J. Math. Soc. Japan
Vol. 27, No. 2, 1975

Ergodicity and capacity of information channels with noise sources[1)]

By Yatsuka NAKAMURA

(Received May 12, 1973)

§1. Introduction.

The information channel is defined as a sort of conditional distributions on a direct product space of an input alphabet space and an output alphabet space, which are direct product spaces of countable copies of finite sets, conditioned by a Borel field of an input alphabet space (cf. Feinstein [3] and Hinchin [5]). The channel defined in this way is the most abstract and general one. However, many actual communication channels are imagined to have noise sources. The channel of additive noise is a typical one of such cases.

In this paper, we shall clarify a relation between ergodicity of such a channel and that of its noise source, and study about the channel capacity for these channels.

§2. Preliminary.

Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces with measurable transformations S, T on X and Y respectively, and Π be a set of S-invariant probability measures on X. Assume Π to be non-empty. An element p in Π is called *the input source. The channel* (from X to Y) is a numerical function ν on $X\times\mathcal{Y}$ which satisfies the followings:

(i) for any $x\in X$, $\nu_x(\cdot)$ is a probability measure on $\mathcal{Y}$,

(ii) for any $F\in\mathcal{Y}$, $\nu.(F)$ is a measurable function on X,

and

(iii) $\nu_{Sx}(F)=\nu_x(T^{-1}F)$ for any $x\in X$ and $F\in\mathcal{Y}$.

An output source q derived from an input source p and a channel ν is defined by

$$q(F)=\int_X \nu_x(F)p(dx) \qquad (F\in\mathcal{Y})$$

and denoted by $q(\cdot)=q(\cdot\,; p, \nu)$, which is a T-invariant probability measure on Y. *A compound source* r derived from an input source p and a channel ν is defined by

---

1) This work is partially supported by the Sakkokai Foundation.

$$r(C)=\int_X \nu_x(C_x)p(dx) \qquad (C\in\mathfrak{X}\times\mathcal{Y})$$

where C_x is an x-section of C. The compound source r is an $S\times T$-invariant probability measure on a product measurable space $(X\times Y, \mathfrak{X}\times\mathcal{Y})$, and denoted by $r(\cdot)=r(\cdot\,;p,\nu)$. A channel ν is said to be *ergodic* if ergodicity of p implies ergodicity of $r(\cdot)=r(\cdot\,;p,\nu)$. A set of all ergodic input sources in Π is denoted by Π_e.

A quadruplet $(X, \mathfrak{X}, p, S)$ is called *a dynamical system* (for arbitrarily fixed p in Π). We can define *the entropy* $h_p(S)$ of the measure preserving transformation S relative to the source p by

$$h_p(S)=\sup_{\mathcal{A}}\lim_n \frac{1}{n}H(\mathcal{A}\vee S^{-1}\mathcal{A}\vee\cdots\vee S^{-n+1}\mathcal{A})$$

where $\mathcal{A}$ is a measurable finite partition of X and the joint '$\vee$' is defined by the following:

$$\mathcal{A}_1\vee\mathcal{A}_2=\{A_1\cap A_2:\ A_1\in\mathcal{A}_1,\ A_2\in\mathcal{A}_2\}$$

for any finite paritions $\mathcal{A}_1$ and $\mathcal{A}_2$, which is also a finite partition, and where $H(\mathcal{A})=-\sum_{A\in\mathcal{A}}p(A)\log p(A)$ is the entropy of finite partition $\mathcal{A}$[2)]. (These arguments are seen in [1]). When there exists a finite partition $\mathcal{A}_0$ which generates $\mathfrak{X}$ in the sense of $\vee_{i=1}^{\infty}S^{-i}\mathcal{A}_0=\mathfrak{X} \bmod p$, then by the Kolmogorov-Sinai theorem

$$h_p(S)=\lim_n \frac{1}{n}H(\mathcal{A}_0\vee S^{-1}\mathcal{A}_0\vee\cdots\vee S^{-n+1}\mathcal{A}_0)\,.$$

We say X, Y and V are *finite alphabet spaces* if for some finite sets A, B and D,

$$X=A^I\,,\qquad Y=B^I\qquad\text{and}\qquad V=D^I$$

where $I=\{0, \pm1, \pm2, \cdots\}$ and A^I, B^I and D^I are direct product measurable spaces with Borel fields $\mathfrak{X}$, $\mathcal{Y}$ and $\mathcal{V}$ respectively generated by all cylinder sets. For $x\in A^I$, $x_i\in A$ denote the i-th coordinate of x. And $[x_i^0x_{i+1}^0\cdots x_{i+k}^0]$ is a (thin) cylinder set, i.e.,

$$[x_i^0x_{i+1}^0\cdots x_{i+k}^0]=\{x\in A^I:\ x_i=x_i^0,\ \cdots,\ x_{i+k}=x_{i+k}^0\}\,.$$

These notations are also adopted to both Y and V. We choose the shift operators for the transformations S, T and P in this case, i.e.,

$$(Sx)_i=x_{i+1}\,,\qquad (Ty)_i=y_{i+1}\qquad\text{and}\qquad (Pv)_i=v_{i+1}\,.$$

The time 0 partition $\mathfrak{X}_0$ of X is defined by

2) The base of the logarithm is assumed to be 2.

$$\mathfrak{X}_0 = \{[x_0] : x \in A^I\}$$

and similarly denote $\mathcal{Y}_0$ and $\mathcal{V}_0$ for Y and V.

§ 3. Channels with noise sources.

Now, let us consider another measurable space $(V, \mathcal{V})$ and a measurable transformation P acting on V. And let ϕ be a measurable mapping from the direct product measurable space $(X \times V, \mathfrak{X} \times \mathcal{V})$ to $(Y, \mathcal{Y})$, which satisfies

$$\text{(I)} \qquad \phi(Sx, Pv) = T\phi(x, v) \qquad \text{for all} \quad x \in X \text{ and } v \in V .$$

Then for any P-invariant probability measure $s(\cdot)$ on V (let us call it *a noise source*), writing

$$\nu_x(F) = \int_V \chi_F(\phi(x, v)) s(dv) , \qquad (x \in X, \ F \in \mathcal{Y})$$

we see that ν is a channel from X to Y. Since $\chi_F(\phi(x, v))$ is a jointly measurable function on $X \times V$, by the Fubini theorem, for each fixed $F \in \mathcal{Y}$ the function $\nu_x(F)$ of $x \in X$ is measurable on X. And it is clear that for each fixed $x \in X$ $\nu_x(\cdot)$ is a probability measure on $(Y, \mathcal{Y})$. Moreover, the formulae

$$\begin{aligned}
\nu_x(T^{-1}F) &= \int_V \chi_{T^{-1}F}(\phi(x, v)) s(dv) = \int_V \chi_F(T\phi(x, v)) s(dv) \\
&= \int_V \chi_F(\phi(Sx, Pv)) s(dv) = \int_V \chi_F(\phi(Sx, Pv)) s(P^{-1}dv) \\
&= \int_V \chi_F(\phi(Sx, v)) s(dv) = \nu_{Sx}(F)
\end{aligned}$$

imply that ν satisfies (iii).

Such a channel ν will be called an *integration channel* determined by the pair (ϕ, s). For ergodicity of such channels, let us give the following proposition.

PROPOSITION 1. *If a direct product measure $p \times s$ is ergodic for every $p \in \Pi_e$, then the integration channel ν determined by (ϕ, s) is ergodic.*

PROOF. Let $C \in \mathfrak{X} \times \mathcal{Y}$ be an $S \times T$-invariant set, i.e., $(S \times T)^{-1}C = C$. Then

$$\begin{aligned}
T^{-1}C_{Sx} &= T^{-1}\{y : (Sx, y) \in C\} = \{y : (Sx, Ty) \in C\} \\
&= \{y : (x, y) \in C\} = C_x .
\end{aligned}$$

And so, putting $f(x, v) = \chi_{C_x}(\phi(x, v))$, we get

$$\begin{aligned}
f(Sx, Pv) &= \chi_{C_{Sx}}(\phi(Sx, Pv)) = \chi_{C_{Sx}}(T\phi(x, v)) \\
&= \chi_{T^{-1}C_{Sx}}(\phi(x, v)) = \chi_{C_x}(\phi(x, v)) = f(x, v) ,
\end{aligned}$$

and the invariant function f takes values 0 or 1. Hence

$$r(C)=\int_X \nu_x(C_x)p(dx)=\int_X\int_V \chi_{C_x}(\phi(x, v))s(dv)p(dx)$$

$$=\int_{X\times V} f(x, v)p\times s(dx, dv)=0 \quad \text{or} \quad 1,$$

which shows that ν is ergodic. Q. E. D.

The above proposition shows that if the noise source s is weakly mixing then the channel ν is ergodic (cf. [4] p. 39). And by the proof we know that ergodicity of $p\times s$ implies ergodicity of r.

Next, let us assume additionally the following conditions for the function $\phi(x, v)$:

(II) for any $x\in X$, $\phi(x, v)=\phi(x, v')$ implies $v=v'$,

(III) putting $\lambda(x, v)=(x, \phi(x, v))$, $H\in\mathfrak{X}\times\mathcal{V}$ implies

$$\lambda(H)=\{\lambda(x, v)\,;\,(x, v)\in H\}\in\mathfrak{X}\times\mathcal{Y}.$$

THEOREM 1. *Let ϕ be a function satisfying* (I), (II) *and* (III). *Then an integration channel determined by (ϕ, s) is ergodic if and only if $p\times s$ is ergodic on $X\times V$ for all $p\in\Pi_e$.*

PROOF. Sufficiency is clear by Proposition 1. Let λ be a mapping defined in (III). Since λ is one-to-one from $X\times V$ into $X\times Y$, $\lambda^{-1}\lambda(H)=H$ for all $H\in\mathfrak{X}\times\mathcal{V}$. Now we assume that $H\in\mathfrak{X}\times\mathcal{V}$ is an $S\times P$-invariant set, then

$$\lambda(H)=\lambda((S\times P)^{-1}H)=\{\lambda(x, v):\,(Sx, Pv)\in H\}$$

$$=\{(x, \phi(x, v)):\,(Sx, Pv)\in H\}$$

$$=\{(x, y):\,y=\phi(x, v) \text{ and } (Sx, Pv)\in H\}$$

$$\subseteqq\{(x, y):\,Ty=\phi(Sx, Pv) \text{ and } (Sx, Pv)\in H\}$$

$$\subseteqq\{(x, y):\,(Sx, Ty)\in\lambda(H)\}=(S\times T)^{-1}\lambda(H).$$

Hence, the ergodicity of r implies

$$p\times s(H)=p\times s(\lambda^{-1}\lambda(H))=\int_V\int_X \chi_{\lambda^{-1}\lambda(H)}(x, v)p(dx)s(dv)$$

$$=\int_V\int_X \chi_{\lambda(H)}(\lambda(x, v))p(dx)s(dv)=\int_V\int_X \chi_{\lambda(H)}(x, \phi(x, v))p(dx)s(dv)$$

$$=\int_X\int_V \chi_{\lambda(H)_x}(\phi(x, v))p(dx)s(dv)=\int_X \nu_x(\lambda(H)_x)p(dx)$$

$$=r(\lambda(H))=0 \quad \text{or} \quad 1,$$

which shows that $p\times s$ is ergodic. Q. E. D.

The following corollaries are immediate.

COROLLARY 1. *The compound source $r=r(\cdot\,;p, \nu)$ is ergodic if and only if $p\times s$ is ergodic.*

COROLLARY 2. *If* $(X, \mathfrak{X})=(V, \mathcal{V})$, $S=P$ *and* Π_e *is a set of all S-invariant ergodic measure on* X, *then a channel determined by* (ϕ, s), *where* ϕ *satisfies* (I), (II) *and* (III), *is ergodic if and only if* s *is weakly mixing on* V.

Indeed, since the direct product measure $s\times s$ on $X\times X$ is ergodic if and only if s is weakly mixing ([**4**]), we get the result.

COROLLARY 3. *Let* $(X, \mathfrak{X})=(Y, \mathcal{Y})=(V, \mathcal{V})$ *is a measurable group with a group operation* $\cdot$ *commuting with* $S=T=P$, *and let* $y=\phi(x, v)=x\cdot v$, *then the integration channel determined by* (ϕ, s) *is ergodic if and only if* s *is weakly mixing.*

If X, Y, V are complete separable metric spaces and $\mathfrak{X}, \mathcal{Y}, \mathcal{V}$ are Borel fields on them, then the Kuratowski theorem (cf. [**6**]) permits us to omit the condition (III). Hence we get:

COROLLARY 4. *Let* $X=A^I$, $Y=B^I$ *and* $V=D^I$, *where* $A=\{0, 1, 2, \cdots, l-1\}$, $D=\{0, 1, 2, \cdots, m-1\}$ *and* $B=\{0, 1, 2, \cdots, l+m-2\}$. *Let* ϕ_a *be* $\phi_a(i, j)=i+j$. *Then we can construct the integration channel determined by* (ϕ, s) *where* ϕ *is defined by* $\phi(x, v)_i=\phi_a(x_i, v_i)$. *This channel is ergodic if and only if* $p\times s$ *is ergodic for all* $p\in \Pi_e$.

Let us call the channel obtained in Corollary 4, *a channel of additive noise.*

Next, we shall characterize the integration channel when the function ϕ is given. It can be proved that $\phi(x, F)=\{\phi(x, v): v\in F\}\in \mathcal{Y}$ for all $x\in X$ and $F\in \mathcal{V}$ assuming (I), (II) and (III). Because the image of $X\times F$ under the function λ is $\mathfrak{X}\times\mathcal{Y}$-measurable by the condition (III), and an x-section of the above image set $\lambda(X\times F)_x=\phi(x, F)$ is $\mathcal{Y}$-measurable.

PROPOSITION 2. *Let* ϕ *be a measurable mapping from* $X\times V$ *to* Y *satisfying* (I), (II) *and* (III). *A channel* ν *from* X *to* Y *is an integration channel determined by* (ϕ, s) *for some noise source* s, *if and only if the following conditions are satisfied;*

i) $\nu_x(\phi(x, V))=1$ *for all* $x\in X$, *and*

ii) $\nu_x(\phi(x, F))=\nu_{x'}(\phi(x', F))$ *for all* $x, x'\in X$ *and* $F\in \mathcal{V}$.

PROOF. Let ν be a channel satisfying the conditions i) and ii). Putting $s(F)=\nu_x(\phi(x, F))$, we see that it is a P-invariant probability measure on $(V, \mathcal{V})$ independent of $x\in X$. Moreover

$$\nu_x(E)=\nu_x(E\cap \phi(x, V))=\nu_x(\phi_x\phi_x^{-1}(E))=\nu_x(\phi(x, F))$$

$$=s(\phi_x^{-1}(E))=\int_V \chi_E(\phi(x, v))s(dv)$$

where $\phi_x(\cdot)=\phi(x, \cdot)$ and $F=\phi_x^{-1}(E)$. Hence ν is an integration channel.

The converse is obvious. Q.E.D.

§ 4. Capacity of some integration channels.

In this section we assume the finite alphabet spaces $X=A^I$, $Y=B^I$ and $V=D^I$. For the output source $q(\cdot)=q(\cdot : p, \nu)$ and the compound source $r(\cdot) =r(\cdot\,; p, \nu)$, the entropies $h_p(S)$, $h_q(T)$ and $h_r(S\times T)$ can be defined as in § 2. When $h_p(S)<+\infty$ and $h_q(T)<+\infty$, it is possible to define the *transmission rate* R_p by

$$R_p=h_p+h_q-h_r\,.$$

The *stationary capacity* C of a channel ν from X to Y is defined by

$$C=\sup_{p\in\Pi} R_p\,.$$

Putting $\Pi'=\{p\in\Pi_e : r(\cdot\,; p, \nu) \text{ is ergodic}\}$, the *ergodic capacity* C_e of a channel ν is defined by

$$C_e=\sup_{p\in\Pi'} R_p\,. \qquad (\text{We put } C_e=0 \text{ if } \Pi' \text{ is empty})\,.$$

Let ϕ_0 be a mapping from a direct product set $A^{m+1}\times D$ to B (m is a non-negative integer), satisfying the following condition (a):

(a) $\qquad \phi_0(a_0a_1\cdots a_m, d)=\phi_0(a_0a_1\cdots a_m, d')$ implies $d=d'$ in D.

Then we can construct the mapping $\hat{\phi}$ from $X\times Y$ to Y by

$$\hat{\phi}(x, v)_i=\phi_0(x_{i-m}x_{i-m+1}\cdots x_i, v_i)\,.$$

Clearly $\hat{\phi}$ is a measurable mapping from $X\times V$ to Y and

$$\begin{aligned}\hat{\phi}(Sx, Pv)_i&=\phi_0((Sx)_{i-m}\cdots(Sx)_i, (Pv)_i)\\ &=\phi_0(x_{i-m+1}\cdots x_{i+1}, v_{i+1})=\hat{\phi}(x, v)_{i+1}=(T\hat{\phi}(x, v))_i\,.\end{aligned}$$

Hence we can define an integration channel ν determined by the mapping $\hat{\phi}$ and a noise source s on D^I. The mapping $\hat{\phi}$ satisfies the conditions (II) and (III), for (II) is clear and (III) follows from the Kuratowski theorem. The integration channel defined as above is clearly an m-memory channel, i. e.,

$$\nu_x([y_i\cdots y_j])=\nu_{x'}([y_i\cdots y_j]) \qquad (i\leqq j)$$

if

$$[x_{i-m}x_{i-m+1}\cdots x_j]=[x'_{i-m}x'_{i-m+1}\cdots x'_j]\,.$$

THEOREM 2. *For the integration channel determined by* $(\hat{\phi}, s)$, *the transmission rate is obtained by*

$$R_p=h_q-h_s\,.$$

PROOF. As we can prove easily

$$h_r=\lim_n \frac{1}{n} \sum_{x_{1-m}\cdots x_n} \sum_{y_1\cdots y_n} r(([x_{1-m}\cdots x_0]\times Y)\cap[(x_1, y_1)\cdots(x_n, y_n)])$$
$$\cdot\log r(([x_{1-m}\cdots x_0]\times Y)\cap[(x_1, y_1)\cdots(x_n, y_n)]),$$

we get

$$R_p=h_p+h_q-h_r$$
$$=h_q+\lim_n \frac{1}{n} \sum_{x_{1-m}\cdots x_n} \sum_{y_1\cdots y_n} p([x_{1-m}\cdots x_n])\nu_x([y_1\cdots y_n])\log \nu_x([y_1\cdots y_n]).$$

Now putting

$$M_i(a_0 a_1\cdots a_n, b)=\{v\in D^I : \phi_0(a_0\cdots a_n, v_i)=b\},$$

we see

$$\nu_x([y_1\cdots y_n])=\int_V \chi_{[y_1\cdots y_n]}(\hat{\phi}(x, v))s(dv)$$
$$=s(M_1(x_{1-m}\cdots x_1, y_1)\cap\cdots\cap M_n(x_{n-m}\cdots x_n, y_n)).$$

Denote

$$B_0=\{\phi_0(x_{i-m}\cdots x_i, d) : d\in D\}.$$

Then for every $y_i\in B_0$ there exists one and only one $[v_i]\in\mathcal{V}_i=S^{-i}\mathcal{V}_0$ such that $M_i(x_{i-m}\cdots x_i, y_i)=[v_i]$. If $y_i\in B\backslash B_0$, then $M_i(x_{i-m}\cdots x_i, y_i)=\emptyset$. Therefore

$$R_p=h_q-\lim_n \frac{1}{n} \sum_{x_{1-m}\cdots x_n} p([x_{1-m}\cdots x_n])H(\mathcal{V}_0\vee P^{-1}\mathcal{V}_0\vee\cdots\vee P^{-n+1}\mathcal{V}_0)$$
$$=h_q-\lim_n \frac{1}{n} H(\mathcal{V}_0\vee P^{-1}\mathcal{V}_0\vee\cdots\vee P^{-n+1}\mathcal{V}_0)=h_q-h_s.$$

Q. E. D.

As the class Π, let us choose the set of all S-invariant probability measures on $X=A^I$. Then:

THEOREM 3. *For the integration channel determined by* $(\hat{\phi}, s)$, *the stationary capacity* C *is achieved by some ergodic source* $p_0\in\Pi_e$, *i.e.,* $C=R_{p_0}$.

PROOF[3]. The finite alphabet space A^I is a compact metric space by the Tychonoff product topology. By the Riesz-Markov-Kakutani representation theorem, the set Π of input sources can be imbedded in the positive part of the unit sphere of $C^*(A^I)$, the conjugate space of the Banach space $C(A^I)$ of all real valued continuous functions of A^I, and the set Π is compact convex in $C^*(A^I)$ with the weak* topology. As the channel ν is of finite memory, we can derive (see Umegaki [7] p. 60) that

$$\frac{1}{n}H(\mathcal{Y}_0\vee T^{-1}\mathcal{Y}_0\vee\cdots\vee T^{-n+1}\mathcal{Y}_0)$$

is a real valued continuous function on Π. Furthermore

3) The proof is a reformation of Breiman [2], in which he proved that the ergodic capacity and the stationary capacity coincide for finite memory, finitely correlated channels.

$$h_q=\inf_n\left\{\frac{1}{n}H(\mathcal{Y}_0\vee T^{-1}\mathcal{Y}_0\vee\cdots\vee T^{-n+1}\mathcal{Y}_0)\right\}$$

is a well known formula ([3], [5]), which shows that h_q is upper semicontinuous on Π when p is varied. The remaining part of the proof is same as Breiman [2]. Q. E. D.

COROLLARY 1. *If ν is a channel of additive noise defined in Corollary 4 of Theorem 1, then $C=R_{p_0}$ for some ergodic source p.*

Next, we assume that $A=B=D$ and is a finite group. Put $\phi_1(a, d)=a\cdot d$, the product in this group. The channel determined by the ϕ_1 is called a *channel of productive noise.* For this channel, Theorem 3 is also valid. However, the more clarified expression is given in the following:

THEOREM 4. *For a channel of productive noise, the capacity C is expressed by*

$$C=\log(\text{Card}\,A)-h_s\,,$$

where Card A *is the cardinarity of a set A.*

PROOF. Putting $N=\text{Card}\,A$, we consider a Bernoulli-source[4] $\tilde{p}$ on A^I determined by an N-dimensional probability vector $(1/N, 1/N, \cdots, 1/N)$. Then, for the output source $q(\cdot)=q(\cdot\,; p, \nu)$,

$$\begin{aligned}\tilde{q}([y_1\cdots y_n])&=\sum_{x_1\cdots x_n}\nu_x([y_1\cdots y_n])\tilde{p}([x_1\cdots x_n])\\&=\frac{1}{N^n}\sum_{x_1\cdots x_n}\nu_x([y_1\cdots y_n])\\&=\frac{1}{N^n}\sum_{x_1\cdots x_n}s([x_1^{-1}\cdot y_1, x_2^{-1}\cdot y_2, \cdots, x_n^{-1}\cdot y_n])=\frac{1}{N^n},\end{aligned}$$

where the last equality follows from the fact that $x_i^{-1}\cdot y_i$ moves all over A when x_i is varied. Hence q is also a Bernoulli measure and $h_p=\log N$. Therefore,

$$C=\sup_p(h_q-h_s)\leqq\log N-h_s=R_{\tilde{p}}\leqq C\,.$$

Q. E. D.

THEOREM 5. *For a channel of productive noise, the noise source is ergodic, if and only if $C=C_e=\log(\text{Card}\,A)-h_s$.*

PROOF. Necessity: Let p be the same Bernoulli source as defined in the above proof. Put $r(\cdot)=r(\cdot\,; p, \nu)$. It suffices to prove that r is ergodic, which is clear from the remark under Proposition 1.

4) A Bernoulli-source p determined by a probability vector $(p_1p_2\cdots p_N)$ is a source which gives a probability to any thin cylinder $[x_1x_2\cdots x_n]$ $(x_i\in A=\{a_1a_2\cdots a_N\})$, in such a way

$$p([x_1x_2\cdots x_n])=p_{i_1}p_{i_2}\cdots p_{i_n}$$

where $x_j=a_{i_j}$ in A and p_{i_j} is an element of the vector $(p_1p_2\cdots p_N)$.

Sufficiency: If $C=C_e=0$, then $h_s=\log N$ by Theorem 3 and $s(\cdot)$ is a Bernoulli measure, hence is ergodic. If $C=C_e>0$, then there exists an ergodic source p_0 and $r_0(\cdot)=r(\cdot\,;p_0,\nu)$ is ergodic. Then by Corollary 1 of Theorem 1, $\tilde{p}\times s$ must be ergodic, and which implies ergodicity of $s(\cdot)$. Q. E. D.

We can expect some applications of this theory. For example, models of burst errors are given by integration channels, choosing m-fold Markov chains as noise sources. By these models we will be able to faithfully represent many types of errors in various communication channels.

Bibliography

[1] P. Billingsley, Ergodic theory and information, John Wiley and Sons, Inc., 1965.

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

[3] A. Feinstein, Foundations of information theory, McGraw-Hill, 1958.

[4] P.R. Halmos, Lectures on ergodic theory, Math. Soc. Japan, 1956.

[5] A.I. Hinchin, Mathematical foundations of information theory, Dover, 1958 (English translation).

[6] K.R. Parthasarathy, Probability measures on metric spaces, Academic Press, 1967.

[7] H. Umegaki, Representations and extremal properties of averaging operators and their application to information channels, J. Math. Analysis Appl., 25 (1969), 41–73.

Yatsuka NAKAMURA
Department of Information Engineering
Faculty of Engineering
Shinshu University
Wakasato, Nagano
Japan