Kybernetika

Štefan Šujan

A local structure of stationary perfectly noiseless codes between stationary non-ergodic sources. I. General considerations

A LOCAL STRUCTURE OF STATIONARY
PERFECTLY NOISELESS CODES
BETWEEN STATIONARY NON-ERGODIC SOURCES
I: General Considerations

ŠTEFAN ŠUJAN

The structure of stationary and perfectly noiseless codes between stationary non-ergodic sources is related with ergodic decomposition of invariant probability measures. It is shown that such codes share the structure of ergodic fibres so that any such code is in fact a class of local codes between corresponding pairs of ergodic subsources. Conversely, the local structure is also sufficient for existence of global perfectly noiseless codes. It is proved that local structure is preserved also in case of finitary codes. Applications of these results will be presented in Part II of the paper.

INTRODUCTION

Many results of ergodic and information theories can be formulated as isomorphism theorems for certain classes of ergodic sources. The aim of this paper is to develop a general method of extension of such isomorphism theorems to aperiodic non-ergodic sources. In particular, extensions will be investigated (in Part II) of isomorphisms described in the following examples.

Example 1. Let S be the class of all ergodic sources over arbitrary countable alphabets which have finite entropy. Each member of S has an isomorphic representation as an ergodic source over a finite alphabet, as follows from Krieger's finite generator theorem [10].

Example 2. Let S be the class of all stationary memoryless sources over arbitrary finite alphabets, all with the same entropy. Any two members of S are isomorphic [11], and the isomorphism can be chosen even as a mod 0 homeomorphism [7].

Example 3. Let S be the class of all aperiodic non-ergodic sources over arbitrary finite alphabets with entropies less than the Shannon capacity of a fixed ergodic and weakly continuous channel. Each member of S has an isomorphic representation as an ergodic invulnerable source [8].

361

In language of information theory, an isomorphism is a perfectly noiseless code, that is, a stationary code such that we can exactly recover the original source from the encoded source. Classical source coding theorems which trace back to Shannon assert that we can get the probability of decoding error as small as we please by taking the block length sufficiently large. However, perfectly noiseless coding is not available within the traditional block coding approach. The problem is that increase of block length inserts considerable changes into the structure of block codes so that one cannot hope for a "limiting" code with zero probability of decoding error. The situation becomes completely different in case of stationary codes. Ornstein [11] observed that it is possible to make a good stationary code much better in exchange to only moderate changes in its structure. In this way, a converging sequence of ever better codes obtains, and its limit is the desired perfectly noiseless code.

All three examples can be considered as applications of Ornstein's coding technique (for Example 1 this was shown by Smorodinsky [17]) except the continuous version of Example 2 for which a different coding technique was developed in [7]. This fact places also some limitations from the point of view of our basic aim. Within Ornstein's approach, stationary codes are constructed from block codes using an auxiliary binary encoding which tells us when to use block code (see [16]; this idea is generally valid even Ornstein himself did not mention block codes at all, and used a marriage lemma in order to choose a good map from n-tuples to n-tuples; cf. [12]). In order we can control the probability of decoding error we have to distinguish, in the encoded sequence, between parts comming from that auxiliary binary encoding and parts which result from applications of the block code (again, Ornstein himself did not construct both encoder and decoder but used a Baire category argument in order to prove invertibility of the limiting code. Subsequent improvements on both encoder and decoder in spirit of Ornstein's technique are employed in [8].). For this we must ensure convergence of relative frequencies of blocks or at least of individual letters to corresponding probabilities. But this means we have to require either total ergodicity or at least ergodicity of the original source.

One way to overcome this difficulty is based on the fact that any ergodic source can be obtained from a stationary source by conditioning with respect to the σ-field of all invariant events. Within Example 1, Denker [3] proved that the only tools needed to derive Krieger's theorem (in the ergodic case) are Rohlin lemma and Shannon-McMillan theorem. If one was able to prove universal versions of these two assertions then the desired extension would follow quite immediately (apart from some technical difficulties connected with absence of universal bounds on speed of convergence). Extension of Krieger's theorem in this spirit are given in [4].

A general formulation of the idea of universality motivated by recent developments on universal source coding has been given by Kieffer and Rahe [9]. They consider classes of ergodic sources and the aim is to prove existence of a single code which performs optimally for any member of a specified class.

Our approach is based on a different idea. We consider mixtures of ergodic compo-

nents as possible candidates for the "true" source statistics, i.e., we always have a class of ergodic sources together with a weighting prior. Thus, our intention is to prove existence of perfectly noiseless codes for mixtures. Surprisingly enough, any such code is also universal in the following sense. Any perfectly noiseless code between two mixtures of ergodic subsources is, in fact, a class of local perfectly noiseless codes between corresponding pairs of ergodic subsources (Theorem 1). In particular, if with positive prior probability, there exist subsources of one mixture having no partners in the second mixture (e.g., for entropy reasons) then a perfectly noiseless code between the two mixtures cannot exist. Conversely, given a set of pairs of ergodic sources such that, in any pair, one source is a perfectly noiseless coding of the other one, we can construct a perfectly noiseless code for the corresponding mixtures (Theorem 2). Theorems 3 and 4 show that this local structure is preserved also in case of finitary perfectly noiseless codes, i.e., of codes which are also almost everywhere homeomorphic maps.

Part II of the paper will be devoted to applications of these general ideas to a variety of problems arising both in ergodic theory and information, including those ones described in the above three examples.

1. PRELIMINARIES

Throughout the paper we shall deal with stationary sources over countable (i.e., finite or countably infinite) alphabets. An ergodic theorist can imagine a stationary source as the structure obtained from a countable partition of a Lebesgue probability space with an automorphism.

Let A be a countable set; if A is finite, let $|A|$ denote its cardinality. Let $\mathscr{A}$ stand for the σ-field of all subsets of A; $\mathscr{A}$ is considered as the σ-field of all Borel subsets of A, hence, we always consider A as a topological space relative to the discrete topology. Put $Z = \{\ldots, -1, 0, 1, \ldots\}$ and $N = \{1, 2, \ldots\}$. As usually, A^Z designates the space of all doubly-infinite sequences $u = (u_i)_{i \in Z}$ with $u_i \in A$. We let $\mathscr{A}^Z$ denote the product σ-field, i.e., the least σ-field containing the class $\mathscr{V}(A)$ of all thin cylinders. By definition, a set $E \subset A^Z$ is in $\mathscr{V}(A)$ if for some $i \in Z$ and $n \in N$ there is a vector $\boldsymbol{u} \in A^n$ such that

$$(1) \qquad E = \{u \in A^Z : (u_i, \ldots, u_{i+n-1}) = \boldsymbol{u}\} \,.$$

A cylinder in A^Z is any set of the form (1) with the singleton $\{\boldsymbol{u}\}$ replaced by an arbitrary subset of A^n. Alternatively, $\mathscr{A}^Z$ is the σ-field of all Borel subsets of A^Z relative to the product topology which makes A^Z a complete, separable, and metrizable space. A metric on A^Z can be defined, for example, by the formula

$$(2) \qquad d_A(u, u') = \begin{cases} \max \{(1 + |i|)^{-1} : i \in Z, u_i \neq u'_i\} & \text{if} \quad u \neq u' \,, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from the definition of the product topology in A^Z that any cylinder is simultaneously open and closed, hence A^Z is a totally disconnected space. We let T_A denote the shift-transformation in A^Z:

$$(3) \qquad (T_A u)_i = u_{i+1} \quad \text{for} \quad u \in A^Z , \quad i \in Z .$$

Then T_A is a homeomorphism of A^Z so that both T_A and T_A^{-1} are $\mathscr{A}^Z$-measurable. We let $\mathscr{I}(A)$ denote the σ-field of all invariant events, in symbols,

$$(4) \qquad \mathscr{I}(A) = \{ E \in \mathscr{A}^Z : T_A E = E \} .$$

Further, let $U = (U_i)_{i \in Z}$ denote the sequence of one-dimensional projections, viz.

$$(5) \qquad U_i(u) = u_i ; \quad u \in A^Z , \quad i \in Z .$$

If $i \in Z$, $n \in N$, and $u \in A^Z$, we put $u_i^n = (u_i, ..., u_{i+n-1})$ and $u^n = u_0^n$; the same applies to U.

Let $P(A)$ denote the set of all probability measures on $(A^Z, \mathscr{A}^Z)$. The weak topology on $P(A)$ is the weakest topology such that the map

$$\mu \mapsto \int_{A^Z} f \, d\mu : P(A) \to (-\infty, \infty)$$

is continuous for any bounded continuous function f on A^Z. Since A^Z is totally disconnected, a sequence $(\mu_n)_{n \in N}$ weakly converges to μ $(\mu, \mu_n \in P(A))$ if and only if

$$(6) \qquad \lim_{n \to \infty} \mu_n(V) = \mu(V) , \quad V \in \mathscr{V}(A) .$$

Denote by $M(A)$ the convex set of all T_A-invariant measures in $P(A)$, and by $E(A)$ the set of all extreme points of $M(A)$. Then $\mu \in E(A)$ if and only if μ is T_A-ergodic, i.e., if $\mu(E) \in \{0, 1\}$ for all $E \in \mathscr{I}(A)$ (cf. (4)). If $\mu \in M(A)$ then the whole structure will be denoted by $[A, \mu]$ or by $[A, \mu, U]$ and called a *stationary source*. The sequence U is the corresponding stationary process whose distribution is μ, in symbols:

$$(7) \qquad \text{dist}\,(U) = \mu .$$

If $\mu \in E(A)$ then the stationary source $[A, \mu]$ is called *ergodic*. The source $[A, \mu]$ is said to be *aperiodic* if μ is a non-atomic measure, i.e., if $\mu\{u\} = 0$ for all $u \in A^Z$.

Let $(B, \mathscr{B})$ be also a discrete countable alphabet. Any measurable map $\bar{\varphi} : A^Z \to B^Z$ is said to be a *code*. A code $\bar{\varphi}$ is called *stationary* if

$$(8) \qquad \bar{\varphi} \circ T_A = T_B \circ \bar{\varphi} .$$

A stationary code $\bar{\varphi} : A^Z \to B^Z$ gives rise to a measurable map $\varphi : A^Z \to B$, viz.

$$(9) \qquad \varphi(u) = \bar{\varphi}(u)_0 , \quad u \in A^Z ,$$

and the map φ induces a measurable partition of A^Z:

$$(10) \qquad \mathscr{P}_\varphi = \{ \varphi^{-1}\{b\} : b \in B \} , \quad \varphi^{-1}\{b\} \in \mathscr{A}^Z .$$

364

Conversely, let $\mathscr{P}$ be any countable measurable partition of A^Z. The formula

(11) $$\bar{\varphi}(u)_i = P \quad \text{if} \quad T_A^i u \in P\,, \quad P \in \mathscr{P}$$

defines a stationary code $\bar{\varphi} : A^Z \to \mathscr{P}^Z$ such that $\mathscr{P}_\varphi = \mathscr{P}$.

Next suppose that $[A, \mu, U]$ and $[B, \varkappa, V]$ (where dist $(V) = \varkappa$, see (5) and (6)) are two stationary sources. A stationary code $\bar{\varphi} : A^Z \to B^Z$ is said to be *perfectly noiseless* (with respect to $[A, \mu]$ and $[B, \varkappa]$) if $\varkappa = \mu\bar{\varphi}^{-1}$ and there exists a measurable set $A^* \subset A^Z$ with $\mu(A^*) = 0$ such that $\bar{\varphi}$ is one-to-one on $A^Z \setminus A^*$. It follows as in [12], Appendix A, that there exists a stationary code $\bar{\psi} : B^Z \to A^Z$ such that $\bar{\psi}(\bar{\varphi}u) = u$ and $\bar{\varphi}(\bar{\psi}v) = v$ almost everywhere. This justifies the clause perfectly noiseless for we then have an endocer/decoder pair $(\bar{\varphi}, \bar{\psi})$ such that

(12) $$\text{Prob}\,[U_0 \neq (\bar{\psi}(\bar{\varphi}U))_0] = 0\,.$$

Remark 1. In what follows we usually neglect sets of measure zero, e.g., we often understand relations like (8) as being valid only almost everywhere. For the same reasons we can consider an everywhere defined and almost everywhere one-to-one map equivalently as an almost everywhere defined and one-to-one map. Consequently, a perfectly noiseless code is but a mod 0 isomorphism between the dynamical systems $(A^Z, \mathscr{A}^Z, \mu, T_A)$ and $(B^Z, \mathscr{B}^Z, \varkappa, T_B)$. Recall from [1] that the two systems are *mod 0 isomorphic* if there exist sets $A^* \in \mathscr{A}^Z$, $B^* \in \mathscr{B}^Z$ and a map $\bar{\varphi} : A^* \to B^*$ such that

(i) $\mu(A^*) = \varkappa(B^*) = 1$, $\bar{\varphi}$ is injective;
(ii) A^* and B^* are invariant;
(iii) if $E \subset A^*$ then $E \in \mathscr{A}^Z$ if and only if $\bar{\varphi}E \in \mathscr{B}^Z$; in this case $\mu(E) = \varkappa(\bar{\varphi}E)$; and
(iv) $\bar{\varphi}(T_A u) = T_B(\bar{\varphi}u)$ for al $u \in A^*$

The concept of mod 0 isomorphism makes sense in much more general measure spaces [1]. In sequence spaces we shall use both concepts. As to the mod 0 nomenclature consult [15] or [14].

A countable measurable partition $\mathscr{P}$ of A^Z is said to be a *generator* (with respect to a given stationary source $[A, \mu]$) if

(13) $$\sigma\Big(\bigcup_{i \in Z} T_A^i\, \sigma(\mathscr{P})\Big) = \mathscr{A}^Z \bmod 0\,.$$

This says that the two σ-fields give rise to isomorphic measure algebras under μ; see [6]. It is well-known and easy to check that a stationary code $\bar{\varphi} : A^Z \to B^Z$ is perfectly noiseless if and only if the partition $\mathscr{P}_\varphi$ (cf. (10)) is a generator.

A stationary source $[B, \varkappa]$ is sais to be a *factor* of the stationary source $[A, \mu]$ if there is a stationary code $\bar{\varphi} : A^Z \to B^Z$ such that

(14) $$\varkappa(F) = \mu(\bar{\varphi}^{-1}F)\,, \quad F \in \mathscr{B}^Z\,.$$

If $\mu \in \mathbf{M}(A)$ is a product measure, the source $[A, \mu]$ is called *memoryless*. Any factor of a memoryless source is called a *Bernoulli source*. Observe that a Bernoulli

source is an information theoretic counterpart of a *generalized Bernoulli shift*, for we allow countable infinite alphabets as well (see § 9 of [12] or [17] for Ornstein theory in case of countable partitions).

2. ERGODIC DECOMPOSITION AND ASSOCIATED LEBESGUE SPACES

Let $[A, \mu]$ be a stationary source over a countable alphabet A. Recall from [18] that a point $u \in A^Z$ is *regular* if there is a measure $\mu_u \in E(A)$ uniquely determined by u via the relations

$$(15) \qquad \mu_u(V) = \lim_{n \to \infty} n^{-1} \sum_{j=0}^{n-i} I_V(T_A^j u) ; \quad V \in \mathscr{V}(A) ,$$

where I_V stands the indicator function of the set V. The set R_A of all regular points is measurable, T_A-invariant, and $\mu(R_A) = 1$ for any $\mu \in M(A)$. The *ergodic decomposition* says that any $\mu \in M(A)$ has (mod 0) unique representation in the form

$$(16) \qquad \mu(E) = \int_{R_A} \mu_u(E) \, \mu(du) , \quad E \in \mathscr{A}^Z .$$

The uniqueness follows from essential uniqueness of regular conditional probabilities relative to the σ-field $\mathscr{I}(A)$. Put

$$(17) \qquad R_A(u) = \{u' \in R_A : \mu_{u'} = \mu_u\} ; \quad u \in R_A .$$

Then

$$(18) \qquad \mu_{u'}(R_A(u)) = \begin{cases} 1 & \text{if } u' \in R_A(u) \text{ (and hence } u \in R_A(u')), \\ 0 & \text{otherwise.} \end{cases}$$

It follows that an ergodic source $[A, \mu_u, U^u]$ is a source such that its process $U^u = (U_i^u)_{i \in Z}$ is allowed to have trajectories only inside of the set $R_A(u)$. Thus, the set of trajectories of the process U which has the property that $\text{dist}(U) \in M(A) \setminus E(A)$ splits into mod 0 mutually exclusive sets of trajectories corresponding to the ergodic processes U^u. In particular, U can have a trajectory which meets two or more different sets $R_A(u)$ only with probability zero.

This simple observation is crucial to all what will follow. On the other hand, (17) shows there may exist many regular points giving rise to the same measure μ_u (e.g., if $u' \in R_A(u)$ then $T_A^i u' \in R_A(u)$ for all $i \in Z$. Less trivially, if $u \in R_A$ and u' differs from u on a subset of Z having density zero, then $u' \in R_A$ and even $u' \in R_A(u)$; see (15)).

Hence, it seems useful to factorize somehow the set R_A. This can be done in many ways. One possibility is to use (15) and imbed the set R_A into $[0, 1]^N$ via the map $u \mapsto (\mu_u(V_n); n \in N)$, where $\mathscr{V}(A) = \{V_1, V_2, \ldots\}$. This way is convenient when one wants to preserve topological properties. However, our desire is to preserve the structure of ergodic decomposition. This can be done quite easily within the framework of Lebesgue spaces [15], and causes only minor additional difficulties.

So, let $\left(\mathscr{F}_m^A, m^A\right)$ denote the completion of A^Z relative to the measure $\mu \in \mathbf{M}(A)$ [6]. Then $\left(A^Z, \mathscr{F}_m^A, m^A\right)$ is a Lebesgue space. Moreover, if the source $[A, \mu]$ is aperiodic then this Lebesque space becomes isomorphic to the Lebesgue space $(I, \mathscr{L}, \lambda)$ of the unit interval $I = [0, 1]$. The family

(19)
$$\varrho_A = \{R_A(u) : u \in R_A\} \cup \{A^Z \smallsetminus R_A\}$$

is a Rohlin measurable partition (cf. [14, 15]; we use the term Rohlin measurable in order to point out the difference with ordinary measurable partitions which have been considered in the preceding section). In particular, the factor space $\left(Q^A, \mathscr{B}^A, m_0^A\right)$ of $\left(A^Z, \mathscr{F}_m^A, m^A\right)$ relative to ϱ_A is again a Lebesgue space, and there exists a canonical system of measures $\left(m_\xi; \xi \in Q^A\right)$ associated with ϱ_A. Our choice of ϱ_A entails that

(20)
$$m_\xi = \mu_u \quad \text{on} \quad \xi \cap \mathscr{F}_m^A \quad \text{for} \quad u \in \xi, \quad \xi \in Q^A .$$

Moreover, the ergodic decomposition (16) assumes on the canonical form

(21)
$$m^A(\tilde{E}) = \int_{Q^A} m_\xi(\tilde{E} \cap \xi) \, m_0^A(\mathrm{d}\xi), \quad \tilde{E} \in \mathscr{F}_m^A .$$

We shall call the elements $\xi \in Q^A$ the *ergodic fibres* of the source $[A, \mu]$ and reserve the name *ergodic components* for the sources $[A, \mu_u]$, $u \in R_A$. We shall use both descriptions of the structure of ergodic decomposition in what follows.

The following representation of the fibre structure was established by Rohlin (see [15, § 4.3], a similar representation as given also by Oxtoby [13]). We include it for purpose of later reference.

Lemma 1. Let $[A, \mu]$ be a stationary aperiodic source over a countable alphabet. Then there exists a mod 0 isomorphism g^A from the product space $\left(Q^A, \mathscr{B}^A, m_0^A\right) \times$ $\times (I, \mathscr{L}, \lambda)$ into $\left(A^Z, \mathscr{F}_m^A, m^A\right)$ which sends $\{\xi\} \times I$ onto ξ for all $\xi \in Q^A$.

In what follows, we let $B^Z, \mathscr{B}^Z, R_B, \varrho_B$ etc., denote the corresponding objects for the alphabet $(B, \mathscr{B})$. Also, v and η serve as generic symbols for elements of B^Z and Q^B, respectively. Finally, Γ^A and Γ^B designate the natural projections which send almost all points from A^Z and B^Z into ergodic fibres containing them.

3. A LOCAL STRUCTURE IS NECESSARY

Throughout this section we are given two fixed stationary aperiodic sources $[A, \mu]$ and $[B, \varkappa]$. The main result is that any perfectly noiseless code between the two sources has a "local" structure in the following sense:

Theorem 1. Let $[A, \mu]$ and $[B, \varkappa]$ be two mod 0 isomorphic stationary and aperiodic sources, and let $\bar{\varphi}: A^Z \to B^Z$ denote the associated perfectly noiseless code.

(a) For m_0^A-almost all $\xi \in Q^A$ there is a unique $\eta \in Q^B$, and for m_0^B-almost all $\eta \in Q^B$ there is a unique $\xi \in Q^A$ such that $\bar{\varphi}(\xi) = \eta$.

(b) Let us put for $\xi \in Q^A \bmod m_0^A$ and for $u \in \xi \bmod m_\xi$,

$$\bar{\varphi}_\xi(u) = \bar{\varphi}(u) \, . \tag{22}$$

If $\xi \in Q^A$ is fixed and $\eta \in Q^B$ corresponds to ξ by assertion (a) (this takes place for almost all ξ) then $\bar{\varphi}_\xi$ is a $\bmod 0$ isomorphism between the ergodic sources $[A, m_\xi]$ and $[B, m_\eta]$.

(c) Let $F: Q^A \times A^Z \to B^Z$ be almost everywhere defined by the property that

$$F(\xi, u) = \bar{\varphi}_\xi(u) \, . \tag{23}$$

Then F is $(\mathcal{B}^A \times \mathcal{F}_m^A, \mathcal{F}_m^B)$-measurable.

Remark 2. In assertion (b), the maps ψ_ξ can be considered either as Borel measurable or as measurable with respect to completed Borel σ-fields. Standard arguments show that each Borel measurable isomorphism can be $\bmod 0$ uniquely extended to an isomorphism relative to the completed σ-fields. Conversely, the latter isomorphism, restricted to Borel sets of full measure remains measurable with respect to Borel σ-fields.

Before proving the theorem, let us prove two auxiliary lemmas.

Lemma 2. With probability one, $\bar{\varphi}$ maps regular points into regular points; in symbols:

$$\mu\{u \in R_A : \bar{\varphi}(u) \in R_B\} = 1 \, . \tag{24}$$

Proof. We have that $\{u \in R_A : \bar{\varphi}(u) \in R_B\} = R_A \cap \bar{\varphi}^{-1} R_B$. Since $R_A \in \mathscr{I}(A)$ and $R_B \in \mathscr{I}(B)$, $R_A \cap \bar{\varphi}^{-1} R_B \in \mathscr{I}(A)$ for $\bar{\varphi}$ is a stationary code. (More precisely, we can find a set $E \in \mathscr{I}(A)$ such that $\mu((R_A \cap \bar{\varphi}^{-1} R_B) \Delta E) = 0$, where Δ stands for symmetric difference, because $\bar{\varphi}$ satisfies (8) only $\bmod 0$. However, arguments of this kind will be frequently omitted.) If $u \in R_A$ then $\mu_u \in E(A) \subset M(A)$ so that $\mu_u(R_A) = 1$. Since $\mu_u \in M(A)$, $\mu_u \bar{\varphi}^{-1} \in M(B)$; in particular, we have that $\mu_u(\bar{\varphi}^{-1} R_B) = 1$. Consequently,

$$\mu_u(R_A \cap \bar{\varphi}^{-1} R_B) = 1 \, , \quad u \in R_A \, ,$$

and an application of (16) concludes the proof. $\qquad \square$

Lemma 3.

$$\mu\{u \in R_A : \mu_u \bar{\varphi}^{-1} = \mu_{\bar{\varphi}(u)}\} = 1 \, . \tag{25}$$

Proof. By Lemma 2, $\bar{\varphi}(u) \in R_B$ for μ-almost all $u \in R_A$. Hence, for μ-almost all u, the measure $\mu_{\bar{\varphi}(u)}$ is well-defined and ergodic. Let $V \in \mathscr{V}(B)$. Using the individual ergodic theorem and stationarity of $\bar{\varphi}$ it follows that $\mu_{\bar{\varphi}(u)}(V) = \mu_u(\bar{\varphi}^{-1} V)$. Hence

$$\mu\{u \in R_A : \mu_{\bar{\varphi}(u)}(V) = \mu_u \bar{\varphi}^{-1}(V)\} = 1 \, ,$$

and since $\mathscr{V}(B)$ is countable, (25) follows. $\qquad \square$

Proof of Theorem 1. E_0 denote the set of all regular points $u \in R_A$ with the following property: there exist points u', $u'' \in R_A(u)$ and points v', $v'' \in R_B$ with $v' \notin R_B(v'')$ such that $v' = \bar\varphi(u')$ and $v'' = \bar\varphi(u'')$. We claim that $\mu(E_0) = 0$. Indeed, if $\mu(E_0) > 0$, we can pick u', u'', v' and v'' as above. Since u', $u'' \in R_A(u)$, $\mu_{u'} = \mu_{u''}$. Hence $\mu_{u'}\bar\varphi^{-1} = \mu_{u''}\bar\varphi^{-1}$, i.e. $\mu_{v'} = \mu_{\bar\varphi(u')} = \mu_{\bar\varphi(u'')} = \mu_{v''}$, and this contradicts the fact that $v' \notin R_B(v'')$. In a similar way, we can prove that $\varkappa(F_0) = 0$, where $F_0 \subset R_B$ is defined exactly as E_0 was. Hence $\mu(A^Z \setminus E_0) = \varkappa(B^Z \setminus F_0) = 1$. Using canonical representations of m^A and m^B (cf. (21) and (20)) we can translate the above conclusions into the setting of associated factor spaces, and thereby obtain conclusion (a) of Theorem 1.

In order to prove (b) suppose that $\bar\varphi_\xi$ is defined by (22). Since $\xi \in Q^A$ we find a $u^0 \in R_A$ so that $\xi = R_A(u^0)$. Let $v^0 \in R_B$ be a point such that $\mu_{v^0} = \mu_{u^0}\bar\varphi^{-1}$. With probability one, $\bar\varphi(u) \in \eta = R_B(v^0)$ for all $u \in \xi = R_A(u^0)$ (see Lemmas 2 and 3). Moreover, $m_\xi = \mu_{u^0}$ on $\xi \cap \mathscr{F}_m^A$, $m_\eta = \mu_{v^0}$ on $\eta \cap \mathscr{F}_m^B$. Let $\bar\varphi$ denote a $(\mathscr{F}_m^A, \mathscr{F}_m^B)$-measurable extension of the isomorphism $\bar\varphi$ (see Remark 2). If $\bar\varphi_\xi : \xi \to \eta$ is defined using this version of $\bar\varphi$, we get that $\bar\varphi_\xi$ is measurable and $m_\xi\bar\varphi_\xi^{-1} = m_\eta$. We claim that

$$m_0^A\{\xi \in Q^A : \bar\varphi_\xi \text{ is } m_\xi\text{-a.e. injective}\} = 1 .$$

In fact, let $\tilde E \in \mathscr{F}_m^A$ be the set on which the above version of $\bar\varphi$ is injective; $m^A(\tilde E) = 1$. Using (21) and observing that $m_\xi(\tilde E \cap \xi) \leqq 1$ for all ξ, we see that

$$m_0^A\{\xi \in Q^A : m_\xi(\tilde E \cap \xi) = 1\} = 1 .$$

But $\tilde E \cap \xi$ is the set on which $\bar\varphi_\xi$ is injective, and this proves the claim. Since $\bar\varphi$ was stationary and all ergodic fibres ξ and η, considered as subsets of sequence spaces, are invariant, $\bar\varphi_\xi$ is stationary (for almost all fibres $\xi \in Q^A$). This proves (b).

Now let us prove (c). Pick $V \in \mathscr{V}(B)$ and consider the set $F^{-1}V = \{(\xi, u) : \bar\varphi_\xi(u) \in V\}$. Since $\bar\varphi_\xi$ is defined (almost everywhere) on ξ, we have the formula

$$(26) \qquad F^{-1}V = \left[\bigcup_{\xi \in Q^A}(\{\xi\} \times \xi)\right] \cap (Q^A \times \bar\varphi^{-1}V)$$

valid mod 0. The set $Q^A \times \bar\varphi^{-1}V$ is measurable (to this end consider again the version of $\bar\varphi$ introduced in the proof of assertion (b)) so that it remains to prove measurability of the set

$$G = \bigcup_{\xi \in Q^A}(\{\xi\} \times \xi) .$$

If Q^A is (mod 0) countable then there is nothing to prove. So suppose Q^A uncountable, and let E^c designate the complement of a set E. We shall prove that G^c is measurable. First observe that

$$(27) \qquad m^A\left(\bigcup_{\xi \in Q^A}\{u \in R_A : u \in \xi\}\right) = \mu(R_A) = 1 .$$

Next

$$G^c = \bigcap_{\xi \in Q^A}(\{\xi\} \times \xi)^c .$$

By (27) we see that $\big(\{\xi\} \times \xi\big)^c$ is a null set for almost all $\xi \in Q^A$ so that G^c is a subset of a null set in a Lebesgue space. Consequently, G^c, hence G itself, must be measurable, and conclusion (c) follows from this and from (26). $\qquad\square$

Remark 3. One can wonder why we are interested in joint measurability of the map $(\xi, u) \to \bar{\varphi}_\xi(u)$. We shall see in the next section that this property is crucial for a construction of global codes. The point is that we shall deal with measures transported by a family of local codes, not by a single global code.

4. THE LOCAL STRUCTURE IS SUFFICIENT

This section contains the abstract kernel of an idea due to Winkelbauer [19] who used it for the purpose of extension of Krieger's codes to the aperiodic non-ergodic case. The main result shows that the local structure derived in Theorem 1 is also sufficient for existence of a global perfectly noiseless code, and the proof will proceed via a measure theoretic construction.

Theorem 2. Let $[A, \mu]$ and $[B, \varkappa]$ be two aperiodic stationary sources over countable discrete alphabets. Suppose that

(i) for m_0^A-almost all fibres $\xi \in Q^A$, the map $\bar{\varphi}_\xi : \xi \to \eta$, $\eta = \bar{\varphi}_\xi(\xi)$, is a mod 0 isomorphism between $[A, m_\xi]$ and $[B, m_\eta]$, and

(ii) the map $(\xi, u) \mapsto \bar{\varphi}_\xi(u) : Q^A \times A^Z \to B^Z$ is $(\mathcal{B}^A \times \mathcal{F}_m^A, \widehat{\mathcal{F}_m^B})$-measurable.

Then there exists a stationary code $\bar{\varphi} : A^Z \to B^Z$ which is perfectly noiseless with respect to $[A, \mu]$ and $[B, \varkappa]$ (and such that its local components are precisely $(\bar{\varphi}_\xi; \xi \in Q^A)$).

P r o o f. Let E_0 be the set constructed above in the proof of Theorem 1. If $Q_0 = = \Gamma^A(R_A \smallsetminus E_0)$, then $m_0^A(Q_0) = 1$ (recall that Γ^A is the natural projection). For if $m_0^A(Q_0) < 1$ then for fibres $\xi \in A^A$ forming a set of positive m_0^A measure, $\bar{\varphi}_\xi$ would map ξ into a set intersecting at least two different ergodic fibres from Q^B. Hence, $\bar{\varphi}_\xi(\xi)$ and therefore ξ itself would carry at least two mutually singular ergodic measures, a contradiction. If $\xi \in Q_0$ and $u \in \xi$, we put

$$(28) \qquad\qquad \psi\xi = \{v \in R_B : \mu_v = \mu_u \bar{\varphi}_\xi^{-1}\}.$$

ψ is well-defined; that is, it does not depend on the particular choice of $u \in \xi$. We now use assumption (ii) which ensures that for any $F \in \mathcal{B}^Z$ the map

$$\xi \mapsto m_\xi\big(\xi \cap \bar{\varphi}_\xi^{-1}F\big)$$

is $\mathcal{B}^A$-measurable so that we can define a measure $\varkappa'$ on $\mathcal{B}^Z$ by the properties that

$$(29) \qquad\qquad \varkappa'(F) = \int_{Q^A} m_\xi\big(\xi \cap \bar{\varphi}_\xi^{-1}F\big) m_0^A(\mathrm{d}\xi)\,, \quad F \in \mathcal{B}^Z\,.$$

Using (21) and (20) for $[B, \varkappa]$ we see that the completion of $\mathscr{B}^Z$ with respect to $\varkappa'$ is but $(\mathscr{F}_m^B, m^B)$. Let $m_0^B = m_0^A \psi^{-1}$. A slight generalization of the fact that, in Lebesgue spaces, and almost everywhere injective measurable map is invertible (see [19, Lemma 5.3]) shows that ψ is a mod 0 isomorphism between the spaces $(Q^A, \mathscr{B}^A, m_0^A)$ and $(Q^B, \mathscr{B}^B, m_0^B)$. Moreover, (28) and the definition of $(m_\eta; \eta \in Q^B)$ show that m_0^B is the right measure (i.e., that one induced by $\varkappa$), for

$$m^B(\tilde{F}) = \int_{Q^B} m_\eta(\tilde{F} \cap \eta)\, m_0^B(\mathrm{d}\eta)\,, \quad \tilde{F} \in \mathscr{F}_m^B\,.$$

At the same time, because of aperiodicity Lemma 1 applies and yields mod 0 isomorphisms $g^A : Q^A \times I \to A^Z$ and $g^B : Q^B \times B^Z$. Observe that until now all isomorphisms concern only probability spaces, not the transformations acting on them. The final part of the proof is devoted to showing how g^A, g^B, and ψ can be used in order to get a stationary coding. Put (mod 0)

$$\bar{\psi}(\eta, t) = (\psi^{-1}\eta, \tilde{\varphi}_\eta(t))\,, \quad \eta \in Q^B\,, \quad t \in I\,,$$

where

$$\tilde{\varphi}_\eta(t) = (g_\xi^A)^{-1}\, \bar{\varphi}_\xi^{-1}[g_\eta^B(t)]\,,$$

and g_ξ^A, g_η^B denote the sections of g^A and g^B, respectively. It is easy to check that $\bar{\psi}$ is a mod 0 isomorphism. Consequently, the map $g^A \circ \bar{\psi} \circ (g^B)^{-1}$ is a mod 0 isomorphism between the Lebesgue spaces $(B^Z, \mathscr{F}_m^B, m^B)$ and $(A^Z, \mathscr{F}_m^A, m^A)$. Put

$$\bar{\varphi} = [g^A \circ \bar{\psi} \circ (g^B)^{-1}]^{-1} \mid R_A \setminus E_0\,,$$

where $R_A \setminus E_0$ is considered as a measurable space relative to the σ-field of all its Borel subsets. Then the set $\bar{\varphi}(R_A \setminus E_0)$ is a Borel set of full measure in B^Z, $\bar{\varphi}$ is (Borel) measurable, and $\varkappa = \mu\bar{\varphi}^{-1}$. By construction, the relation

$$\bar{\varphi}_\xi \circ T_A = T_B \circ \bar{\varphi}_\xi$$

is valid mod 0 with respect to m_ξ, and for all ergodic fibres relative to the measure m_0^A. Using (21) we see that $\bar{\varphi} \circ T_A = T_B \circ \bar{\varphi}$ mod 0, and the proof is complete.

5. FINITARY PERFECTLY NOISELESS CODES

Extensions of Ornstein theory to classical (differentiable) dynamical systems, as reported in [12], make it possible to investigate ergodic properties of many mechanical systems. These physical applications provide strong evidence in favour of codes enjoying also certain continuity properties (see the discussion at the end of [5] and related results in [2]).

Given two stationary sources $[A, \mu]$ and $[B, \varkappa]$, a perfectly noiseless stationary code is said to be *finitary*, if it is a mod 0 homeomorphism. This means (see Remark 1) that there exist measurable sets A^* and B^* and a map $\bar{\varphi} : A^* \to B^*$ as in the definition

of the mod 0 isomorphism with the additional property that $\bar{\varphi}$ is a homeomorphism (this makes sense for A^* and B^* are Borel subsets of A^Z and B^Z, respectively). As already pointed out, any stationary code $\bar{\varphi} : A^Z \to B^Z$ can be considered as a sequential code, i.e.,

$$(30) \qquad \bar{\varphi}(u)_i = \varphi(T_A^i u) ; \quad u \in A^Z , \quad i \in Z ,$$

where $\bar{\varphi}$ and φ are connected by (9). In general, the value $\varphi(u)$, and hence also the values $\bar{\varphi}(u)_i$, depend on the entire sequence u. However, if $\bar{\varphi}$ is mod 0 continuous then for almost all $u \in A^Z$, the code length is finite. Indeed, if we want to determine $v = \bar{\varphi}(u)$, we have to look at the coordinates $u_{-n}, u_{-n+1}, \ldots, u_0, \ldots, u_n$ for successive $n = 0, 1, 2, \ldots$ until we find a sequence such that

$$\{u' \in A^Z : u'_i = u_i, |i| \leq n\} \subset \bar{\varphi}^{-1}\{v \in B^Z : v_0 = b\}$$

for some $b \in B$. Due to the continuity assumption, the latter set is mod 0 a countable union of cylinders so that the above will happen at some finite n (of course, n depends on u) for almost all $u \in A^Z$. Then we can set $\bar{\varphi}(u)_0 = b$ and by shifting the procedure we can determine $\bar{\varphi}(u)_i$ for all $i \in Z$.

If $\bar{\varphi}$ is a mod 0 homeomorphism, exactly the same reasoning applies to $\bar{\varphi}^{-1}$. Thus, a perfectly noiseless code $\bar{\varphi} : A^Z \to B^Z$ is finitary if and only if the 0-th coordinate $\bar{\varphi}(u)_0$ (and $\bar{\varphi}^{-1}(v)_0$) can be decided by knowing only some finite segment of u (of v), and this happens for μ-almost all $u \in A^Z$ (for $\varkappa$-almost all $v \in B^Z$). Again, we shall call such a code alternatively a finitary isomorphism.

Now suppose that $[A, \mu]$ and $[B, \varkappa]$ are two aperiodic stationary sources which are finitarily isomorphic, and let $\bar{\varphi} : A^Z \to B^Z$ denote the corresponding code. By Theorem 1, $\bar{\varphi}$ has a local structure. The question is what can be said about that local structure under the additional knowledge that $\bar{\varphi}$ is finitary. Let us start with two auxiliary assertions.

Lemma 4. The map $u \to \mu_u : R_A \to \mathbf{E}(A)$ is continuous with respect to the topology induced on R_A by the metric d_A (see (2)) and weak topology on $\mathbf{E}(A)$ (see (6)).

Proof. Suppose that $u^{(n)}, u$ are in R_A such that

$$\lim_{n \to \infty} d_A(u^{(n)}, u) = 0 .$$

Put $j(n) = \min \{|i| : i \in Z, \ u_i^{(n)} \neq u_i\}$. It follows from the definition of d_A that $\lim j(n) = \infty$. We have to prove that for any $V \in \mathscr{V}(A)$,

$$(31) \qquad \lim_{n \to \infty} \mu_{u^{(n)}}(V) = \mu_u(V) .$$

So fix $V \in \mathscr{V}(A)$. Since V depends only on finitely many coordinates, the fact that $\lim j(n) = \infty$ implies that the difference $|\mu_{u^{(n)}}(V) - \mu_u(V)|$ becomes as close to zero as we please if n is large enough. $\qquad\square$

Lemma 5. For any $u \in R_A$, the set $R_A(u)$ (cf. (17)) is closed.

Proof. We have that

(32) $$R_A(u) = \bigcap_{V \in \mathscr{V}(A)} \{u' \in R_A : \mu_{u'}(V) = \mu_u(V)\} \, .$$

For fixed V, 1_V is continuous as V is open and closed. By Lemma 4, the map $\tilde{\mu}_V :$ $: u \to \mu_u(V)$ from R_A into $[0, 1]$ is continuous. But

$$\{u' \in R_A : \mu_{u'}(V) = \mu_u(V)\} = \tilde{\mu}_V^{-1}\{t\} \, , \quad t = \mu_u(V) \, .$$

Hence, $R_A(u)$ is closed, in fact, owing to (32) it is an intersection of closed sets. $\square$

Lemma 5 is all what we need in order to prove a finitary version of Theorem 1.

Theorem 3. Suppose $[A, \mu]$, $[B, \varkappa]$, and $\bar{\varphi}$ are as in Theorem 1. If $\bar{\varphi}$ is finitary then so are the local isomorphisms $\bar{\varphi}_\xi$, $\xi \in Q^A$.

Proof. Let $E_0 \in \mathscr{A}^Z$ and $F_0 \in \mathscr{B}^Z$ be such that $\mu(E_0) = \varkappa(F_0) = 1$ and $\bar{\varphi}$ is a homeomorphism between E_0 and F_0. We can and we do assume that $E_0 \subset R_A$ and $F_0 \subset R_B$. Now

$$1 = \mu(E_0) = \int_{R_A} \mu_u(E_0) \, \mu(du) = \varkappa(F_0) = \int_{R_B} \mu_v(F_0) \, \varkappa(dv)$$

so that

$$\mu\{u \in R_A : \mu_u(E_0 \cap R_A(u)) = 1\} = 1 \, ,$$

$$\varkappa\{v \in R_B : \mu_v(F_0 \cap R_B(v)) = 1\} = 1 \, .$$

If $v \in R_B$ is $\bar{\varphi}(u)$ for some $u \in R_A$, then $\bar{\varphi}(R_A(u)) = R_B(v)$. Fix u, v in this manner. By Lemma 5, the set $E_0 \cap R_A(u)$ is closed in E_0 (with respect to the relative topology on E_0). Since a restriction of a continuous map defined on a topological space is continuous with respect to the relative topology, $\bar{\varphi} \mid R_A(u)$ is almost everywhere continuous. Similarly, $\bar{\varphi}^{-1} \mid R_B(v)$ is almost everywhere continuous, and

$$\bar{\varphi}(E_0 \cap R_A(u)) = \bar{\varphi}(E_0) \cap \bar{\varphi}(R_A(u)) = F_0 \cap R_B(v) \, ,$$

$$\bar{\varphi}^{-1}(F_0 \cap R_B(v)) = \bar{\varphi}^{-1} F_0 \cap \bar{\varphi}^{-1} R_B(v) = E_0 \cap R_A(u) \, .$$

Thus, $\bar{\varphi}$ considered as a map from $R_A(u)$ to $R_B(v)$, $v = \bar{\varphi}(u)$, is a mod 0 homeomorphism with respect to measures μ_u and μ_v. It remains to translate these conclusions into the setup of associated Lebesgue spaces. But this is a straight forward task so that we omit this step. $\square$

We do not know whether it is possible to modify the construction performed in Section 4 in order to get a finitary version of Theorem 2. On the other hand, using Theorem 2, it is quite easy to give an existence proof.

Theorem 4. Let $[A, \mu]$ and $[B, \varkappa]$ be two aperiodic stationary sources over countable discrete alphabets. Suppose that

(i) for m_0^A-almost all fibres $\xi \in Q_A$, the map $\bar{\varphi}_\xi : \xi \to \eta$ $\eta = \bar{\varphi}_\xi(\xi)$, is a finitary isomorphism between $[A, m_\xi]$ and $[B, m_\eta]$; and
(ii) the map $(\xi, u) \mapsto \bar{\varphi}_\xi(u): : Q^A \times A^Z \to B^Z$ is $(\mathscr{B}^A \times \mathscr{F}_m^A, \mathscr{F}_m^B)$-measurable.
Then there exists a perfectly noiseless finitary code $\bar{\varphi} : A^Z \to B^Z$.

Proof. Suppose, contrary to the conclusion, there is no finitary perfectly noiseless code between $[A, \mu]$ and $[B, \varkappa]$. Thus, for any isomorphism $\bar{\varphi}$ (which exists by Theorem 2) we can find either a set $G \in \mathscr{A}^Z$ of positive measure on which $\bar{\varphi}$ is not continuous or a set $H \in \mathscr{B}^Z$ of positive measure on which $\bar{\varphi}^{-1}$ is not continuous. Let us consider the first case. Then (16) applied to G yields

$$\mu\{u \in R_A : \mu_u(G) > 0\} > 0 .$$

Let $u \in R_A$ be such that $\mu_u(R_A)(u) \cap G) = \mu_u(G) > 0$. Let $\xi = R_A(u)$ and $\eta = R_B(\bar{\varphi}_\xi(u))$. By assumption, $\bar{\varphi}_\xi$ is mod 0 continuous. Hence, if $\tilde{G} \in \xi \cap \mathscr{F}_m^A$ corresponds to $G \cap R_A(u)$, then $m_\xi(\tilde{G}) = 0$. But

$$m_\xi(\tilde{G}) = \mu_u(G \cap R_A(u)) > 0 ,$$

a contradiction. Consequently, there must be a perfectly noiseless code $\bar{\varphi}$ which is also mod 0 continuous. Let E_0 be the set on which $\bar{\varphi}$ is both continuous and injective. Then $\mu(E_0) = 1$ and so $\varkappa(\bar{\varphi}(E_0)) = 1$, too. Repeating the above reasoning we see that $\bar{\varphi}^{-1} : \bar{\varphi}(E_0) \to E_0$ is also continuous. In other words, $\bar{\varphi}$ is a mod 0 homeomorphism. □

REFERENCES

[1] P. Billingsley: Ergodic Theory and Information. Wiley, New York 1965.
[2] R. Bowen: Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. (Lecture Notes in Mathematics 470.) Springer - Verlag, Berlin—Heidelberg—New York 1975.
[3] M. Denker: Finite generators for ergodic, measure-preserving transformations. Z. Wahrsch. verw. Gebiete 29 (1974), 45—55.
[4] M. Denker, Ch. Grillenberger, and K. Sigmund: Ergodic Theory on Compact Spaces. (Lecture Notes in Mathematics 527.) Springer - Verlag, Berlin—Heidelberg—New York 1976.
[5] M. Denker and M. Keane: Almost topological dynamical systems. Israel J. Math. 34 (1979), 139—160.
[6] P. R. Halmos: Measure Theory. D. Van Nostrand, Princeton, N. J. 1950.
[7] M. Keane and M. Smorodinsky: Bernoulli schemes of the same entropy are finitarily isomorphic. Ann. of Math. 109 (1979), 387—406.
[8] J. C. Kieffer: Zero-error stationary coding over stationary channels. Z. Wahrsch. verw. Gebiete 56 (1981), 113—126.
[9] J. C. Kieffer and M. Rahe: Selecting universal partitions in ergodic theory. Ann. Probab. 9 (1981), 705—709.
[10] W. Krieger: On entropy and generators of measure-preserving transformations. Trans. Amer. Math. Soc. 149 (1970), 453—464, Erratum, ibid. 168 (1972), 519.
[11] D. S. Ornstein: Bernoulli shifts with the same entropy are isomorphic. Adv. in Math. 4 (1970), 337—352.

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

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

[14] W. Parry: Entropy and Generators in Ergodic Theory. W. A. Benjamin, New York—Amsterdam 1969.

[15] V. A. Rohlin: On the basic concepts of measure theory (in Russian). Matem. Sborn. *25 (67)* (1949), 107—150.

[16] P. C. Shields: Stationary coding of processes. IEEE Trans. Inform. Theory *IT-25* (1979), 293—291.

[17] M. Smorodinsky: Ergodic Theory, Entropy. (Lecture Notes in Mathematics 214.) Springer - Verlag, Berlin—Heidelberg—New York 1971.

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

[19] K. Winkelbauer: On the existence of finite generators for invertible measure-preserving transformations. Comment. Math. Univ. Carolinae *18* (1977), 789—812.

RNDr. Štefan Šujan, CSc., 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.