KYBERNETIKA — VOLUME 21 (1985), NUMBER 1

SOME FUNCTIONALS ON SETS OF STATIONARY CODES

ŠTEFAN ŠUJAN

Functionals of Kieffer-Rahe type measuring the degree of non-invertibility of stationary codes are defined, and their continuity properties are studied. General approximation theorems in the spirit of Sinai's weak isomorphism theorem and Ornstein's isomorphism theorem are derived.

1. INTRODUCTION

The theory of Bernoulli process and the associated coding techniques [3] found numerous applications in ergodic and information theories (see, e.g., the survey [6]). In this paper we address the question of abstract (topological) backgrounds of these techniques, without touching the problems connected with actual constructions of the corresponding codes (as to that cf. [6]). The main tool will be certain functionals on sets of stationary codes called Kieffer-Rahe functionals following [2], where similar quantities have been introduced for the purpose of universal coding in ergodic theory. We study various continuity properties of these functionals which result, in particular, in abstract formulations of Sinai's and Ornstein's theorems.

2. KIEFFER-RAHE FUNCTIONALS

Let A be a finite set with $\|A\|$ elements. If $\mathscr{A} = \{E : E \subset A\}$, then $(A^Z, \mathscr{A}^Z)$ $(Z = \{\ldots, -1, 0, 1, \ldots\}$ will denote the measurable space consisting of A^Z, the set of all doubly-infinite sequences $x = (x_i; i \in Z)$ with $x_i \in A$, and $\mathscr{A}^Z$, the usual product σ-field of subsets of A^Z. We let $\mathbf{M}(A)$, $\mathbf{E}(A)$, and $\mathbf{B}(A)$ denote the sets of all T_A-invariant, T_A-invariant and ergodic, and Bernoulli probability measures on $(A^Z, \mathscr{A}^Z)$, where T_A is the usual (two-sided) shift on A^Z [6].

We make the following notational convention. If $\mu_1, \ldots, \mu_n$ are probability measures on sequence spaces, and if an object, $0(\mu_1, \ldots, \mu_n)$, is defined, then we shall write also $0(X_1, \ldots, X_n)$, where X_i are processes with $\operatorname{dist}(X_i) = \mu_i$, $1 \leqq i \leqq n$ (e.g., if $\bar{d}(\mu_1, \mu_2)$ is the $\bar{d}$-distance between μ_1, $\mu_2 \in \mathbf{M}(A)$ and if X_1, X_2 are stationary processes with the state space A, $\operatorname{dist}(X_1) = \mu_1$, and $\operatorname{dist}(X_2) = \mu_2$, then we shall write $\bar{d}(\mu_1, \mu_2) = \bar{d}(X_1, X_2)$).

If B is another finite set, $\mathscr{Z}_A(B)$ will denote the set of all measurable partitions $P = (P^b; b \in B)$ of $A^{\mathbf{Z}}$ indexed by B. Also, F_{AB} will stand for the set of all stationary codes $\bar{\varphi} : A^{\mathbf{Z}} \to B^{\mathbf{Z}}$. By definition, $\bar{\varphi} \in F_{AB}$ if

(1) $$\bar{\varphi}^{-1}\mathscr{B}^{\mathbf{Z}} \subset \mathscr{A}^{\mathbf{Z}} \quad \text{and} \quad \bar{\varphi} \circ T_A = T_B \circ \bar{\varphi} \, .$$

The formulae

(2) $$P_\varphi = (\varphi^{-1}\{b\}; b \in B) \, , \quad \text{where} \quad \bar{\varphi} \in F_{AB} \, , \quad \varphi(x) = (\bar{\varphi}x)_0 \; ;$$

(3) $$(\bar{\varphi}_P x)_i = b \quad \text{if} \quad T_A^i x \in P^b \, , \quad \text{where} \quad P = (P^b; b \in B) \in \mathscr{Z}_A(B)$$

establish a one-to-one correspondence between equivalence classes (resulting from mod 0 identifications) of F_{AB} and $\mathscr{Z}_A(B)$ relative to any fixed $\mu \in \mathbf{M}(A)$. This means that two codes $\bar{\varphi}$, $\bar{\psi} \in F_{AB}$ are equivalent if $\bar{\varphi} = \bar{\psi} \bmod_\mu 0$, i.e., if $\mu\{x: \bar{\varphi}x \neq \bar{\psi}x\} = 0$. Two partitions P, $Q \in \mathscr{Z}_A(B)$ are equivalent if $P = Q \bmod_\mu 0$, i.e., if $|P - Q|_\mu = 0$, where

(4) $$|P - Q|_\mu = \tfrac{1}{2} \sum_{b \in B} \mu(P^b \,\Delta\, Q^b) \, ,$$

and Δ stands for the set-theoretical symmetric difference. Let

(5) $$P_A = (\{x \in A^{\mathbf{Z}}: x_0 = a\} \; ; \quad a \in A) \, .$$

Then $\bigvee \{T_A^i P_A; \; i \in Z\} = \sigma(\bigcup\{T_A^i P_A: i \in Z\}) = \mathscr{A}^{\mathbf{Z}}$, i.e., P_A is a generator in the strict sense. For $j \in \{0\} \cup N$ $(N = \{1, 2, \ldots\})$ put

(6) $$P_{A,j} = \bigvee_{i=-j}^{j} T_A^i P_A \, .$$

For each j, let $Q_j^{(1)}$, $Q_j^{(2)}$, … be an enumeration of all those partitions of $B^{\mathbf{Z}}$ whose atoms are finite-dimensional cylinders, and $\|Q_j^{(i)}\| = \|P_{A,j}\|$, $i \in N$. For any $\mu \in \mathbf{M}(A)$ and $P \in \mathscr{Z}_A(B)$ we define

(7) $$G(\mu, P) = \sum_{j=0}^{\infty} 2^{-j-1} \inf_{i \geq 1} |\bar{\varphi}_P^{-1} Q_j^{(i)} - P_{A,j}|_\mu \, .$$

If ν, $\nu' \in \mathbf{M}(B)$, let $\bar{d}(\nu, \nu')$ denote their $\bar{d}$-distance (see [3] or [6]). If $\mu \in \mathbf{M}(A)$, $\nu \in \mathbf{M}(B)$, and $P \in \mathscr{Z}_A(B)$, then we define

(8) $$F(\mu, \nu, P) = G(\mu, P) + \bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) \, .$$

F and G will be called *Kieffer-Rahe functionals* (in [2], G was not considered explicitly, and F was considered in the special case when $F(\mu, \nu, P) = F(\mu, \nu_\mu, P)$, where $\mu \mapsto \nu_\mu$ was a certain correspondence).

It is well-known that a countable (possibly finite) measurable partition P of $A^{\mathbf{Z}}$

is a generator (with respect to T_A and μ, $\mu \in \mathbf{M}(A)$), i.e.

$$\bigvee_{i \in Z} T_A^i P = \mathscr{A}^Z \bmod_\mu 0 , \tag{9}$$

if and only if the code $\bar{\varphi}_P$ $(B = P)$ is invertible. This may be expressed in terms of G as well:

Lemma 1. A finite measurable partition P of A^Z is a generator with respect to T_A and μ; $\mu \in \mathbf{M}(A)$, if and only if $G(\mu, P) = 0$.

We omit the standard proof (based on the idea of the proof that an almost everywhere injective measurable mapping between two standard probability spaces is automatically invertible cf., e.g., [3, Appendix A]).

Lemma 2. Let $\mu \in \mathbf{M}(A)$, $P^{(n)}$, $P \in \mathscr{Z}_A(B)$ be such that $|P^{(n)} - P|_\mu \to 0$. Then $G(\mu, P^{(n)}) \to G(\mu, P)$. In other words, the function $P \to G(\mu, P)$ is continuous relative the semimetric (4) on $\mathscr{Z}_A(B)$.

Proof. This follows from the fact that, for fixed i and j,

$$\begin{aligned} &\left| |\bar{\varphi}_{P^{(n)}}^{-1} Q_j^{(i)} - P_{A,j}|_\mu - |\bar{\varphi}_P^{-1} Q_j^{(i)} - P_{A,j}|_\mu \right| \leqq \\ &\leqq \tfrac{1}{2} \sum_k \mu[\bar{\varphi}_{P^{(n)}}^{-1} Q_j^{(i)}(k) \, \Delta \bar{\varphi}_P^{-1} Q_j^{(i)}(k)] , \end{aligned}$$

where the summation is over all atoms of $Q_j^{(i)}$. In fact, it remains to observe that if $P \in \mathscr{Z}_A(B)$ then the set $\varphi_P^{-1} Q_j^{(i)}(k)$ may be expressed in terms of taking unions, intersections, and shifts of the atoms of P (we do not present cumbersome general expressions and give only an example: if $Q_j^{(i)}(k) = \{y \in B^Z : y_0 \in \{b_1, b_2\}, y_1 = b_3\}$ then $\bar{\varphi}_P^{-1} Q_j^{(i)}(k) = (P^{b_1} \cup P^{b_2}) \cap T_A^{-1} P^{b_3}$). □

Lemma 3. Let $\mu \in \mathbf{M}(A)$ and $\nu \in \mathbf{M}(B)$. Then
(a) the function $P \mapsto F(\mu, \nu, P)$ is continuous relative to the semimetric (4) on $\mathscr{Z}_A(B)$; and
(b) the shifts (T_A, μ) and (T_B, ν) are mod 0 isomorphic (cf. [8] for the definition) if and only if there is a $P \in \mathscr{Z}_A(B)$ such that $F(\mu, \nu, P) = 0$.

Proof. Both assertions are implicitly contained in [2]. In order to prove (a), use Lemma 2 together with the following facts:

$$\bar{d}(\mu\bar{\varphi}_P^{-1}, \mu\bar{\varphi}_Q^{-1}) \leqq |P - Q|_\mu \quad \text{(cf. [3])};$$

$$|d(x, y) - d(x, z)| \leqq d(y, z)$$

(the latter is valid for any semimetric, and is a simple consequence of the triangle inequality). In light of Lemma 1, part (b) just recalls definitions in a different form. □

The key steps of Ornstein's coding technique may be conveniently expressed in terms of F and G:

Lemma 4. (a) Let $\mu \in \mathbf{E}(A)$ and $\nu \in \mathbf{B}(B)$ satisfy $h_\mu(T_A) \geqq h_\nu(T_B)$ (h stands for entropy [1]). For any $\varepsilon > 0$ there is a $\delta > 0$ such that for any $Q \in \mathscr{X}_A(B)$ satisfying $\bar{d}(\mu\bar{\varphi}_Q^{-1}, \nu) < \delta$ we can find a $P \in \mathscr{X}_A(B)$ with $|P - Q|_\mu < \varepsilon$ and $\bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) = 0$.

(b) Suppose, in addition, that $\mu \in \mathbf{B}(A)$ and $h_\mu(T_A) = h_\nu(T_B)$. For any $\varepsilon > 0$ there is a $\delta > 0$ such that for any $Q \in \mathscr{X}_A(B)$ satisfying $F(\mu, \nu, Q) < \delta$ we can find a $P \in \mathscr{X}_A(B)$ with $|P - Q|_\mu < \varepsilon$ and $F(\mu, \nu, P) = 0$.

Part (a) says that having a coding sufficiently close to the Bernoulli measure ν we can change that coding slightly, and thereby get a conding that yields ν exactly, i.e., under which ν becomes a factor (= a stationary coding) of μ (thus, strong Sinai's theorem is valid; cf. [7] for its information-theoretic implications).

Part (b) does not follows immediately from (a). Indeed, the possibility of constructing a sequence of partitions which become ever "more generating", and the encoded measures become ever "closer" to $\nu \in \mathbf{B}(B)$ is the core of Ornstein's technique (see [3], Lemmas I.4.7, I.4.10, Proposition I.4.11).

3. CONTINUITY PROPERTIES

On $\mathbf{M}(B)$ we place the weak topology. As well-known [4], it is metrizable, and we define the weak metric d_w by

$$d_w(\nu, \nu') = \sum_{n=1}^{\infty} 2^{-n} \sum_{\mathbf{b} \in B^n} |\nu^n(\mathbf{b}) - (\nu')^n(\mathbf{b})| , \tag{10}$$

where $\nu^n(\mathbf{b}) = \nu\{y \in B^Z: y_0^{n-1} = \mathbf{b}\}$, $y_0^{n-1} = (y_0, \ldots, y_{n-1})$; and similarly for ν'.

Theorem 1. Let $\mu \in \mathbf{B}(A)$, $\nu \in \mathbf{B}(B)$, and $P \in \mathscr{X}_A(B)$. Then $F(\mu, \nu, P) = 0$ if and only if for any $\varepsilon > 0$ there is a $\delta > 0$ so that $F(\mu, \nu', P) < \varepsilon$ for any $\nu' \in \mathbf{E}(B)$ satisfying $d_w(\nu, \nu') < \delta$ and $|h(\nu) - h(\nu')| < \delta$ ($h(\nu)$ stands as an abbreviation for $h_\nu(T_B)$).

Proof. First let us prove sufficiency of the latter condition. Using it choose a sequence $\nu^{(n)} \in \mathbf{E}(B)$ such that

(a) $d_w(\nu^{(n)}, \nu) \to 0$,

(b) $|h(\nu^{(n)}) - h(\nu)| \to 0$, and

(c) $F(\mu, \nu^{(n)}, P) \to 0$.

Since $\nu \in \mathbf{B}(B)$, (T_B, ν) is finitely determined [3, 6] so that (a) and (b) entail

(d) $\bar{d}(\nu^{(n)}, \nu) \to 0$.

Since $F(\mu, \nu^{(n)}, P) = G(\mu, P) + \bar{d}(\mu\bar{\varphi}_P^{-1}, \nu^{(n)})$, the only way in which (c) may happen is that $G(\mu, P) = 0$. By (d) and (c),

$$\bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) \leqq \bar{d}(\mu\bar{\varphi}_P^{-1}, \nu^{(n)}) + \bar{d}(\nu^{(n)}, \nu) \to 0 .$$

Hence $\bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) = 0$ so that $F(\mu, \nu, P) = 0$.

We prove necessity indirectly. Suppose $F(\mu, \nu, P) = 0$ but the condition is false. Then we find an $\varepsilon_0 > 0$ such that for any $\delta > 0$ there is a $\nu' \in \mathbf{E}(B)$ so that $d_w(\nu', \nu) < \delta$, $|h(\nu') - h(\nu)| < \delta$, but $F(\mu, \nu', P) \geqq \varepsilon_0$. Hence, pick a sequence $\nu^{(n)} \in \mathbf{E}(B)$

with $\bar{d}(\nu^{(n)}, \nu) \to 0$ (use again bernoullicity of ν) and $F\mu, \nu^{(n)}, P) \geqq \varepsilon_0$ for any $n \in N$.) Using triangle inequality,

$$G(\mu, P) + \bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) + \bar{d}(\nu, \nu^{(n)}) \geqq F(\mu, \nu^{(n)}, P) \geqq \varepsilon_0 .$$

Take the limit as $n \to \infty$:

$$F(\mu, \nu, P) = G(\mu, P) + \bar{d}(\mu\bar{\varphi}_P^{-1}, \nu) \geqq \varepsilon_0 ;$$

a contradiction. □

Remark 1. Let $\mu \in \mathbf{B}(A)$, $\nu \in \mathbf{B}(B)$, and $P \in \mathscr{Z}_A(B)$ be such that $F(\mu, \nu, P) = 0$. Then $\bar{\varphi}_P \in F_{AB}$ is invertible so that $\bar{\varphi}_P^{-1} \in F_{BA}$ is a well-defined stationary code. If $P^{-1} \in \mathscr{Z}_B(A)$ corresponds to $\bar{\varphi}_P^{-1}$ (cf. (2)), then $F(\nu, \mu, P^{-1}) = 0$. Hence, Theorem 1 can be formulated also dually, i.e., for fixed ν and for changing μ, respectively.

Theorem 2. (Sinai's Theorem). Let $X \in \mathbf{M}(A)$ and $Y \in \mathbf{M}(B)$ be such that there exist processes $X^{(n)} \in \mathbf{M}(A)$, $Y^{(n)} \in \mathbf{M}(B)$, and partitions $Q^{(n)} \in \mathscr{Z}_A(B)$ $(n \in N)$ such that

(a) $\bar{d}(X^{(n)}, X) \to 0$,
(b) $\bar{d}(Y^{(n)}, Y) \to 0$,
(c) $|Q^{(n)} - P|_X \to 0$ for some $P \in \mathscr{Z}_A(B)$, and
(d) $\bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, Y^{(n)}) \to 0$ as $n \to \infty$.

Then $\bar{d}(\bar{\varphi}_P X, Y) = 0$, i.e., $Y = \bar{\varphi}_P X \bmod 0$. That is, Y is a factor of X.

Proof. On the first step we show
(e) $\bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y) \to 0$.
We have

$$\bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, Y) \leqq \bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, Y^{(n)}) + \bar{d}(Y^{(n)}, Y) \to 0$$

by (b) and (d), hence
(f) $\bar{d}(\bar{\varphi}_{Q^{(n)}} X^{(n)}, Y) \to 0$.
Next
(g) $\bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y) \leqq \bar{d}(\bar{\varphi}_{Q^{(n)}}X, \bar{\varphi}_{Q^{(n)}}X^{(n)}) + \bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, Y)$.
In order to evaluate the middle term in (g) we employ the following equivalent description of $\bar{d}$ (see [3, Appendix C]). Let $U = (U_i; i \in Z)$ and $V = (V_i; i \in Z)$ be two stationary processes over the same finite alphabet C. Then $\bar{d}(U, V) < \varepsilon$ if and only if on some probability space there exist jointly stationary processes $\tilde{U}$ and $\tilde{V}$, each with the state space C, such that dist $(\tilde{U})$ = dist (U), dist $(\tilde{V})$ = dist (V) and

$$\operatorname{Prob}[\tilde{U}_0 \neq \tilde{V}_0] < \varepsilon .$$

Now fix a finite set D and a code $\bar{\varphi} \in F_{CD}$. Then $(\bar{\varphi}\tilde{U}, \bar{\varphi}\tilde{V})$ is again jointly stationary, dist $(\bar{\varphi}\tilde{U})$ = dist $(\bar{\varphi}U)$, dist $(\bar{\varphi}\tilde{V})$ = dist $(\bar{\varphi}V)$, and

$$\operatorname{Prob}[(\bar{\varphi}\tilde{U})_0 \neq (\bar{\varphi}\tilde{V})_0] \leqq \operatorname{Prob}[\tilde{U}_0 \neq \tilde{V}_0]$$

(for $\bar{\varphi}$ only can "past together" different letters). Hence
(h) $\bar{d}(\bar{\varphi}U, \bar{\varphi}V) \leqq \bar{d}(U, V)$.

Use (h) for $n \in N$, $U = X^{(n)}$, $V = X$, $\bar{\varphi} = \bar{\varphi}_{Q^{(n)}}$:
(i) $\bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, \bar{\varphi}_{Q^{(n)}}X) \leqq \bar{d}(X^{(n)}, X)$; $n \in N$.
Combining (i), (a), (f), and (g) results in (e). But

$$\bar{d}(\bar{\varphi}_P X, Y) \leqq \bar{d}(\bar{\varphi}_P X, \bar{\varphi}_{Q^{(n)}}X) + \bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y) \to 0$$

since (e) applies to the second summand, while from (c) it follows that

$$\bar{d}(\bar{\varphi}_P X, \bar{\varphi}_{Q^{(n)}}X) \leqq |P - Q^{(n)}|_X \to 0 .$$

□

Theorem 3. (Isomorphism Theorem). Suppose $X \in \mathbf{M}(A)$, $Y \in \mathbf{M}(B)$, and $P \in \mathscr{X}_A(B)$ be such that there exist sequences $Y^{(n)} \in \mathbf{M}(B)$ and $Q^{(n)} \in \mathscr{X}_A(B)$ satisfying
(a) $\bar{d}(Y^{(n)}, Y) \to 0$,
(b) $|Q^{(n)} - P|_X \to 0$, and
(c) $F(X, Y^{(n)}, Q^{(n)}) \to 0$ as $n \to \infty$.
Then $F(X, Y, P) = 0$.

Proof. Apply Theorem 2 to the constant sequence $X^{(n)} = X$ $(n \in N)$; we get $\bar{d}(\bar{\varphi}_P X, Y) = 0$. From (c) it follows that $G(X, Q^{(n)}) \to 0$. Use (b) and Lemma 2 to conclude that $G(X, Q^{(n)}) \to G(X, P)$; hence $G(X, P) = 0$, too. □

Remark 2. The proofs of Theorems 2 and 3 are easy because the major point of Ornstein's coding technique – the construction of a converging sequence of partitions – is already built in our hypotheses ((c) of Theorem 2, (b) of Theorem 3). A remark of Ornstein (see [3], remark after Proposition I.4.8) throws light upon which conditions could entail convergence. In fact, if a sequence $Q^{(n)}$ of partitions is ever better in the sense of F, then this alone does not lead to the conclusion that the $Q^{(n)}$'s approach a generator P. The reason is that this approach may be too slow. But if we could prove, say, that $\sum |Q^{(n+1)} - Q^{(n)}|_X < \infty$, then we would obtain convergence. But the latter condition says essentially that the elements of the sequence $\{Q^{(n)}; n \in N\}$ are not allowed to differ much from each other.

We separate the idea suggested in the above remark in the form of the following lemma. Let

(11) $$\varrho_\mu(P, Q) = |P - Q|_\mu ; \quad \mu \in \mathbf{M}(A) , \quad P, Q \in \mathscr{X}_A(B) \quad (\text{cf. } (4)) ,$$

(12) $$d_e(v, v') = d_w(v, v') + |h(v) - h(v')| ; \quad v', v' \in \mathbf{M}(B) .$$

Lemma 5. Let $\mu \in \mathbf{B}(A)$ and $v \in \mathbf{B}(B)$. Let $\mathscr{X} \subset \mathscr{X}_A(B)$ be a ϱ_μ-totally bounded set. Suppose for each $\varepsilon > 0$ there is a $\delta > 0$ such that for any $v' \in \mathbf{E}(B)$ satisfying $d_e(v, v') < \delta$ there is a partition $Q \in \mathscr{X}$ with $F(\mu, v', Q) < \varepsilon$. Then there is a $P \in \mathscr{X}_A(B)$ satisfying $F(\mu, v, P) = 0$.

Proof. Use the assumptions and find sequences $v^{(n)} \in \mathbf{E}(B)$ and $Q^{(n)} \in \mathscr{X}_A(B)$ such that
(a) $\bar{d}(v^{(n)}, v) \to 0$ and
(b) $F(\mu, v^{(n)}, Q^{(n)}) \to 0$

((a) is obtained as in the proof of Theorem 1). From (a) and (b) we have

(c) $F(\mu, \nu, Q^{(n)}) \leq G(\mu, Q^{(n)}) + \bar{d}(\mu\bar{\varphi}_{Q^{(n)}}^{-1}, \nu^{(n)}) + \bar{d}(\nu^{(n)}, \nu) =$
$= F(\mu, \nu^{(n)}, Q^{(n)}) + \bar{d}(\nu^{(n)}, \nu) \to 0 .$

In particular,

(d) $\bar{d}(\mu\bar{\varphi}_{Q^{(n)}}^{-1}, \nu) \to 0$

(note that this is the same as (e) in the proof of Theorem 2). We claim that the sequence $\{Q^{(n)}\}$ is ϱ_μ-Cauchy (cf. 11)), i.e.

(e) $\lim_{m,n\to\infty} |Q^{(n)} - Q^{(m)}|_\mu = 0 .$

If C is a finite set, let R_C denote the set of all regular points, i.e. of all sequences $u \in C^Z$ such that there exists a measure $\mu \in \mathbf{E}(C)$ so that

(f) $\mu(V) = \lim_{n\to\infty} n^{-1} \sum_{i=0}^{n-1} 1_V(T_C^i u)$

for all elementary cylinders $V \subset C^Z$, where 1_V stands for the indicator function of V. If $\mu \in \mathbf{E}(C)$, let $R(\mu) = \{u \in R_C : \mu = \mu_u\}$, where μ_u is the unique measure in $\mathbf{E}(C)$ determined by the right-hand side of (f). If $\mu \in \mathbf{E}(C)$, then we have $\mu(R(\mu)) = 1$ and $\mu(R(\nu)) = 0$ if $\nu \in \mathbf{E}(C)$, $\nu \neq \mu$ (see [8] or [5, 6] for details).

Since for any $n \in N$, $\mu\bar{\varphi}_{Q^{(n)}}^{-1} \in \mathbf{E}(B)$ (in fact, it is even bernoulli; cf. [3]), we can express the $\bar{d}$-distance between encoded processes as follows:

(g) $\bar{d}(\mu\bar{\varphi}_{Q^{(n)}}^{-1}, \mu\bar{\varphi}_{Q^{(m)}}^{-1}) = \inf_{\substack{y' \in R(\mu\bar{\varphi}_{Q^{(n)}}^{-1}) \bmod 0 \\ y'' \in R(\mu\bar{\varphi}_{Q^{(m)}}^{-1}) \bmod 0}} \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H(y'_k, y''_k) ,$

where d_H is the usual Hamming distance on B. Let $\mu = \operatorname{dist}(X)$. For each $n \in N$, $\operatorname{dist}(\bar{\varphi}_{Q^{(n)}}X) = \mu\bar{\varphi}_{Q^{(n)}}^{-1}$, and we may suppose (by changing the underlying probability space if necessary) that the pair process $(\bar{\varphi}_{Q^{(n)}}X, \bar{\varphi}_{Q^{(m)}}X)$ is jointly ergodic. Now

$$2 \operatorname*{Prob}_\mu [(\bar{\varphi}_{Q^{(n)}}X)_0 \neq (\bar{\varphi}_{Q^{(m)}}X)_0] = \sum_{\substack{b', b'' \in B \\ b' \neq b''}} \operatorname*{Prob}_\mu [(\bar{\varphi}_{Q^{(n)}}X)_0 = b',$$
$$(\bar{\varphi}_{Q^{(m)}}X)_0 = b''] = 2|Q^{(n)} - Q^{(m)}|_\mu$$

(the latter equality follows from (2)); i.e.

(h) $|Q^{(n)} - Q^{(m)}|_\mu = \operatorname*{Prob}_\mu [(\bar{\varphi}_{Q^{(n)}}X)_0 \neq (\bar{\varphi}_{Q^{(m)}}X)_0] .$

Let

$$[b', b''] = \{(y', y'') \in B^Z \times B^Z : y'_0 = b', y''_0 = b''\} .$$

Using joint ergodicity of $(\bar{\varphi}_{Q^{(n)}}X, \bar{\varphi}_{Q^{(m)}}X)$ and the pointwise ergodic theorem we get

(i) $\operatorname*{Prob}_\mu [(\bar{\varphi}_{Q^{(n)}}X)_0 \neq (\bar{\varphi}_{Q^{(m)}}X)_0] = \frac{1}{2} \sum_{\substack{b', b'' \in B \\ b' \neq b''}} \lim_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} 1_{[b', b'']}(T_B^k \bar{\varphi}_{Q^{(n)}}x, T_B^k \bar{\varphi}_{Q^{(m)}}x)$

for μ-almost all $x \in R(\mu)$. We wish to prove that if (g) approaches zero then so does (i). However, there arises a difficulty, for (g) involves infimum over sequences from two distinct sets, while (i) concerns only sequences $x \in R(\mu)$. Therefore, we must proceed in a less straight way. Suppose (i) is not valid. Then we find a $\gamma_0 > 0$ and two different letters b'_0, $b''_0 \in B$ so that for μ-almost all $x \in R(\mu)$,

(j) $\lim_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} 1_{[b_0', b_0'']}(T_B^k \bar{\varphi}_{Q^{(n)}} x, T_B^k \bar{\varphi}_{Q^{(m)}} x) \geqq \gamma_0$

for infinitely many $n \in N$ and for infinitely many $m \in N$. In (j), a one appears in the sum only if $(\bar{\varphi}_{Q^{(n)}} x)_k = b'_0$ and $(\bar{\varphi}_{Q^{(m)}} x)_k = b''_0$. By the definition of Hamming distance, for any n, $m \in N$, the inequality

$$\lim_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} 1_{[b_0', b_0'']}(T_B^k \bar{\varphi}_{Q^{(n)}} x, T_B^k \bar{\varphi}_{Q^{(m)}} x) \leqq \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, (\bar{\varphi}_{Q^{(m)}} x)_k)$$

takes place for μ-almost all $x \in R(\mu)$. Hence, it follows from (j) that

(k) $\inf_{x \in R(\mu) \bmod 0} \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, (\bar{\varphi}_{Q^{(m)}} x)_k) \geqq \gamma_0$

for infinitely many $n \in N$ and infinitely many $m \in N$. Let $n \in N$. If $y \in R(\mu \bar{\varphi}_{Q^{(n)}}^{-1})$ then, with probability one, $y = \bar{\varphi}_{Q^{(n)}} x$ for some $x \in R(\mu)$ (see [5], Lemma 3). Hence, the abovementioned difficulty appears in case when y', y'' yielding "nearly" the infimum in (g) do not arise from the same $x \in R(\mu)$. In order to overcome this we use (d) rewritten in the form

$$(1) \qquad \inf_{\substack{x \in R(\mu) \bmod 0 \\ y \in R(\nu) \bmod 0}} \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, y_k) \to 0$$

as $n \to \infty$. Now, a simple consequence of the triangle inequality is the following one:

$$\limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, (\bar{\varphi}_{Q^{(m)}} x)_k) \leqq \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, y_k) + {} + \limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(m)}} x)_k, y_k).$$

Now pick $x \in R(\mu)$, $y \in R(\nu)$, and $n_0 \in N$ so that for $n \geqq n_0$

$$\limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, y_k) < \tfrac{1}{2}\gamma_0$$

(this is possible by (1)). If $m \geqq n_0$, too, we get the existence of an $x \in R(\mu)$ with

$$\limsup_{K\to\infty} K^{-1} \sum_{k=0}^{K-1} d_H((\bar{\varphi}_{Q^{(n)}} x)_k, (\bar{\varphi}_{Q^{(m)}} x)_k) < \gamma_0$$

for all n, $m \geqq n_0$, and this contradicts (k). Using (g) and (i) we get the claimed (e).

Now $\{Q^{(n)}; n \in N\}$ is a ϱ_μ-Cauchy sequence, and each $Q^{(n)} \in \mathscr{Z}$, where $\mathscr{Z}$ is ϱ_μ-totally bounded. Consequently, we can find a subsequence n_k and a partition $P \in \mathscr{Z}_A(B)$ (in fact, P belongs to the ϱ_μ-closure of $\mathscr{Z}$) so that

(m) $\lim_{k\to\infty} |Q^{(n_k)} - P|_\mu = 0$.

By Lemma 3(a), $F(\mu, \nu, Q^{(n_k)}) \to F(\mu, \nu, P)$. On the other hand, (c) implies $F(\mu, \nu, Q^{(n)}) \to 0$. Consequently, we get $F(\mu, \nu, P) = 0$ as desired. □

Next let us generalize Theorems 2 and 3. Since the proof of (e) in the proof of Theorem 2 does not employ the assumption that $|Q^{(n)} - P|_X \to 0$, the following result is implicitly contained in Theorem 2:

Theorem 4. Suppose the hypotheses of Theorem 2 are satisfied except that (c) is weakened to
(c') $\{Q^{(n)}; n \in N\} \subset \mathscr{Z}$, where $\mathscr{Z} \subset \mathscr{Z}_A(B)$ is ϱ_μ-totally bounded; $\mu = \operatorname{dist}(X)$.
Then there is a partition $P \in \mathscr{Z}_A(B)$ such that $\bar{d}(\bar{\varphi}_P X, Y) = 0$.

It is intuitively clear that if two processes $X \in \mathbf{M}(A)$ and $Y \in \mathbf{M}(B)$ allow for approximations by processes $X^{(n)} \in \mathbf{M}(A)$ and $Y^{(n)} \in \mathbf{M}(B)$ such that $X^{(n)}$ and $Y^{(n)}$ become ever "more isomorphic" when n grows (and the corresponding codes do not differ too much; see Remark 2), then X and Y themselves should be isomorphic. However, Theorem 3 gives us only a weaker result in that X was kept fixed and merely Y was approximated by a sequence $Y^{(n)}$. Lemma 5 allows us to remove this asymmetry.

Indeed, add to (a) and (b) of Theorem 3 the condition

$$\bar{d}(X^{(n)}, X) \to 0 \tag{13}$$

and modify (c) of Theorem 3 to

$$F(X^{(n)}, Y^{(n)}, Q^{(n)}) \to 0\,. \tag{14}$$

Then (13) and (14) entail

$$\begin{aligned}\bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y^{(n)}) &\leqq \bar{d}(\bar{\varphi}_{Q^{(n)}}X, \bar{\varphi}_{Q^{(n)}}X^{(n)}) + \bar{d}(\bar{\varphi}_{Q^{(n)}}X^{(n)}, Y^{(n)}) \leqq \\ &\leqq \bar{d}(X, X^{(n)}) + F(X^{(n)}, Y^{(n)}, Q^{(n)}) \to 0\,.\end{aligned}$$

Since

$$\bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y) \leqq \bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y^{(n)}) + \bar{d}(Y^{(n)}, Y)\,,$$

we get again

$$\lim_{n \to \infty} \bar{d}(\bar{\varphi}_{Q^{(n)}}X, Y) = 0\,. \tag{15}$$

We see that (15) was obtained without the assumption that $|Q^{(n)} - P|_X \to 0$ for some $P \in \mathscr{Z}_A(B)$. Hence, by Lemma 5 we get the following generalization of Theorem 3:

Theorem 5. Let $X \in \mathbf{M}(A)$ and $Y \in \mathbf{M}(B)$ be such that there exist sequences $X^{(n)} \in \mathbf{M}(A)$, $Y^{(n)} \in \mathbf{M}(B)$, and $Q^{(n)} \in \mathscr{Z}_A(B)$ such that
(a) $\bar{d}(X^{(n)}, X) \to 0$,
(b) $\bar{d}(Y^{(n)}, Y| \to 0$,
(c) $F(X^{(n)}, Y^{(n)}, Q^{(n)}) \to 0$, and
(d) $\{Q^{(n)}; n \in N\} \subset \mathscr{Z} \subset \mathscr{Z}_A(B)$, where $\mathscr{Z}$ is a ϱ_μ-totally bounded set, $\mu = \operatorname{dist}(X)$.
Then there exists a $P \in \mathscr{Z}_A(B)$ with $F(X, Y, P) = 0$.

Remark 3. The only condition concerning entropies in Theorem 5 is that $h(X^{(n)}) \to$

$\to h(X)$ and $h(Y^{(n)}) \to h(Y)$ (this follows from $\bar{d}$-continuity of entropy) and, of course, $h(X) = h(Y)$. In particular, we may pick $h(X^{(n)}) \downarrow 0$, $h(Y^{(n)}) \downarrow 0$, i.e., $h(X) = h(Y) = 0$. In other words, Theorem 5 is valid equally well also for zero-entropy processes. It is an easy exercise to prove that, under the conditions of Theorem 5, the approaches to zero of $h(X^{(n)})$ and $h(Y^{(n)})$ must be at the same rate. The formula

$$\lim_{n\to\infty} H(X_0^{n-1})/H(Y_0^{n-1}) = \lim_{n\to\infty} h(X^{(n)})/h(Y^{(n)}) \tag{16}$$

then shows that the speed of convergence in $n^{-1} H(X_0^{n-1}) \downarrow 0$ is a new isomorphism invariant for zero-entropy processes. However, a detailed investigation on the isomorphism problem for zero-entropy processes exceeds the frame of the present paper.

ACKNOWLEDGMENT

The author is indebted to an anonymous referee for his constructive criticism concerning the first draft of this paper.

(Received April 19, 1984.)

REFERENCES

[1] P. Billingsley: Ergodic Theory and Information. J. Wiley, New York—London—Sydney 1965.

[2] J. C. Kieffer and M. Rahe: Selecting universal partitions in ergodic theory. Ann. Probab. *9* (1981), 705—709.

[3] D. S. Ornstein: Ergodic Theory, Randomness, and Dynamical Systems. Yale Univ. Press, New Haven—London 1974.

[4] K. R. Parthasarathy: Probability Measures on Metric Spaces. Academic Press, New York 1967.

[5] Š. Šujan: A local structure of stationary perfectly noiseless codes between stationary non-ergodic sources I. Kybernetika *18* (1982), 361—375.

[6] Š. Šujan: Ergodic theory, entropy, and coding problems of information theory. Kybernetika *19* (1983), supplement, 68 pp.

[7] Š. Šujan: Sinai's theorem and entropy compression. Problems Control Inform. Theory *12* (1983), 419—428.

[8] K. Winkelbauer: On the asymptotic rate of non-ergodic information sources. Kybernetika *6* (1970), 127—148.

RNDr. Štefan Šujan, CSc., *Joint Institute for Nuclear Research, Laboratory of Computing Techniques and Automation, Head Post Office, P. O. Box 79, 101000 Moscow. USSR.*
Permanent address: *Ústav merania a meracej techniky CEFV SAV (Institute of Measurement and Measuring Technique, Electro-Physical Research Centre, Slovak Academy of Sciences), 842 19 Bratislava. Czechoslovakia.*