1. The triple $(X, \nu, \mathfrak{Y})$ is called *stationary channel*, with the homeomorphism $S\otimes T$, if the function $\nu(x, V)$ is defined on the product set $X\times\mathfrak{Y}$ such that:

(C 1) *for each fixed* $V\in\mathfrak{Y}$, $\nu(\cdot, V)$ *is a Borel measurable function over* $(X, \mathfrak{X})$,

(C 2) *for each fixed* $x\in X$, $\nu(x, \cdot)$ *is a probability regular measure over* $(Y, \mathfrak{Y})$,

(C 3) *the function* $\nu(\cdot, \cdot)$ *is stationary with respect to* $S\otimes T$, *i.e.,*

$$\nu(Sx, TV) = \nu(x, V) \qquad \text{for every} \quad x\in X \text{ and } V\in\mathfrak{Y}.$$

The function $\nu(\cdot, \cdot)$ is called *channel distribution*, X and Y are called *input* and *output* spaces, and Z *compound space*, respectively. The conditions (C 1), (C 2) and (C 3) are simultaneously corresponding to the conditions (m 1), (m 2) and (m 3) in the stationary memory channel, cf. § 2.

For every $\xi\in\boldsymbol{L}(X)$, putting

$$\xi'(V)=\int_X \nu(x, V)\, d\xi(x), \ V\in\mathfrak{Y}$$

and

$$\xi''(U\times V)=\int_U \nu(x, V)\, d\xi(x), \ U\in\mathfrak{X}, \ V\in\mathfrak{Y},$$

then $\xi'\in\boldsymbol{L}(Y)$ and ξ'' is uniquely extended to a bounded measure over $(Z, \mathfrak{Z})$, denote it again by ξ'' which belongs to $\boldsymbol{L}(Z)$. Whence it holds that for every bounded Borel measurable function $f(z)$ $(=f(x, y))$

$$\int_X\int_Y f(x, y)\,\nu(x, dy)\, d\xi(x)=\int_Z f(z)\, d\xi''(z). \tag{4}$$

Denote K' and K'' the mappings $\xi\rightarrow\xi'$ and $\xi\rightarrow\xi''$ on $\boldsymbol{L}(X)$ into $\boldsymbol{L}(Y)$ and $\boldsymbol{L}(Z)$, respectively. Then the following will be proved:

THEOREM 1. *The mappings K' and K'' are non-negative, bounded linear transformations with norm one on $\boldsymbol{L}(X)$ into $\boldsymbol{L}(Y)$ and on $\boldsymbol{L}(X)$ into $\boldsymbol{L}(Z)$ respectively, satisfying*

$$||K'\xi||_1=||K''\xi||_1=||\xi||_1 \qquad \textit{for every} \quad \xi\in\boldsymbol{L}^{+}(X) \tag{5}$$

and

$$(6) \qquad TK'=K'S \quad \text{and} \quad (S\otimes T)K''=K''S.$$

These K' and K'' will be called the *channel transformations* associated with the channel distribution ν. This functional formulation of K' and K'' was firstly given by Echigo and Nakamura [5] from the point of veiw of the theory of von Neumann algebra.

Proof. The linearity and non-negative definiteness of K' and K'' are obvious from the definitions of them. For every $\xi\in \boldsymbol{L}(X)$ there exists a real measurable function $\theta(z)$ on Z such that

$$d|K''\xi|(z)=\exp(i\theta(z))\, d(K''\xi)(z), \qquad i=\sqrt{-1}.$$

Hence

$$||K''\xi||_1=\int_Z d|K''\xi|(z)=\int_Z \exp(i\theta(z))\, d(K''\xi)(z)$$

$$=\int_X\int_Y \exp(i\theta(x, y))\nu(x, dy)d\xi(x)\leqq\int_X d|\xi|(x)=||\xi||_1.$$

While for $\xi\in \boldsymbol{L}^+(X)$,

$$||K''\xi||_1=(K''\xi)(X, Y)=\int_X \nu(x, Y)d\xi(x)=\xi(X)=||\xi||_1$$

and $||K''||=1$, similarly $||K'||=1$ is obtained, and (5) is proved. (6) follows from that: for every $U\in\mathfrak{X}$ and $V\in\mathfrak{Y}$,

$$((S\otimes T)K''\xi)(U\times V)=(S\otimes T)\xi''(U\times V)=\xi''(S^{-1}U\times T^{-1}V)=\int_{S^{-1}U}\nu(x, T^{-1}V)d\xi(x)$$

$$=\int_{S^{-1}U}\nu(Sx, V)d\xi(x)=\int_U \nu(x, V)d\xi(S^{-1}x)=(K''S\xi)(U\times V)$$

and

$$(TK'\xi)(V)=(S\otimes T)K''\xi(X\times V)=(K''S\xi)(X\times V)=(K'S\xi)(V).$$

It follows from Theorem 1 that

COROLLARY 1. 1. *For the transformations K' and K'', it holds that*

$$(7) \qquad K'\boldsymbol{P}(X)\subset\boldsymbol{P}(Y) \quad \text{and} \quad K''\boldsymbol{P}(X)\subset\boldsymbol{P}(Z)$$

and

$$(8) \qquad K'\boldsymbol{P}(X, S)\subset\boldsymbol{P}(Y, T) \quad \text{and} \quad K''\boldsymbol{P}(X, S)\subset\boldsymbol{P}(Z, S\otimes T).$$

5. Transmission functional and its integral representation.

Assume that the spaces X and Y are *totally disconnected with the bases of clopen sets* $\mathfrak{X}_0$ of X and $\mathfrak{Y}_0$ of Y, respectively. If the partition $\mathfrak{P}$ is a subfamily of

$\mathfrak{X}_0$ then $\mathcal{G}$ is called a *clopen partition.* For any fixed pair of clopen partitions $\mathcal{G}$ of X and $\mathcal{Q}$ of Y, denote $\mathcal{K}=\mathcal{G}\otimes\mathcal{Q}$ the clopen partitions $\{U\times V;\ U\in\mathcal{G},\ V\in\mathcal{Q}\}$ of $Z=X\times Y$. Then it can be introduced three entropy functionals

$$(9)\qquad H(\xi)=H(\xi,\mathcal{G},S),\ H_1(\eta)=H(\eta,\mathcal{Q},T)\ \text{and}\ H_2(\zeta)=H(\zeta,\mathcal{K},S\otimes T)$$

over the spaces $\boldsymbol{L}(X,S)$, $\boldsymbol{L}(Y,T)$ and $\boldsymbol{L}(Z,S\otimes T)$, respectively. These functionals H, H_1 and H_2 are represented by the integrals of the bounded non-negative Borel measurable functions $h(x)=h(x,\mathcal{G},S)$, $h_1(y)=h(y,\mathcal{Q},T)$ and $h_2(z)=h(z,\mathcal{K},S\otimes T)$, which are universally determined by the partitions $\mathcal{G}$, $\mathcal{Q}$ and $\mathcal{K}$, resp., and by the heomeomorphisms S, T and $S\otimes T$, resp. (called universal entropy functions, cf. [16], the final part of § 4). Furthermore these h, h_1 and h_2 are invariant relative to each homeomorphism. The integral reprentations are expressed such that (cf. [16], § 4)

$$(10)\qquad H(\xi)=\int_X h(x)d\xi(x),\ H_1(\eta)=\int_Y h_1(y)d\eta(y)\ \text{and}\ H_2(\zeta)=\int_Z h_2(z)d\zeta(z)$$

for every $\xi\in\boldsymbol{L}(X,S)$, $\eta\in\boldsymbol{L}(Y,T)$ and $\zeta\in\boldsymbol{L}(Z,S\otimes T)$. Putting

$$(11)\quad H'(\xi)=H_1(\xi'),\ H''(\xi)=H_2(\xi''),\ \Re(\xi;\mathcal{G},\mathcal{Q})\ \text{(or simply}\ \Re(\xi))=H(\xi)+H'(\xi)-H''(\xi)$$

where $\xi'=K'\xi$ and $\xi''=K''\xi$, then these H, H', H'' and $\Re$ are bounded non-negative linear functionals over $\boldsymbol{L}(X,S)$ and $\Re$ will be called the *transmission functional* of the channel (X,ν,Y) associated with $(\mathcal{G},\mathcal{Q})$ on $(Z,\mathfrak{Z})$.

These descriptions yeild the amounts of entropies in the usual sense of the stationary memory channel (A^I,ν,B^I). That is, let $[a_0]$ ($a_0\in A_0=$the alphabet of the 0-th coordinate of A^I) be the message symbol of single element a_0 in the notation of § 2. Then, putting $\mathcal{G}=\{[a_0];\ a_0\in A_0\}$, $\mathcal{G}$ is a clopen partition of A^I and the field $\mathfrak{A}_0$ of cylinder sets in A^I is generated by the sequence of clopen partitions $\{S^n\mathcal{G};\ n=0,\pm1,\cdots\}$, and similarly putting $\mathcal{Q}=\{[b_0];\ b_0\in B_0\}$, the field $\mathfrak{B}_0$ is generated by $\{T^n\mathcal{Q};\ n=0,\pm1,\cdots\}$. Therefore the entropy $H(p)$ of the input stationary source p (on A^I) equals to $H(p,\mathcal{G},S)$ in the present sense, and similarly it holds for the entropy of the output source q (on B^I) and for the transmission rate $\Re(p)$ of the input source p.

Now we go back to the present case. Putting

$$(12)\qquad h'(x)=\int_Y h_1(y)\nu(x,dy)\quad\text{and}\quad h''(x)=\int_Y h_2(x,y)\nu(x,dy),$$

they are bounded Borel measurable, and by (4)

$$\int_X h'(x)d\xi(x)=\int_Y h_1(y)d\xi'(y)\quad\text{and}\quad\int_X h''(x)\,d\xi(x)=\int_Z h_2(z)d\xi''(z).$$

Then we obtain the integral representation theorem[2] of Parthasarathy's type (cf. [11]):

2) Somewhat after than Parthasarathy, the integral representation of the transmission rate in memory channel has been also proved by Jacobs [9] in which his proof was done by the similar manner under Krylof-Bogokiouboff's theorem. In the preceding paper [16], we have applied their method to a general case which will yield the present case.

THEOREM 2. *The transmission functional* $\mathfrak{R}(\cdot;\ \mathfrak{F}, \mathfrak{G})$ *is bounded non-negative and linear over the Banach space* $\boldsymbol{L}(X, S)$ *and there exists universally an* S*-invariant bounded Borel measurable function* $\mathfrak{r}(x)$ *on* X *such that*

$$\mathfrak{R}(\xi;\ \mathfrak{F}, \mathfrak{G})=\int_X \mathfrak{r}(x)d\xi(x) \qquad \textit{for every} \quad \xi\in\boldsymbol{L}(X, S), \tag{13}$$

where the function $\mathfrak{r}(x)$ *is defined by* $\mathfrak{r}(x)=h(x)+h'(x)-h''(x)$.

Proof. The S-invariance of h'' follows from $h_2(Sx, y)=h_2(x, T^{-1}y)$ and

$$h''(Sx)=\int_Y h_2(Sx, y)\nu(Sx, dy)=\int_Y h_2(x, y)\nu(x, dy)=h''(x)$$

and similarly for $h'(x)$. It remains only to show the non-negative definiteness of $\mathfrak{R}(\cdot;\ \mathfrak{F}, \mathfrak{G})$:

Put the finite partitions $\mathfrak{F}_n^\circ=\{U_1, \cdots, U_{n_1}\}$, $\mathfrak{G}_n^\circ=\{V_1, \cdots, V_{n_2}\}$ and $\mathcal{K}_n^\circ(=\mathfrak{F}_n^\circ\otimes\mathfrak{G}_n^\circ)$ $=\{U_i\times V_j;\ U_i\in\mathfrak{F}_n^\circ,\ V_j\in\mathfrak{G}_n^\circ\}$ (cf. §3). Whence for any fixed $p\in\boldsymbol{P}(X, S)$

$$\begin{aligned}\textstyle\sum_{i,j}p''(U_i\times V_i)\log p\otimes p'(U_i\times V_j)&=\textstyle\sum_{i,j}p''(U_i\times V_j)\,[\log p(U_i)+\log p'(V_i)]\\ &=\textstyle\sum_i p(U_i)\log p(U_i)+\sum_j p'(V_j)\log p'(V_j)\\ &=\textstyle\sum_{i,j}p\otimes p'(U_i\times V_j)\,[\log p(U_i)+\log p'(V_j)]\\ &=\textstyle\sum_{i,j}p\otimes p'(U_i\times V_j)\log p\otimes p'(U_i\times V_j)\end{aligned}$$

where in these equalities the indeces i and j with $p(U_i)=p'(V_j)=0$ are neglected, and by the linearity of $H(\cdot, \mathcal{K}, S\otimes T)$

$$\begin{aligned}&\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G})=H([p\otimes p'-p''], \mathcal{K}, S\otimes T)\\ &=-\lim_{n\to\infty}\frac{1}{n}\textstyle\sum_{i,j}p''(U_i\times V_j)\,[\log p\otimes p'(U_i\times V_j)-\log p''(U_i\times V_j)]\end{aligned}$$

and this is non-negative, because, for every $a_{ij}>0$ and $b_{ij}>0$ with $\sum a_{ij}=\sum b_{ij}$, $\sum a_{ij}[\log a_{ij}-\log b_{ij}]\geqq 0$. For general $\xi\in\boldsymbol{L}^+(X, S)$, it reduces to the above.

Now, we discuss the exact form of Parthasarathy. *Let* $\mathfrak{B}_X$ *be the* σ*-subfield*[3] *generated by* $\{S^n\mathfrak{F};\ n=0, \pm1, \pm2, \cdots\}$ and let $\boldsymbol{P}'_S$ be the set of all S-invariant probability regular measures over $(X, \mathfrak{B}_X)$, where the regularity is defined within open or closed sets belonging to $\mathfrak{B}_X$. Then $\boldsymbol{P}'_S$ consists of the measures p_1, restriction of $p\in\boldsymbol{P}(X, S)$ over $(X, \mathfrak{B}_X)$, $p_1=(p|\mathfrak{B}_X)$ say, and if $p_1\in\boldsymbol{P}'_S$ is ergodic over $(X, \mathfrak{B}_X)$ relative

3) $\mathfrak{B}_X$ is a proper σ-subfield of $\mathfrak{X}$, because X is not assumed to be separable.

to S then $p'=(p|\mathfrak{B}_X)$ for certain ergodic $p\in\boldsymbol{P}(X, S)$.[4] Whence by the definition of the entropy functional $H(\cdot)$, $H(p)=H(\rho)$ holds for p, $\rho\in\boldsymbol{P}(X, S)$ within $(p|\mathfrak{B}_X)=(\rho|\mathfrak{B}_X)$, and hence $H(p_1)$ $(p_1\in P'_S)$ is defined by $H(p)$ for $p\in\boldsymbol{P}(X, S)$, $p_1=(p|\mathfrak{B}_X)$. In particular, since m_r $(r\in R)$ are ergodic over $(X, \mathfrak{B}_X)$ (where m_r and R are defined by §4 in [16]), there corresponds an ergodic $\bar{m}_r\in\boldsymbol{P}(X, S)$ such as $m_r=(\bar{m}_r|\mathfrak{B}_X)$, and hence $\mathfrak{R}(m_r)$ $(=\mathfrak{R}(m_r;\ \mathfrak{F}, \mathfrak{G}))$ is defined by the amount $\mathfrak{R}(\bar{m}_r)$. For the partitions $\mathfrak{G}$ of Y and $\mathcal{K}$ $=\mathfrak{F}\otimes\mathfrak{G}$ of Z, similarly *define the σ-fields $\mathfrak{B}_Y$ (over Y) and $\mathfrak{B}_Z$ (over Z).* Then we obtain the theorem under the condition (C 1′) which contains (m 1) as a special case:

(C 1′) *For each fixed $V\in\mathfrak{B}_Y$, $\nu(\cdot V)$ is measurable over $(X, \mathfrak{B}_X)$.*

THEOREM 3. *Under the additional condition* (C 1′), *the function $\mathfrak{R}(m_r)$ of $r\in R$ is $\mathfrak{B}_X$-measurable on R and satisfies*

$$\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G})=\int_R \mathfrak{R}(m_r;\ \mathfrak{F}, \mathfrak{G})\, dp(r) \qquad \textit{for every}\quad p\in\boldsymbol{P}(X, S). \tag{14}$$

Proof. Since the functions $h_1(y)$ and $h_2(z)$ are $\mathfrak{B}_Y$ and $\mathfrak{B}_Z$-measurable (cf. [16], Theorem 5 and its proof), the condition (C 1′) implies $\mathfrak{B}_X$-measurability of $h'(x)$ and $h''(x)$, and hence $\mathfrak{r}(x)$ is so. Again by [16] (§4, Lemma 3), (14) follows from

$$\int_X \mathfrak{r}(x)\, dp(x)=\int_R\int_X \mathfrak{r}(x)\, d\bar{m}_r(x)\, dp(r) =\int_R\int_X \mathfrak{r}(x)\, dm_r(x)\, dp(r).$$

While, another integral representation of the transmission functional is given by means of the function $h_2(z)$ on Z:

THEOREM 4. *For every $p\in\boldsymbol{P}(X, S)$, putting $\mu_p=p\otimes p'-p''$, then it is a signed measure $\mu_p\in\boldsymbol{L}(Z, S\otimes T)$ and satisfies*

$$\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G})=\int_Z h_2(z)\, d\mu_p(z) \tag{15}$$

where $h_2(z)=h(z, \mathcal{K}, S\otimes T)$, $\mathcal{K}=\mathfrak{F}\otimes\mathfrak{G}$, is the universal entropy function over Z associated with $\mathfrak{F}\otimes\mathfrak{G}$ and $S\otimes T$.

Proof. Since both $p\otimes p'$ and p'' belong to $\boldsymbol{P}(Z, S\otimes T)$ and $H_2(\cdot)=H(\cdot, \mathcal{K}, S\otimes T)$

---

4) Let $\boldsymbol{C}_{\mathfrak{F}}$ be the linear subspace of $C(X)$ generated by $\{C_U\}_U$, U being the sets belonging to the field of clopen sets $\mathfrak{U}$ generated by $\{S^n\mathfrak{F};\ n=0, \pm 1, \cdots\}$, as in §4, page 22 of [16]. Then, by Riesz theorem, $p_1\epsilon\boldsymbol{P}'_S$ is indentified with a non-negative linear functional, with norm one, over $\boldsymbol{C}_{\mathfrak{F}}$ (cf. Lemma 1, [16]), and hence, by Hahn-Banach theorem, p_1 has non-negative definite extensions over $C(X)$ preserving its norm. Putting $\boldsymbol{P}_1$ the set of all such extensions of p_1, then $\boldsymbol{P}_1$ is invariant under S. Furthermore, $\boldsymbol{P}_1$ is weakly* compact and convex, and hence, by the fixed point theorem S-invariant p exists in $\boldsymbol{P}_1$, i.e., $p_1=(p|\mathfrak{B}_X)$ for some $p\epsilon\mathbf{P}(X, S)$.

If p_1 is ergodic over $(X, \mathfrak{B}_X)$, then, putting E_1 the set of all extreme points in $\boldsymbol{P}(X, S)\frown\boldsymbol{P}_1$, every $p\epsilon E_1$ is ergodic over $(X, \mathfrak{X})$. Indeed, if there exist σ, $\rho\epsilon\boldsymbol{P}(X, S)$ such that p $=\alpha\sigma+\beta\rho$ on $C(X)$ $(\alpha, \beta\geqq 0, \alpha+\beta=1)$ and hence so on $\boldsymbol{C}_{\mathfrak{F}}$. But by the ergodicity of p_1, $p=\sigma$ $=\rho$ on $C_{\mathfrak{F}}$. Therefore σ, $\rho\epsilon\boldsymbol{P}(X, S)\frown\boldsymbol{P}_1$ and hence $p\epsilon E_1$ implies $p=\sigma=\rho$ on $C(X)$.

is linear over $\boldsymbol{L}(Z, S\otimes T)$,

$$\mathfrak{R}(p;\ \mathfrak{S},\ \mathfrak{L})=H(p,\ \mathfrak{S},\ S)+H(p',\ \mathfrak{L},\ T)-H(p'',\ \mathcal{K},\ S\otimes T)=H_2(p\otimes p')-H_2(p'')$$

$$=H_2(p\otimes p'-p'')=H_2(\mu_p)=\int_Z h_2(z)\, d\mu_p(z).$$

6. Condition of continuity of channel and its capacity.

In the previous sections 4 and 5, for the stationary channel (X, ν, Y) the conditions corresponding to the finite memory was not assumed. Now it will be introduced a condition on the channel (X, ν, Y), which contains the conditions on the finite memory.

As in §5, assume the *totally disconnectedness of X and Y with the clopen bases $\mathfrak{X}_0$ and $\mathfrak{Y}_0$*, respectively.

DEFINITION 2. The stationary channel (X, ν, Y) is called *continuous*, if the following condition (C 4) is satisfied:

(C 4) *for every fixed set $V\in\mathfrak{Y}_0$, the functions $\nu(\cdot, V)$ is continuous on X.*

In the memory channel (A^I, ν, B^I), (C 4) *corresponds to the condition* (m 4). More precisely, (A^I, ν, B^I) having only the conditions (m 2) and (m 4), satisfies always (C 4). Indeed, as stated in §2, both the spaces A^I and B^I are totally disconnected with the bases $\mathfrak{A}_0$ and $\mathfrak{B}_0$ consisting of their finite dimensional cylinder sets, respectively. Let $V=[b_m\cdots b_n]$ be a finite message in B^I, then, by (m 4), $\nu(\mathfrak{a}, V)=\nu(\mathfrak{a}', V)$ holds for every pair $\mathfrak{a}, \mathfrak{a}'\in A^I$ with $a_k=a'_k$ $(k=m-l, \cdots, n)$. Hence, for fixed $\mathfrak{a}\in A^I$, taking $U(\mathfrak{a})=[a_{m-l}\cdots a_n]$ as a neighbouhood of $\mathfrak{a}$ in A^I, then

$$\nu(\mathfrak{a}, V)-\nu(\mathfrak{a}', V)=0 \qquad \text{for every} \quad \mathfrak{a}'\in U(\mathfrak{a}). \tag{16}$$

Furthermore, since every cylinder set $V\in\mathfrak{B}_0$ is expressed by finite union of disjoint finite messages in B^I, by (m 2), we can find a neighbourhood $U(\mathfrak{a})$ of $\mathfrak{a}$ in A^I such that the equality (16) holds for the V. Therefore (m 2) and (m 4) imply (C 4), and in particular, *in every stationary finite memory channel, the channel distribution $\nu(\mathfrak{x}, V)$ is continuous of $\mathfrak{x}\in A^I$ for each fixed finite message $V=[b_m\cdots b_n]$.*

In general, relative to the associated transformations K' and K'' the following is proved.

THEOREM 5. *If the stationary channel (X, ν, Y) is continuous, then the channel transformations K' and K'' are continuous with respect to the weak* topologies on $\boldsymbol{L}(X)$, $\boldsymbol{L}(Y)$ and $\boldsymbol{L}(Z)$, that is, if $\xi_\alpha\to\xi$ in $\boldsymbol{L}(X)$ and uniformly bounded: $\|\xi_\alpha\|_1\leqq M$, then $K'\xi_\alpha\to K'\xi$ in $\boldsymbol{L}(Y)$ and $K''\xi_\alpha\to K''\xi$ in $\boldsymbol{L}(Z)$, where the convergences are weakly as conjugate spaces, (i.e. weak* convergences).*

Proof. For every $U\in\mathfrak{X}_0$ and $V\in\mathfrak{Y}_0$, the functions $C_U(\cdot)$ (the characteristic function of U) and $\nu(\cdot, V)$ belong to $C(X)$. Therefore, for such sets U and V, and for a net $\{\xi_\alpha\}\subset\boldsymbol{L}(X)$ weakly* converging to $\xi\in\boldsymbol{L}(X)$ with $\|\xi_\alpha\|_1\leqq M$, the following holds:

$$(K''\xi_\alpha)(U\times V)=\xi_\alpha''(U\times V)=\int_X C_U(x)\nu(x, V)d\xi_\alpha(x)=<C_U\nu(\cdot, V), \xi_\alpha>$$

$$\to<C_U\nu(\cdot, V), \xi>=\int_X C_U(x)\nu(x, V)\,d\xi(x)=\xi''(U\times V)=(K''\xi)(U\times V)$$

where $<, >$ is defined in (1′), and hence

(17) $$<C_W, K''\xi_v>\to<C_W, K''\xi> \quad \text{for every} \quad W\in\mathfrak{Z}_0\ (=\mathfrak{X}_0\otimes\mathfrak{Y}_0).$$

Since, by Theorem 1, $||K''\xi_\alpha||_1\leqq||\xi_\alpha||_1\leqq M$ for all indeces α and since the set of linear combinations of C_W, $W\in\mathfrak{Z}_0$, is uniformly dense in $C(Z)$, (17) implies that $\xi_\alpha''=K''\xi_\alpha$ converges weakly* to $\xi''=K''\xi$, that is, $<f, \xi_\alpha''>\to<f, \xi''>$ for every $f\in C(Z)$. The continuity of K' follows from that of K''.

By this proof, it follows under the same assumptions on (X, ν, Y)

COROLLARY 5.1. *K' and K'' arc sequentially weakly* continuous over the respective Banach spaces $\boldsymbol{L}(X)$, $\boldsymbol{L}(Y)$ and $\boldsymbol{L}(Z)$.*

The concept of the stationary capacity, as defined in Feinstein [6], of the finite memory channel can be similarly defined for the channel (X, ν, Y):

DEFINITION 3. For a pair of finite partitions $\mathfrak{F}$ of X and $\mathfrak{G}$ of Y, put

(18) $$C_s(\mathfrak{F}, \mathfrak{G}) \text{ (or simply } C_s)=\sup\{\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G});\ p\in\boldsymbol{P}(X, S)\}$$

and it is called the *stationary capacity*, relative to $(\mathfrak{F}, \mathfrak{G})$, of the channel (X, ν, Y).

The existence of the stationary capacity $C_s(\mathfrak{F}, \mathfrak{G})$ will be proved as the following:

THEOREM 6. *If the channel distuibution $\nu(\cdot, \cdot)$ satisfies either* (C 1′) *or* (C 4), *then,*

(19) $$C_s(\mathfrak{F}, \mathfrak{G})=\sup\{\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G});\ p\in\boldsymbol{P}_e(X)\}$$

where $\boldsymbol{P}_e(X)$ is the set of all ergodic measures on X. Particularly, under (C 4), *the capacity C_s can be achieved on $\boldsymbol{P}(X, S)$.*

Proof. Under the condition (C 1′), the equality (19) follows from Theorem 3. While, when (C 4) is satisfied, by Theorem 5, the associated channel transformations K' and K'' are weakly* continuous on $\boldsymbol{P}(X, S)$. Hence, by the inverse mapping of K', every weakly* open subset in $\boldsymbol{P}(Y, T)$ is mapped onto an open subset in $\boldsymbol{P}(X, S)$, and similarly so for K''. Besides, by Theorem 4 in [16], the functional $H(\cdot)$, $H_1(\cdot)$ and $H_2(\cdot)$ are weakly* upper-semicontinuous over $\boldsymbol{P}(X, S)$, $\boldsymbol{P}(Y, T)$ and $\boldsymbol{P}(Z, S\otimes T)$, respectively. Therefore, combining with the fact just above, $H'(p)=H_1(K'p)$ and $H''(p)=H_2(K'p)$ are weakly* upper-semicontinuous on $\boldsymbol{P}(X, S)$, and so is $\mathfrak{R}(p)$ $(=\mathfrak{R}(p;\ \mathfrak{F}, \mathfrak{G}))$. Consequently, by the weak* compactness of $\boldsymbol{P}(X, S)$, $C_s=\mathfrak{R}(p)$ for some $p\in\mathbf{P}(X, S)$. (19) and the achieving of C_s follow immediately from $\boldsymbol{P}(X, S)$ being weakly* convex closure of $\boldsymbol{P}_e(X)$ and the upper-semicontinuity of $\mathfrak{R}(p)$.

It should be noted, that Breiman's theorem (cf. Theorem 1 of [2]) relative to upper-semicontinuous functional implies that *there exists at least one ergodic $p\in\boldsymbol{P}_e(X)$ such that*

$$C_s(\mathfrak{I}, \mathfrak{L})=\Re(p; \mathfrak{I}, \mathfrak{L}).$$

According to Feinstein [6], an ergodic measure $p\in \boldsymbol{P}_e(X)$ is called *admissible* if $p''=K''p\in \boldsymbol{P}(Z, S\otimes T)$ is also ergodic, and denote $\boldsymbol{P}_A(X)$ the set of all admissible $p\in \boldsymbol{P}_e(X)$. Put

$$C_e(\mathfrak{I}, \mathfrak{L})= \sup\{\Re(p; \mathfrak{I}, \mathfrak{L}); p\in \mathbf{P}_A(X)\}$$

and it is called ***ergodic capacity***, relative to $(\mathfrak{I}, \mathfrak{L})$, of the channel (X, ν, Y). Whence $C_e(\mathfrak{I}, \mathfrak{L})\leqq C_s(\mathfrak{I}, \mathfrak{L})$. Besides, there are channels (A^I, ν, B^I) (satisfying (m 1)$\sim$(m 4) but not (m 5)) without having nonempty $\boldsymbol{P}_A(X)$, and for such channels the ergodic capacity is undefined, cf. Feinstein [6]. However it holds that, for every finite memory channel (A^I, ν, B^I) satisfying (m 1)$\sim$(m 5), $\boldsymbol{P}_A(X)=\boldsymbol{P}_e(X)$, cf. Takano [12], and hence $C_e(\mathfrak{I}, \mathfrak{L})=C_s(\mathfrak{I}, \mathfrak{L})$ hold. To describe this for the present stationary channel (X, ν, Y), according to Adler [1], we introduce the following: If the stationary channel (X, ν, Y) satisfies

$$\text{(C 5)} \qquad \lim_{n\to\infty} [\nu(x, T^n V_1\frown V_2)-\nu(x, T^n V_1)\cdot\nu(x, V_2)]=0, \qquad x\in X$$

for every $V_1 V_2\in\mathfrak{Y}_0$, then it will be called *asymptotically independent*. As Adler stated, in the channel (A^I, ν, B^I), (m 5) implies (C 5) for $\alpha\in A^I$ and $V_1, V_2\in\mathfrak{B}_0$.

In the present case we obtain that: If (X, ν, Y) satisfies (C 1)$\sim$(C 3) and (C 5), then every ergodic $p\in\boldsymbol{P}_e(X)$ is transformed to ergodic measures $p'\in\boldsymbol{P}(Y, T)$ and $p''\in\boldsymbol{P}(Z, S\otimes T)$ by the channel transformations K' and K'', respectively, and hence if (X, ν, Y) satisfies (C 1)$\sim$(C 5) then $\boldsymbol{P}_A(X)=\boldsymbol{P}_e(X)$ and $C_e(\mathfrak{I}, \mathfrak{L})=C_s(\mathfrak{I}, \mathfrak{L})$ hold. This is given by Adler's proof without any modifications, which is done by the Halmos' well-known characteristic property of ergodicity of p:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} p(S^{-k}U_1\frown U_2)=p(U_1)\,p(U_2)$$

for every pair of $U_1, U_2\in\mathfrak{X}_0$.

References

[1] Adler, R. L., Ergodic and mixing properties of infinite memory channels. Proc. Amer. Math. Soc. **12** (1961), 924–930.

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

[3] Carleson, L., Basic theorems of information theory. Math. Scand. **6** (1958), 175–180.

[4] Doob, J. L., Stochastic processes. New York. John Wiley (1953).

[5] Echigo, M., and M. Nakamura, A remark of the concept of channels. Proc. Japan Acad. **38** (1962), 307–309.

[6] Feinstein, A., Foundations of information theory. New York, McGraw-Hill Book Co., Inc. (1958).

[7] Feinstein, A., On the coding theorem and its converse for finite memory channels. Information and Control **2** (1959), 25–44.

[8] HALMOS, P. R., Entropy in ergodic theory (Mimeographed lecture notes). University of Chicago (1959).

[9] JACOBS, K., Über die Struktur der mittleren Entropie. Math. Zeitschr. **78** (1962) 33–43.

[10] KHINCHIN, A. I., On the fundamental theorems of information theory. Uspehı Mat. Nauk **11** (1956), 17–75. (in Russian)

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

[12] TAKANO, K., On the basic theorems of information theory. Ann. Inst. Statist. Math. Tokyo **9** (1958), 53–77.

[13] TSAREGRADSKY, I. P., On the capacity of a stationary channel with finite memory. Theory of Probability and its Applications **3** (1958) 79–91.

[14] UMEGAKI, H., Entropy functionals ın stationary channels. Proc. Japan Acad. **38** (1962) 668–672.

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

[16] UMEGAKI, H., General treatment of alphabet-message space and integral representation of entropy. Kōdai Math. Sem. Rep. **16** (1964), 18–26.

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.

Added in proof. After this paper was receıved, the author is informed by a letter of Professor M. Nakamura, that a similar abstract characterization of finite memory channels is also obtained and prepared to discuss ın his lecture at the Osaka Gakugei Daigaku.