147. *Entropy Functionals in Stationary Channels*

By Hisaharu UMEGAKI
Department of Mathematics, Tokyo Institute of Technology
(Comm. by K. KUNUGI, M.J.A., Nov. 12, 1962)

1. The purpose of this note is to introduce a functional analysistic method into the amount of the information on stationary memory channels which contain finite or infinite memory channels (cf. Feinstein [4]). By the method described in this note, the notational complications to introducing the concepts of the entropy, the transmission rate and the stationary or ergodic capacity in a memory channel can be avoided and they will be given by a simple measure theoretic functional forms.

It will be stated that, in § 2, the amount of the entropy of every probability distribution on an information source determins uniquely a bounded stationary linear functional of the subconjugate space of the Banach space of bounded random variables over the source, this is an extension of Breiman's Theorem [1], and that, in § 3, the channel distribution determins a bounded linear transformation between the subconjugate spaces of the Banach spaces of bounded random variables over the input space and the compound space. This treatment of the channel distribution as a linear transformation was previously introduced by Echigo-Nakamura [3] in the case of a kind of continuous channel without memory. In the final part, it will be stated an integral representation theorem of the transmission rate of a probability distribution of the input space. This is a generalization of Parthasarathy's Theorem [7].

2. Let $(X, \mathfrak{X})$ be a measurable space with a measurable transformation T from X onto itself. Assume that there exists a finite subfield $\mathfrak{X}_0$ of $\mathfrak{X}$ such that

$$(1) \qquad \bigvee_{k=-\infty}^{\infty} T^{-k}\mathfrak{X}_0 = \mathfrak{X}$$

where $T^{-k}\mathfrak{X}_0 = \{T^{-k}U; u \in \mathfrak{X}_0\}$ and $\vee_{k\in N} T^{-k}\mathfrak{X}_0$ (N being a set of integers) is the smallest σ-subfield containing the fields $T^{-k}\mathfrak{X}_0$, $k \in N$. The measurable space $(X, \mathfrak{X})$ defined here contains as a special case the input message space of a memory channel, that is the case, X is infinite product space of alphabet, $\mathfrak{X}$ is σ-field generated by all cylinder sets and T is the shift transformation.

Denote $P_s(X)$ the set of all T-stationary probability measures over $(X, \mathfrak{X})$. Then $P_s(X)$ is nonempty and is dominated by a measure $\mu \in P_s(X)$, that is, every $p \in P_s(X)$ is absolutely continuous with respect to μ, say, $p \ll \mu$ for every $p \in P_s(X)$. Let $M(X)$ be the linear space of all bounded complex measurable functions over $(X, \mathfrak{X})$ and $M_s(X)$ be the linear subspace, of $M(X)$, of all T-invariant functions in $M(X)$. Let $L(X)$ be the linear space of all bounded signed measures ξ with $\xi \ll \mu$ and $L_s(X)$ be the space of all T-stationary $\xi \in L(X)$. Then the spaces $M(X)$, $M_s(X)$ and $L(X)$, $L_s(X)$ are Banach spaces with the norms

$$\|f\|_\infty = \mu\text{-}ess.\ sup\ |f(x)| \qquad \text{for } f \in M(X) \text{ or } f \in M_s(X)$$

and

$$\|\xi\|_1 = total\ variation\ of\ \xi \quad \text{for } \xi \in L(X) \text{ or } \xi \in L_s(X),$$

respectively. Then it is not hard to show that $M(X)$ is the conjugate Banach space of $L(X)$, or equivalently $L(X)$ is the subconjugate space of $M(X)$.

Denote

$$(2) \qquad \mathfrak{X}_\infty = \bigvee_{k=1}^{\infty} T^{-k}\mathfrak{X}_0 \quad \text{and} \quad \mathfrak{U} = \bigcup_{n=1}^{\infty} \bigvee_{k=-n}^{n} T^{-k}\mathfrak{X}_0$$

Where $\bigcup$ means the set theoretic union. The family $\mathfrak{U}$ is a subfield but not necessarily σ-subfield of $\mathfrak{X}$. For any probability measure $p \ll \mu$ and for any fixed σ-subfield $\mathfrak{B}$ of $\mathfrak{X}$, denote $P_p(U|\mathfrak{B})$ the conditional probability of a set $U \in \mathfrak{X}$ conditioned by $\mathfrak{B}$. For any positive bounded measure $\xi \in L(X)$ over $(X, \mathfrak{X})$, say $\xi \in L^+(X)$, $\xi_1 = \xi/\|\xi\|_1$ is a probability measure and put

$$P_\xi(U|\mathfrak{B}) = P_{\xi_1}(U|\mathfrak{B}).$$

For the σ-subfield $\mathfrak{B} = \mathfrak{X}_\infty$, define a functional

$$H(\xi) = -\sum_{U \in \mathfrak{X}_0} \int_X P_\xi(U|\mathfrak{B}) \log P_\xi(U|\mathfrak{B}) d\xi(x)$$

for every $\xi \in L_s^+(X)$, where $\sum_{U \in \mathfrak{X}_0}$ means that the summation on U running over all atoms in $\mathfrak{X}_0$. When $(X, \mathfrak{X})$ is input message space of a memory channel, and p is a stationary probability measure, i.e., $p \in P_s(X)$, the amount $H(p)$ coincides with the amount of the entropy of p (cf. Halmos [5]).

Besides, any measure $\xi \in L_s(X)$ is uniquely expressed such as

$$\xi = \xi^{(1)} - \xi^{(2)} + i\xi^{(3)} - i\xi^{(4)} \quad (i = \sqrt{-1})$$

for $\xi^{(k)} \in L_s^+(X)$ $(k = 1, 2, 3, 4)$ and the domain of the functional $H(\cdot)$ is extended over the full space $L_s(X)$ by

$$H(\xi) = H(\xi^{(1)}) - H(\xi^{(2)}) + iH(\xi^{(3)}) - iH(\xi^{(4)}).$$

The functional $H(\cdot)$ is well-defined over $L_s(X)$ and it will be called by *entropy functional* of the measurable space $(X, \mathfrak{X})$. A functional $F(\cdot)$ over $L(X)$ is *T-stationary*, if $F(\xi) = F(T\xi)$ for all $\xi \in L(X)$, where $(T\xi)(U) = \xi(T^{-1}U)$, $U \in \mathfrak{X}$. The following is an extension of the Breiman's convex linear form [1].

THEOREM 1. *The entropy functional $H(\cdot)$ is a bounded linear functional over the Banach space $L_s(X)$ and it is uniquely extended* *to a T-stationary bounded linear functional over the Banach space $L(X)$.*

Since $M(X)$ is the conjugate space of $L(X)$, by Riesz-Markov-Kakutani's Theorem, the following is obtained:

THEOREM 2. *There exists uniquely, within μ a.e., a bounded T-invariant non-negative measurable function $h(x)\in M(X)$ such that*

$$H(\xi)=\int_X h(x)d\xi(x) \text{ for every } \xi\in L(X). \tag{3}$$

This contains the integral representation theorem of Parthasarathy [7].

3. Now we wish to apply Theorem 1 for the channel, which contains every discrete memory channel (cf. Feinstein [4]) as a special case. It will be introduced a concept of stationary channel (X,ν,Y): The input space $(X,\mathfrak{X})$ is the measurable space with T and $\mathfrak{X}_0$ given in § 2, and the output space $(Y,\mathfrak{Y})$ is a measurable space having a measurable transformation S and finite subfield $\mathfrak{Y}_0$ with the corresponding property (1) to the case of $(X,\mathfrak{X},T,\mathfrak{X}_0)$. Furthermore let $\nu(V,x)$ be a function defined over the product family $\mathfrak{Y}\times X$ such that

(i) *For each fixed $V\in\mathfrak{Y}$, $\nu(V,x)$ is a measurable functions on $(X,\mathfrak{X})$,*

(ii) *For each fixed $x\in X$, $\nu(V,x)$ is a probability measure over $(Y,\mathfrak{Y})$*

and

(iii) *$\nu(V,x)$ is stationary, i.e., $\nu(SV,Tx)=\nu(V,x)$ for every $x\in X$ and $V\in\mathfrak{Y}$.*

Whence we call the triple (X,ν,Y) being a *stationary channel.*

Denote $(Z,\mathfrak{Z})$ the product measurable space $(X\times Y,\ \mathfrak{X}\otimes\mathfrak{Y})$, and R the measurable transformation over $(Z,\mathfrak{Z})$ defined by $T\otimes S$, the product transformation of T and S, and furthermore put the product σ-subfields $\mathfrak{Z}_0=\mathfrak{X}_0\otimes\mathfrak{Y}_0$, and $\mathfrak{Z}_\infty=\mathfrak{X}_\infty\otimes\mathfrak{Y}_\infty$. Then the condition (1) satisfies for $(Z,\mathfrak{Z},R,\mathfrak{Z}_0)$. For any $\xi\in L(X)$, putting

$$\xi''(U,V)=\int_U \nu(V,x)d\xi(x) \qquad \text{for } U\in\mathfrak{X} \text{ and } V\in\mathfrak{Y}, \tag{4}$$

ξ'' is uniquely extended to a bounded signed measure over $(Z,\mathfrak{Z})$ (denote it by the same symbol ξ''), and putting

$$\xi'(V)=\int_X \nu(V,x)d\xi(x) \qquad \text{for } V\in\mathfrak{Y}, \tag{5}$$

ξ' is a bounded signed measure over $(Y,\mathfrak{Y})$.

Besides we set up a condition

(i′) *For each fixed $V\in\mathfrak{Y}$, $\nu(V,x)$ is measurable with respect to the subfield $\mathfrak{U}$ defined in* (2).

When the stationary channel (X,ν,Y) satisfies the condition (i′), it will be called *properly finite.* Such a channel contains, as a special case, every finite memory channel in the sense of Khinchin [6] and Takano [8].

Let $L(Y)$, $L(Z)$ and $L_s(Y)$, $L_s(Z)$ be the Banach spaces defined as the $L(X)$ and $L_s(X)$ respectively. The following is a general treatment of Echigo-Nakamura's Theorem [3].

THEOREM 3. *In a stationary channel (X, ν, Y) the mappings $\xi \to \xi'$ and $\xi \to \xi''$ defined by* (4) *and* (5) *are bounded non-negative linear transformations from $L(X)$ (from $L_s(X)$) into $L(Y)$ and $L(Z)$ (into $L_s(Y)$ and $L_s(Z)$), respectively. If the channel (X, ν, Y) is properly finite, then their transformations are weakly* continuous in the sequentially weak* topologies over $L(X)$, $L(Y)$ and $L(Z)$.*

In a stationary channel (X, ν, Y), since the entropy functionals $H(\cdot)$ over $(Y, \mathfrak{Y})$ and $(Z, \mathfrak{Z})$ are defined and since $\xi' \in L(Y)$ and $\xi'' \in L(Z)$ for every $\xi \in L(X)$, it can be defined a functional

$$R(\xi) = H(\xi) + H(\xi') - H(\xi'')$$

and it will be called by *transmission functional* of the channel. Whence we obtain the following:

THEOREM 4. *The transmission functional $R(\cdot)$ is a bounded T-stationary linear functional over $L(X)$ and there exists an essentially unique bounded T-invariant measurable function $r(\cdot)$ on X such that*

$$R(\xi) = \int_X r(x)\, d\xi(x) \text{ for every } \xi \in L(X).$$

If the stationary channel (X, ν, Y) is properly finite and having an *ergodicity property*:

$$(6) \qquad \lim_{n\to\infty} [\nu(T^n V_1 \cap V_2, x) - \nu(T^n V_1, x)\nu(V_2, x)] = 0$$

for every $x \in X$ and every $V_1, V_2 \in \mathfrak{Y}$, then the channel (X, ν, Y) is *admissible* in the sence of Feinstein [4] and hence after proving some continuity properties of the entropy and the transmission functionals, it can be obtained that *the stationary capacity coincides with the ergodic capacity.* This contains the result of Breiman [1], Carleson [2] and Parthasarathy [7].

The detailed proofs of the statements in this note will be published in the Kōdai Mathematical Seminar Report with the allied topics.

References

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

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

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

[4] A. Feinstein: Foundation of Information Theory, New York, McGraw-Hill, (1958).
[5] P. R. Halmos: Entropy in Ergodic Theory, University of Chicago, (1959).
[6] A. Ia. Khintchine: Mathematical Foundations of Information Theory, Dover Publications, Inc., New York (1958).
[7] K. R. Parthasarathy: On the integral representation of the rate of transmission of a stationary channel, Ill. J. Math. **5**, 299–205 (1961).
[8] K. Takano: On the basic theorem of information theory, Ann. Inst. Statist. Math., Tokyo, **9**, 53–77 (1958).