KYBERNETIKA ČÍSLO 4, ROČNÍK 2/1966

The Synchronization Problem of Information Theory

IGOR VAJDA

In the paper a formulation of the synchronization problem is given. There are shown some simple characteristic properties of synchronizing block encoders. Further the paper treats with a rate of convergence of a probability of error associated with synchronizing block encoders provided the channel is memoryless and source independent.

1. INTRODUCTION

Some authors working in the algebraic coding theory studied the synchronization problem which rises up under using of block encoding procedures. Namely, if n-tuples of consecutive letters of a message are encoded to the p-tuples of letters of the input alphabet of a noiseless channel, then the receiver obtaining a sequence of consecutive letters of output alphabet of the channel cannot a priori split up this sequence into p-tuples corresponding to the p-tuples sent over the channel and, consequently, he cannot use the appropriate decoding procedure. If we assume that the communication channel under consideration is the unique communication channel between the sender and the receiver, the encoders satisfying some synchronization requirements must be used. The questions rising up in this case may be sumarized as the synchronization problem of information theory.

The synchronization problem provided the channel is a noisy one was first given and studied by J. Nedoma, [2]. It follows from [2] that in this case the algebraic methods must be replaced by suitable statistical methods. The aim of this paper is to define the synchronisation problem of information theory under general assumptions concerning encoding and decoding procedures as well as sources and channels. Most attention will be paid to the definition of a synchronizing encoder that is the basic concept concerning the problem. Particular attention will be devoted to certain properties of the synchronizing encoders which may be useful for a further study of the problem as well as for applications.

2. NOTATION AND PRELIMINARY DISCUSSION

Throughout the paper the set of all integers will be denoted by I and the set of all positive integers by I^+.

For every non-empty set X and for every $n \in I^+$ we shall define the sets X^n and $\mathfrak{X}$ by

(2.1) $$X^n = \bigotimes_{i=1}^{n} X_i\,, \quad \mathfrak{X} = \bigotimes_{i=-\infty}^{+\infty} X_i\,, \quad X_i = X \quad \text{for all} \quad i \in I\,,$$

where $\otimes$ is the symbol for the Cartesian product. Elements of the sets $\mathfrak{X}$ and X^n will be denoted by $\mathfrak{x}$ and $\mathbf{x}$ respectively, i.e. $\mathfrak{x}$ denotes an infinite-dimensional vector with the i-th coordinate $(\mathfrak{x})_i \in X$, and $\mathbf{x}$ denotes and n-dimensional vector with the coordinates $(\mathbf{x})_1, (\mathbf{x})_2, \ldots, (\mathbf{x})_n$ belonging to X. In the entire paper we shall use the notation:

(2.2) $$(\mathfrak{x})_i^j = ((\mathfrak{x})_i, (\mathfrak{x})_{i+1}, \ldots, (\mathfrak{x})_j) \quad \text{for every} \quad \mathfrak{x} \in \mathfrak{X},\ i, j \in I,\ i \leqq j\,.$$

If $i, j \in I$, $i \leqq j$, then the σ-algebra generated by the class of all sets of the form

$$\{\mathfrak{x} : \mathfrak{x} \in \mathfrak{X}\,, \quad (\mathfrak{x})_i^j = \mathbf{x}\}\,, \quad \mathbf{x} \in X^{j-i+1}$$

will be denoted by $\mathscr{X}_i^j$, the σ-algebra generated by the class

$$\bigcup_{i \leqq j} \mathscr{X}_i^j$$

will be denoted by $\mathscr{X}$, and the σ-algebra generated by the class $\{\{\mathbf{x}\} : \mathbf{x} \in X^n\}$ will be denoted by $\mathscr{X}^n$.

For every non-empty set X and for every $j \in I$ we shall define the coordinate-shift transformation T^j of the set $\mathfrak{X}$ into itself by

(2.3) $$(T^j \mathfrak{x})_i = (\mathfrak{x})_{i+j} \quad \text{for every} \quad i \in I\,, \quad \mathfrak{x} \in \mathfrak{X}\,.$$

It is easy to see that T^j is a measurable transformation of the measurable space $(\mathfrak{X}, \mathscr{X})$ onto itself, for every $j \in I$.

An information source (or, briefly, a source) is described by specifying the following two elements:

(I) A non-empty alphabet C. The measurable space $(\mathfrak{C}, \mathscr{C})$ will be called the space of messages.
(II) A probability measure μ on $\mathscr{C}$.

A communication channel (or, briefly, a channel) is described by specifying the following two elements:

(I) Input alphabet $A \neq \emptyset$ and output alphabet $B \neq \emptyset$ of the channel. The measurable spaces $(\mathfrak{A}, \mathscr{A})$, $(\mathfrak{B}, \mathscr{B})$ will be called spaces of input and output signals of the channel.
(II) A function $\nu = \nu(E \mid \mathfrak{a})$ defined for all $\mathfrak{a} \in \mathfrak{A}$ and $E \in \mathscr{B}$ which is $\mathscr{A}$-measurable for evey fixed $E \in \mathscr{B}$ and is a probability measure on $\mathscr{B}$ for every fixed $\mathfrak{a} \in \mathfrak{A}$.

Since in the sequel the sets A, B, C will be assumed to be fixed, a source will be denoted briefly by μ and a channel by ν.

By saying "encoder $\tilde{\varphi}$" we shall understand the following two elements:

(I) A probability space $(\mathfrak{Y}, \mathscr{Y}, \tilde{\eta})$.

(II) A measurable transformation $\tilde{\varphi}$ of the space $(\mathfrak{C} \otimes \mathfrak{Y}, \mathscr{C} \otimes \mathscr{Y})$ into $(\mathfrak{A}, \mathscr{A})$ such that for every $\mathfrak{c} \in \mathfrak{C}$ the function $\tilde{\varphi}(\mathfrak{c}, \mathfrak{y})$ of $\mathfrak{y} \in \mathfrak{Y}$ is $\mathscr{Y}$-measurable.

The intuitive meaning of an encoder is such that the sender (situated between the source and the space of input signals) chooses for every realized message $\mathfrak{c} \in \mathfrak{C}$ an input signal $\mathfrak{a} = \tilde{\varphi}(\mathfrak{c}, \mathfrak{y})$ depending on the value of the random variable $\mathfrak{y} \in (\mathfrak{Y}, \mathscr{Y}, \tilde{\eta})$. The definition of a random encoder as it is given here is a natural generalisation of the usual definition of an encoder.

By saying "decoder $\tilde{\psi}$" we shall understand the following two elements:

(I) A probability space $(\mathfrak{Z}, \mathscr{Z}, \tilde{\xi})$.

(II) A measurable transformation $\tilde{\psi}$ of the space $(\mathfrak{B} \otimes \mathfrak{Z}, \mathscr{B} \otimes \mathscr{Z})$ into $(\mathfrak{C}, \mathscr{C})$ such that for every $\mathfrak{b} \in \mathfrak{B}$ the function $\tilde{\psi}(\mathfrak{b}, \mathfrak{z})$ of $\mathfrak{z} \in \mathfrak{Z}$ is $\mathscr{Z}$-measurable.

The intuitive meaning of a decoder $\tilde{\psi}$ is such that the receiver observing the sample space $(\mathfrak{B}, \mathscr{B})$ ot output signals takes a decision $\mathfrak{c}' \in \mathfrak{C}$ by the random decision procedure $\tilde{\psi}$. The definition of a random decoder as it is given here is a natural generalisation of the usual definition of the decoder.

By saying "(n, p)-encoder φ" for $n, p \in I^+$ we shall understand the following two elements:

(I) A probability space $(Y, \mathscr{Y}_*, \eta)$.

(II) A measurable transformation φ of the space $(C^n \otimes Y, \mathscr{C}^n \otimes \mathscr{Y}_*)$ into $(A^p, \mathscr{A}^p)$ such that for every $\mathfrak{c} \in C^n$ the function $\varphi(\mathfrak{c}, y)$ of $y \in Y$ is $\mathscr{Y}_*$-measurable.

By saying "(p, n)-decoder ψ" for $n, p \in I^+$ we shall understand the following two elements:

(I) A probability space $(Z, \mathscr{Z}_*, \xi)$.

(II) A measurable transformation ψ of the measurable space $(B^n \otimes Z, \mathscr{B}^n \otimes \mathscr{Z}_*)$ into $(C^n, \mathscr{C}^n)$ such that for every $\mathfrak{b} \in B^n$ the function $\psi(\mathfrak{b}, z)$ of $z \in Z$ is $\mathscr{Z}_*$-measurable.

For every (n, p)-encoder φ and (p, n)-decoder ψ we shall define an encoder $\tilde{\varphi}$ and decoder $\tilde{\psi}$ by

$$(2.4)\qquad (\mathfrak{Y}, \mathscr{Y}, \tilde{\eta}) = \bigotimes_{i=-\infty}^{+\infty} (Y_i, \mathscr{Y}_i, \eta_i), \quad (Y_i, Y_i, \eta_i) = (Y, \mathscr{Y}_*, \eta) \quad \text{for all} \quad i \in I,$$

$$(2.5)\qquad (\mathfrak{Z}, \mathscr{Z}, \tilde{\xi}) = \bigotimes_{i=-\infty}^{+\infty} (Z_i, \mathscr{Z}_i, \xi_i), \quad (Z_i, \mathscr{Z}_i, \xi_i) = (Z, \mathscr{Z}_*, \xi) \quad \text{for all} \quad i \in I,$$

$$(2.6)\qquad (\tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))_{pi+1}^{p(i+1)} = \varphi((\mathfrak{c})_{ni+1}^{n(i+1)}, (\mathfrak{y})_i) \quad \text{for every} \quad \mathfrak{c} \in \mathfrak{C}, \mathfrak{y} \in \mathfrak{Y}, \ i \in I,$$

$$(2.7)\qquad (\tilde{\psi}(\mathfrak{b}, \mathfrak{z}))_{ni+1}^{n(i+1)} = \psi((\mathfrak{b})_{pi+1}^{p(i+1)}, (\mathfrak{z})_i) \quad \text{for every} \quad \mathfrak{b} \in \mathfrak{B}, \ \mathfrak{z} \in \mathfrak{Z}, i \in I.$$

If $w(\mathfrak{c}, \mathfrak{c}')$ for $\mathfrak{c}, \mathfrak{c}' \in \mathfrak{C}$ is a $\mathscr{C} \otimes \mathscr{C}$-measurable non-negative real-valued function serving as a measure of the loss caused by taking decision $\mathfrak{c}'$, whereas $\mathfrak{c}$ is the transmitted message, then in case $\gamma(. \mid \mathfrak{c})$ is a conditional probability measure on $\mathscr{B}$ corresponding to the transmitted message $\mathfrak{c}$, the risk corresponding to $\mathfrak{c}$ and $\tilde{\psi}$ may be expressed as the average value of the loss corresponding to them, i.e. as

(2.8)
$$\int_{\mathfrak{B}} \int_{\mathfrak{Z}} w(\mathfrak{c}, \tilde{\psi}(\mathfrak{b}, \mathfrak{z}))\, d\xi(\mathfrak{z})\, d\gamma(\mathfrak{b} \mid \mathfrak{c}),$$

for every $\mathfrak{c}$ and for every decoder $\tilde{\psi}$.

Remark. In (2.8) there is evidently assumed that $\mathfrak{z} \in \mathfrak{Z}$ and $\mathfrak{b} \in \mathfrak{B}$ are independent random variables.

We shall say that the space of output signals is directly observable by the receiver if, for every encoder $\tilde{\varphi}$, the measure $\gamma(. \mid \mathfrak{c})$ on $\mathscr{B}$ is given by

(2.9)
$$\gamma(E \mid \mathfrak{c}) = \int_{\mathfrak{Y}} \gamma(E \mid \tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y}) \quad \text{for every} \quad E \in \mathscr{B}, \ \mathfrak{c} \in \mathfrak{C}.$$

In information theory it is customary to assume that the space of output signals is directly observable, and the usual properties of channels such as the transmission rate, capacity, ..., are always defined and studied under this assumption. In this case, according to (2.8) and (2.9), the risk under the transmission of the source μ through the channel ν by means of $\tilde{\varphi}$ and $\tilde{\psi}$ may be expressed as

(2.10)
$$r_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}, w) = \int_{\mathfrak{C}} \int_{\mathfrak{B}} \int_{\mathfrak{Z}} w(\mathfrak{c}, \tilde{\psi}(\mathfrak{b}, \mathfrak{z}))\, d\xi(\mathfrak{z})\, d\gamma(\mathfrak{b} \mid \mathfrak{c})\, d\mu(\mathfrak{c}),$$

where $\gamma(. \mid .)$ is defined by (2.9). If we denote by $\mathscr{D}$ a class of decoders and by $\mathscr{E}$ a class of encoders, we may define for every encoder $\tilde{\varphi}$ the non-negative number

(2.11)
$$r_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w) = \inf_{\mathscr{D}} r_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}, w)$$

which may be referred to as a minimum risk under the transmission of the source μ through the channel ν by means of the random encoding procedure $\tilde{\varphi}$. If we put

(2.12)
$$r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w) = \inf_{\mathscr{E}} r_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w)$$

then the number $r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ may be regarded as a minimum risk with respect to the classes $\mathscr{E}$ and $\mathscr{D}$.

3. FORMULATION OF THE PROBLEM

In this section we shall study some questions connected with the transmission of sources through channels under the assumption that the space of output signals is not directly observable by the receiver. Specifying the conditions of the observation we shall see that the problem which will rise up may be regarded as a synchronization problem.

Let $\pi(i)$, $i \in I$, be a sequence of non-negative numbers for which

$$\sum_{i \in I} \pi(i) = 1 . \tag{3.1}$$

Let us assume that in case when an output signal $\mathfrak{b} \in \mathfrak{B}$ is realized, the signal observed by the receiver is $T^i \mathfrak{b} \in \mathfrak{B}$, where $i \in I$ is a random variable given by the probability distribution π on I.

The intuitive motivation of this assumption may be the following one: Let us consider the source μ, the channel ν, and the encoder $\tilde{\varphi}$. Let, for every $\mathfrak{c} \in \mathfrak{C}$, the i-th coordinate of the signal $\tilde{\varphi}(\mathfrak{c},.)$ be sent at the time i measured by the time scale of the sender. Let us assume that, for every such coordinate, the time necessary for the transmission through the channel is zero and let $b^i \in B$ coincides at the output of the channel with the input coordinate $(\tilde{\varphi}(\mathfrak{c},.))_i$ for every $i \in I$. Hence, the output signal $\mathfrak{b} \in \mathfrak{B}$ which coincides with the input signal $\varphi(\mathfrak{c},.)$ is such that $(\mathfrak{b})_i = b^i$ for every $i \in I$. Let the time scale of the receiver be shifted by the value $j \in I$ with respect to the time scale of the sender, i.e. let the i-th time of the sender be $(j + i)$-th time of the receiver, for every $i \in I$. In this case the signal observed by the receiver is not $\mathfrak{b}$, but $\mathfrak{b}' = T^j \mathfrak{b}$.

It is easy to see that if $\mathfrak{c} \in \mathfrak{C}$ is a realized message and i is a realized value of the random parameter i, then the conditional probability measure on the sample space $(\mathfrak{B}, \mathscr{B})$ provided $\tilde{\varphi}$ is an encoder used for transmission is given by

$$\tilde{\gamma}(E \mid \mathfrak{c}, i) = \int_{\mathfrak{Y}} \gamma(T^{-i}E \mid \tilde{\varphi}(\mathfrak{c}, \mathfrak{y}) \, d\tilde{\eta}(\mathfrak{y}) \quad \text{for every} \quad E \in \mathscr{B}, \mathfrak{c} \in \mathfrak{C}, i \in \mathrm{I} . \tag{3.2}$$

In the sequel we shall study composed decision procedures of the following form: if $\mathfrak{b}$ is a realized signal at the output, if $\tilde{\psi}$ is a decoder, and it $\varrho(\mathfrak{b}) \in I$ is a decision concerning parameter i, then the received message $\mathfrak{c}'$ will be defined as

$$\mathfrak{c}' = \tilde{\psi}_\varrho(\mathfrak{b},.) = \tilde{\psi}(T^{-\varrho(\mathfrak{b})}\mathfrak{b},.) . \tag{3.3}$$

We shall assume that in this case the loss corresponding to the message $\mathfrak{c}$ sent over the channel is given by

$$w(\mathfrak{c}, T^{s(i-\varrho(\mathfrak{b}))} \tilde{\psi}_\varrho(\mathfrak{b},.)) , \tag{3.4}$$

where w is the weight function discussed in the preceding section and where $s(i)$, $i \in I$, is a sequence of integers with $s(0) = 0$ (an assumption of consistency with the preceding section).

If $\mathscr{E}$ or $\mathscr{D}$ is a class of admissible encoders or decoders respectively and if the conditions of observation are as above, then the number $r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ cannot serve as

a measure of the risk under the transmission of the source μ through the channel ν
and instead of it the number $\tilde{r}_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ must be used,

$$(3.5) \qquad \tilde{r}_{\mu\nu} = \inf_{\varrho \in \mathscr{R}} \inf_{\mathscr{E}, \mathscr{D}} \tilde{r}_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}_\varrho, w)$$

where $\mathscr{R}$ denotes the class of all $\mathscr{B}$-measurable functions $\varrho : \mathfrak{B} \to I$ and where $\tilde{r}_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}_\varrho, w)$ is the average risk corresponding to the encoder $\tilde{\varphi}$, decoder $\tilde{\psi}$, and decision function ϱ, i.e.

$$(3.6) \quad \tilde{r}_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}_\varrho, w) = \sum_{i \in I} \pi(i) \int_{\mathfrak{C}} \int_{\mathfrak{B}} \int_{\mathfrak{Z}} w(\mathfrak{c}, T^{s(i-\varrho(\mathfrak{b}))} \tilde{\psi}_\varrho(\mathfrak{b}, \mathfrak{z}))\, d\tilde{\xi}(\mathfrak{z})\, d\tilde{\gamma}(\mathfrak{b} \mid \mathfrak{c}, i)\, d\mu(\mathfrak{c})$$

(cf. (3.3)).

The problem of a relation between $\tilde{r}_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ and $r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ can be refered to as a *synchronization problem of information theory*. If we assume that the class $\mathscr{D}$ contains for every $\tilde{\psi} \in \mathscr{D}$ also decoders of the form $T^{s(i-\varrho(\mathfrak{b}))} \tilde{\psi}_\varrho(\mathfrak{b}, .)$ then it is easy to see that $\tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w) \geqq r_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w)$ for every encoder $\tilde{\varphi}$, where $\tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w)$ is defined similarly to $r_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w)$, and consequently $\tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w) \geqq r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w)$ for every $\tilde{\varphi} \in \mathscr{E}$. The problem arises, for which μ, ν, $\mathscr{E}$, and $\mathscr{D}$ there is, for every $\lambda > 0$, an encoder $\tilde{\varphi} \in \mathscr{E}$ and a class $\mathscr{D}_{\tilde{\varphi}}$ of decoders of the form (3.3) such that

$$(3.7) \qquad \tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}_{\tilde{\varphi}}, w) \leqq r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w) + \lambda\,.$$

Let us consider, for example, the following trivial case: Let for every $\tilde{\varphi} \in \mathscr{E}$ there is a decision function ϱ that is $\mathscr{B}$-measurable transformation of the set $\mathfrak{B}$ into I such that, in case $i \in I$ is a is a realized value of the random parameter, $\varrho(\mathfrak{b}) = i$ for any observed signal $\mathfrak{b}$. If we define $\mathscr{D}_{\tilde{\varphi}}$ for any class $\mathscr{D}$ by

$$\mathscr{D}_{\tilde{\varphi}} = \{\tilde{\psi}_\varrho : \tilde{\psi} \in \mathscr{D}\}\,,$$

then it is easy to see that

$$\tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}_{\tilde{\varphi}}, w) = r_{\mu\nu}(\tilde{\varphi}, D, w)$$

and consequently that (3.7) is satisfied for every λ and $\tilde{\varphi} \in \mathscr{E}$.

The remainder of this section will be devoted to a general definition of synchronizing encoder. Let us assume that the weight function w is bounded from above, i.e. let

$$(3.8) \qquad w(\mathfrak{c}, \mathfrak{c}') \leqq w_0 \quad \text{for every} \quad \mathfrak{c}, \mathfrak{c}' \in \mathfrak{C}\,,$$

and let ϱ be any $\mathscr{B}$-measurable decision function: $\mathfrak{B} \to I$. In order to evaluate, for given $\tilde{\varphi} \in \mathscr{E}$, $\tilde{\psi} \in \mathscr{D}$, and w, the risk $\tilde{r}_{\mu\nu}(\varphi, \tilde{\psi}_\varrho, w)$, we shall define a set $\{\tau_i\}$, $i \in I$, of signed measures on $\mathscr{B}$ by

$$(3.9) \quad \tau_i(E) = \int_{\mathfrak{C}} \int_E \int_{\mathfrak{Z}} [w(\mathfrak{c}, T^{s(i)} \tilde{\psi}(T^i \mathfrak{b}, \mathfrak{z})) - w(\mathfrak{c}, \tilde{\psi}(\mathfrak{b}, \mathfrak{z}))]\, d\tilde{\xi}(\mathfrak{z})\, d\gamma(\mathfrak{b} \mid \mathfrak{c})\, d\mu(\mathfrak{c})$$

for every $E \in \mathscr{B}$, $i \in I$. Let us define, for every $i, i' \in I$, the function $\tilde{w}(i, i' \mid \tilde{\varphi}, \tilde{\psi}, w) = 0$ or w_0 according as $\tau_{i-i'}(E)$ is or is not non-positive on $\mathscr{B}$. Next we shall prove that the following inequality holds:

$$\tilde{r}_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}_\varrho, w) \leqq r_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}, w) + r_{\mu\nu}(\varrho, \pi, \tilde{w}) , \tag{3.10}$$

where

$$r_{\mu\nu}(\varrho, \pi, \tilde{w}) = \sum_{i \in I} \pi(i) \int_{\mathfrak{B}} \tilde{w}(i, \varrho(\mathfrak{b}) \mid \tilde{\varphi}, \tilde{\psi}, w) \, \mathrm{d}\gamma T^{-i}(\mathfrak{b}) . \tag{3.11}$$

(In other words, $r_{\mu\nu}(\varrho, \pi, \tilde{w})$ is the average risk corresponding to the decision function ϱ, parameter space (I, π), decision space I, weight function $\tilde{w}$, and conditional probability measures γT^i on $(\mathfrak{B}, \mathscr{B})$ defined for every $i \in I$ by means of

$$\gamma(E) = \int_{\mathfrak{C}} \gamma(E \mid \mathfrak{c}) \, \mathrm{d}\mu(\mathfrak{c}), \; E \in \mathscr{B} , \tag{3.12}$$

where $\gamma(. \mid .)$ is defined by (2.9).)

Proof: Let ϱ be a decision function definedas above and let

$$\mathfrak{B} = \bigcup_{i' \in I} E_{i'}$$

$$E_{i'} = \{\mathfrak{b} : \mathfrak{b} \in \mathfrak{B}, \varrho(\mathfrak{b}) = i'\} .$$

It follows from the definition of $\tilde{w}$ that

$$\sum_{i' \in I} \tau_{i-i'}(E_{i'}) \leqq \int_{\mathfrak{B}} \tilde{w}(i, \varrho(\mathfrak{b}) \mid .) \, \mathrm{d}\gamma T^{-i}(\mathfrak{b}) \quad \text{for every} \quad i \in I ,$$

and the desired inequality (3.10) follows from the following relations:

$$\tilde{r}_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}_\varrho, w) - r_{\mu\nu}(\tilde{\varphi}, \tilde{\psi}, w) = \sum_{i \in I} \pi(i) \int_{\mathfrak{C}} \int_{\mathfrak{B}} \int_{\mathfrak{Z}} [w(\mathfrak{c}, T^{s(i-\varrho(\mathfrak{b}))} \tilde{\psi}(T^{i-\varrho(\mathfrak{b})}\mathfrak{b}, \mathfrak{z})) - w(\mathfrak{c}, \tilde{\psi}(\mathfrak{b}, \mathfrak{z}))] \, \mathrm{d}\xi(\mathfrak{z}) \, \mathrm{d}\gamma(\mathfrak{b} \mid \mathfrak{c}) \, \mathrm{d}\mu(\mathfrak{c}) = \sum_{i \in I} \pi(i) \sum_{i' \in I} \tau_{i-i'}(E_{i'}) .$$

Denoting by $\mathscr{P}$ the set of all distributions π on I, it seems to be suitable to define a synchronizing encoder in the following way:

An encoder $\tilde{\varphi}$ is said to be synchronizing with respect to μ, ν, w, and $\mathscr{D}$ if, for every $\tilde{\psi} \in \mathscr{D}$,

$$\inf_{\varrho \in \mathscr{R}} \sup_{\pi \in \mathscr{P}} r_{\mu\nu}(\varrho, \pi, \tilde{w}) = 0 . \tag{3.13}$$

(As, in accordance with our model, the distribution π cannot be assumed to be known, the minimax condition is used here.)

But, on the other hand, we must respect also the viewpoint of real communication systems. As the memory of real receivers is always finite, the requirement of the $\mathscr{B}$-measurability of decision functions $\varrho \in \mathscr{R}$ in the latter definition is not sufficient. Namely, in case the capacity of the memory is "$h \in I$ concesutive letters of the alphabet B", it is to see that the receiver cannot distinguish between output signals $\mathfrak{b}, \mathfrak{b}'$ in case $(\mathfrak{b})_1^h = (\mathfrak{b}')_1^h$. That is why we shall define a synchronizing encoder as follows:

Definition 1. Let $\mathscr{R}_h \subset \mathscr{R}$, $h \in I^+$, denotes the set of all $\mathscr{B}_1^h$-measurable decision functions ϱ and let μ, ν, w, $\mathscr{E}$, and $\mathscr{D}$ be fixed. We shall say that $\tilde{\varphi} \in \mathscr{E}$ is a synchronizing encoder with respect to μ, ν, and $\mathscr{D}$ if, for every $\tilde{\psi} \in \mathscr{D}$,

$$\lim_{h \to \infty} r_{\mu\nu}^h(\tilde{w}) = 0, \tag{3.14}$$

where

$$r_{\mu\nu}^h(\tilde{w}) = \inf_{\mathscr{R}_h} \sup_{\mathscr{P}} r_{\mu\nu}(\varrho, \pi, \tilde{w}(.,. \mid \tilde{\varphi}, \tilde{\psi}, w)), \quad h \in I^+. \tag{3.15}$$

It is clear that the condition (3.14) is stronger than (3.13).

Now the synchronization problem may be formulated as follows (cf. (3.7) and (3.10)): For which μ, ν, $\mathscr{E}$, and $\mathscr{D}$ there is, for every $\lambda > 0$, a synchronizing encoder $\tilde{\varphi} \in \mathscr{E}$ with respect to μ, ν, and $\mathscr{D}$ such that

$$\tilde{r}_{\mu\nu}(\tilde{\varphi}, \mathscr{D}, w) \leqq r_{\mu\nu}(\mathscr{E}, \mathscr{D}, w) + \lambda. \tag{3.16}$$

4. CONCEPT OF SYNCHRONIZING (n, p)-ENCODER

In what follows we shall rastrict ourselves to the class $\mathscr{E}_{np}$, $n, p \in I^+$, of encoders $\tilde{\varphi}$ which may be obtained from an (n, p)-encoder φ by (2.4) and (2.6), and to the class $\mathscr{D}_{np}$ of decoders $\tilde{\psi}$ which may be obtained from a (p, n)-decoder ψ by (2.5) and (2.7) where, in order to keep a logical meaning of this statements, the sets Y and Z are assumed to be for instance sets of real numbers. Moreover, since there is one-to-one correspondence between an (n, p)-encoder φ and $\tilde{\varphi} \in \mathscr{E}_{np}$, we shall restrict ourselves further on to the study of (n, p)-encoders only. Throughout this section we shall assume that if $i = kp$ for some $k \in I$, then $s(i) = -kn$.

In the sequel the following two cases will be discussed separately:

(ST) Both the source μ and the channel ν are stationary, i.e.

$$\mu(TE) = \mu(E) \quad \text{for every} \quad E \in \mathscr{C}, \tag{4.1}$$

$$\nu(TE \mid T\mathfrak{a}) = \nu(E \mid \mathfrak{a}) \quad \text{for every} \quad E \in \mathscr{B},\ \mathfrak{a} \in \mathscr{A}. \tag{4.2}$$

(NON-ST) The source μ or the channel ν is non-stationary.

Following Khinchin [1] we shall use in case (ST) the weight function $w = w_n$, $n \in I^+$, defined as follows:

$$\begin{aligned} w_n(\mathfrak{c}, \mathfrak{c}') &= 0 \quad \text{if} \quad (\mathfrak{c})_1^n = (\mathfrak{c}')_1^n, \\ w_n(\mathfrak{c}, \mathfrak{c}') &= 1 \quad \text{if} \quad (\mathfrak{c})_1^n \neq (\mathfrak{c}')_1^n. \end{aligned} \tag{4.3}$$

 It is clear that w_n is $\mathscr{C} \otimes \mathscr{C}$-measurable and

(4.4) $$w_n \leqq w_0 = 1 \quad \text{for every} \quad n \in I^+ .$$

In case (NON-ST) we shall use the weight function $w = \bar{w}_n$, $n \in I^+$, where

(4.5) $$\bar{w}_n(\mathfrak{c}, \mathfrak{c}') = \limsup_{k \to \infty} \frac{1}{2k+1} \sum_{i=-k}^{k} w_n(T^{in}\mathfrak{c}, T^{in}\mathfrak{c}')$$

for w_n defined above. It is evident that $\bar{w}_n$ is $\mathscr{C} \otimes \mathscr{C}$-measurable and

(4.6) $$\bar{w}_n \leqq w_0 = 1 \quad \text{for every} \quad n \in I^+ .$$

We shall put for every (n, p)-encoder φ

(4.7) $$e(\varphi, \mu, \nu) = r_{\mu\nu}(\tilde{\varphi}, \mathscr{D}_{np}, w) ,$$

where $w = w_n$ or $\bar{w}_n$ according as the condition (ST) is or is not satisfied. The number $\epsilon(\varphi, \mu, \nu)$ is refered to as the average probability of the incorrect transmission of an n-sequence $(\mathfrak{c})_{ni+1}^{n(i+1)}$, $i \in I$, associated with the (n, p)-encoder φ.

Defining on $I \otimes I$ the weight function

(4.8) $$\tilde{w}_p(i, i') = \left\langle \begin{matrix} 0 \\ 1 \end{matrix} \right. \quad \text{for} \quad p \in I^+$$

depending on whether $i \equiv i' \pmod p$ or $i \not\equiv i' \pmod p$ respectively, we shall prove the following result:

Lemma 1. *For every $\tilde{\varphi} \in \mathscr{E}_{np}$, $\tilde{\psi} \in \mathscr{D}_{np}$, $\varrho \in \mathscr{R}$, $\pi \in \mathscr{P}$, for every channel ν and source μ, and for every sequence $s(i)$, $i \in I$, under consideration the following inequality holds*

(4.9) $$r_{\mu\nu}(\varrho, \pi, \tilde{w}(.\,, .\mid \tilde{\varphi}, \tilde{\psi}, w)) \leqq r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) ,$$

where $w = w_n$ or $\bar{w}_n$ according as the condition (ST) is or is not satisfied.

Proof. (I) let the condition (ST) be satisfied. To prove (4.9) it suffices to show that, in case $i = i' + kp$, $k \in I$, for every $\tilde{\varphi} \in \mathscr{E}_{np}$ and $\tilde{\psi} \in \mathscr{D}_{np}$, the equality $\tilde{w}(i, i' \mid \tilde{\varphi}, \tilde{\psi}, w_n) = 0$ holds or, in view of the definition of $\tilde{w}$, that $\tau_{kp}(E) \leqq 0$, $E \in \mathscr{B}$. To prove the latter inequality it sufficies to prove that

(4.10) $$\int_{\mathfrak{C}} \int_E \int_{\mathfrak{Z}} w_n(\mathfrak{c}, T^{-kn}\tilde{\psi}(T^{kp}\mathfrak{b}, \mathfrak{z}))\, d\xi(\mathfrak{z})\, d\gamma(\mathfrak{b} \mid \mathfrak{c})\, d\mu(\mathfrak{c}) \leqq$$
$$\leqq \int_{\mathfrak{C}} \int_E \int_{\mathfrak{Z}} w_n(\mathfrak{c}, \tilde{\psi}(\mathfrak{b}, \mathfrak{z}))\, d\xi(\mathfrak{z})\, d\gamma(\mathfrak{b} \mid \mathfrak{c})\, d\mu(\mathfrak{c}) , \quad E \in \mathscr{B} ,$$

(cf. the equality $s(kp) = -kn$). The latter inequality (with the sign of equality) immediately follows from the following property of (n, p)-encoders: $\tilde{\psi}(T^{kp}\mathfrak{b}, \mathfrak{z}) = T^{kn}\tilde{\psi}(\mathfrak{b}, T^{-k}\mathfrak{z})$ and from the clear fact that $\xi T^k = \xi$ (cf. (2.5)).

(II) The proof of (4.9) in case (NON-ST) can be given similarly.

According to Lemma 1 and Definition 1, we can give the following definition:

Definition 2. Let μ be a source and ν a channel. We shall say that an (n, p)-encoder φ is synchronizing with respect to μ and ν if

$$\lim_{h \to \infty} r^h_{\mu\nu} = 0 , \tag{4.11}$$

where

$$r^h_{\mu\nu} = \inf_{\varrho \in \mathcal{R}_h} \sup_{\pi \in \mathcal{P}} r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) . \tag{4.12}$$

Formulation of the synchronisation problem: For which source μ and channel ν there is, for every $\lambda > 0$, a positive integer n and a synchronizing (n, n)-encoder φ with respect to μ and ν such that $e(\varphi, \mu, \nu) < \lambda$.

Remark. It was shown in [2] and [7] that if μ is an ergodic source with finite alphabet C and ν is a totally ergodic channel with finite alphabets A, B, and with finite past history, then in case the entropy rate of μ is less than the ergodic capacity of ν there is, for every $\lambda > 0$, $n_0 \in I^+$ such that for every $n > n_0$ there exists a synchronizing (n, n)-encoder φ (with respect to μ and ν) with $e(\varphi, \mu, \nu) < \lambda$. The author has obtained a similar result in case that μ is an arbitrary source and ν is a memoryless channel, including an evaluation of the rate of convergence of $r^h_{\mu\nu}$ in (4.13) to zero (for the obtained synchronizing encoders). A similar result was obtained also in case ν is a non-ergodic channel of a special type.

5. PROPERTIES OF SYNCHRONIZING (n, p)-ENCODERS

Theorem 1. *If μ is a stationary source, ν a stationary channel, and φ an (n,p)-encoder, then the following three conditions are equivalent:*

(I) *φ is a synchronizing encoder with respect to μ and ν.*

(II) *For every $\lambda > 0$ there is $h \in I^+$ and a mapping $\varrho \in \mathcal{R}_h$ of the set $\mathfrak{B}$ into the set of integers $\{0, 1, \ldots, p - 1\}$ such that*

$$\int_{\mathfrak{C}} \int_{\mathfrak{Y}} \nu(\{\mathfrak{b} : \varrho(T^i\mathfrak{b}) \neq i\} \mid \tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y})\, d\mu(\mathfrak{c}) < \lambda \quad \text{for} \quad i = 0, 1, \ldots, p - 1 . \tag{5.1}$$

(III) *$\gamma \perp \gamma T^i$ for $i = 0, 1, \ldots, p - 1$, where γ is defined, for the given φ, by (2.6) and (3.12).*

Proof. The proof will be based on the following two lemmas:

Lemma 2. *Let $(\mathfrak{X}, \mathcal{X})$ be an arbitrary measurable space and γ, $\bar{\gamma}$ probability measures on $\mathcal{X}$. If, for every $\lambda > 0$, there is $E \in \mathcal{X}$ such that $\gamma(E) > 1 - \lambda$, $\bar{\gamma}(E) < \lambda$, then γ, $\bar{\gamma}$ are mutually singular, i.e. $\gamma \perp \bar{\gamma}$.*

Proof. For every $\lambda(n) = (1/2)^n$, $n \in I^+$, there is $E(n) \in \mathscr{X}$ such that

$$\begin{aligned}(5.2) \qquad & \gamma(E(n)) > 1 - (1/2)^n, \\ & \bar{\gamma}(E(n)) < (1/2)^n .\end{aligned}$$

Put

$$F(k) = \bigcup_{n=k}^{\infty} E(n),$$

$$F = \bigcap_{k=1}^{\infty} F(k).$$

It follows from (5.2) that $\gamma(F(k)) = 1$ for all $k \in I^+$ and, therefore, $\gamma(F) = 1$. To prove $\bar{\gamma}(F) = 0$ one uses the relation $F \subset F(k)$ which holds for all $k \in I^+$ and, consequently, which implies the inequality

$$\bar{\gamma}(F) \leqq \sum_{n=k+1}^{\infty} \bar{\gamma}(E(n)) \leqq (1/2)^k \quad \text{for all} \quad k \in I^+ .$$

Lemma 3. *If μ is a stationary source and ν a stationary channel and if φ is an (n, p)-encoder, then, for every $i, j \in I$, $i \equiv j \pmod p$, $\gamma T^i = \gamma T^j$, where the measure γ is defined for the given φ by (2.6) and (3.12).*

Proof. If $E \in \mathscr{B}$, then

$$\gamma T^i(E) = \int_{\mathfrak{C}} \int_{\mathfrak{Y}} \nu(T^i E \mid \tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y})\, d\mu(\mathfrak{c})$$

(cf. (2.4)) and, in accordance with (4.2),

$$\gamma T^i(E) = \int_{\mathfrak{C}} \int_{\mathfrak{Y}} \nu(E \mid T^{-i}\tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y})\, d\mu(\mathfrak{c}) \quad \text{for every} \quad i \in I .$$

If $i \equiv j \pmod p$, then there are $k_1, k_2 \in I$ and $1 \leqq r \leqq p$ such that $i = k_1 p + r$, $j = k_2 p + r$. In view of

$$T^{kp}\tilde{\varphi}(\mathfrak{c}, \mathfrak{y}) = \tilde{\varphi}(T^{kn}\mathfrak{c}, T^k\mathfrak{y})$$

and (2.4), it follows that

$$\gamma T^i(E) = \int_{\mathfrak{C}} \int_{\mathfrak{Y}} \nu(E \mid T^{-r}\tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y})\, d\mu T^{k_1 p}(\mathfrak{c}),$$

$$\gamma T^j(E) = \int_{\mathfrak{C}} \int_{\mathfrak{Y}} \nu(E \mid T^{-r}\tilde{\varphi}(\mathfrak{c}, \mathfrak{y}))\, d\tilde{\eta}(\mathfrak{y})\, d\mu T^{k_2 p}(\mathfrak{c}),$$

and it remains to use (4.1).

Next we prove that (I) $\Rightarrow$ (II). By (3.11) and (4.8) we can write, for every $\varrho \in \mathscr{R}$,

$$r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) = \sum_{i \in I} \pi(i)\, \gamma T^{-i}(\{\mathfrak{b} : \varrho(\mathfrak{b}) \not\equiv i \pmod p\})\,.$$

There is a disjoint decomposition

$$I = \bigcup_{i=0}^{p-1} I_i$$

defined by

$$j \in I_i \Leftrightarrow j \equiv i \pmod p \quad \text{for} \quad i = 0, 1, \ldots, p-1\,.$$

Hence, by Lemma 3,

$$r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) = \sum_{i=0}^{p-1} \pi(I_i)\, \gamma T^{-i}(\{\mathfrak{b} : \varrho(\mathfrak{b}) \neq i\})\,. \tag{5.3}$$

It follows immediately from (5.3), (4.13), and (4.14) that if (I) holds, then there is, for every $\lambda > 0$, $h \in I^+$ and a decision function $\varrho \in \mathscr{R}_h$ such that

$$\gamma T^{-i}(\{\mathfrak{b} : \varrho(\mathfrak{b}) \neq i\}) < \lambda \quad \text{for} \quad i = 0, 1, \ldots, p-1 \tag{5.4}$$

or, equivalently,

$$\gamma(\{\mathfrak{b} : \varrho(T^i\mathfrak{b}) \neq i\}) < \lambda \quad \text{for} \quad i = 0, 1, \ldots, p-1\,. \tag{5.5}$$

The desired validity of (5.1) follows from (3.12).

(II) $\Rightarrow$ (III). Let $\lambda > 0$ be arbitrary. If ϱ satisfies the inequality (5.1), then defining a measurable set E by $E = \{\mathfrak{b} : \varrho(\mathfrak{b}) = 0\}$ one may show that $\gamma(E) > 1 - \lambda$, $\bar{\gamma}(E) < \lambda$ for

$$\bar{\gamma} = \frac{1}{p-1} \sum_{i=1}^{p-1} \gamma T^i$$

(cf. the equality

$$\sum_{i=1}^{p-1} \gamma T^i = \sum_{i=1}^{p-1} \gamma T^{-i}$$

implied by Lemma 3). To prove that (III) holds it remains to apply Lemma 2 and use the fact that γT^i, $i = 1, 2, \ldots, p-1$, is absolutely continuous with respect to $\bar{\gamma}$.

(III) $\Rightarrow$ (I). In case that (III) holds it is easily verified that there is a set $E \in \mathscr{B}$ such that $\gamma(E) = 1$ and that $T^iE \cup T^jE = \emptyset$ for $i, j = 0, 1, \ldots, p-1$; $i \neq j$. Hence, denoting by E_h a cylinder set in $\mathscr{B}$, determined in coordinates $1, 2, \ldots, h$ i.e.

$$E_h = \bigcup_{\boldsymbol{b} \in F} \{\mathfrak{b} : (\mathfrak{b})_1^h = \boldsymbol{b}\}$$

for at least one $F \subset B^h$, we obtain the following result: There exist, for $i = 0, 1, \ldots$

 ..., $p - 1$, sequences $\{E_h^i\}_{h=1}^{\infty}$ of cylinder sets in $\mathscr{B}$ such that

(5.6) $$\lim_{h\to\infty} E_h^i = T^i E ,$$

$$\lim_{h\to\infty} \gamma(T^{-j}E_h^i) = 0 \quad \text{for all} \quad j = 0, 1, \ldots, p - 1 , \quad j \neq i ,$$

and, consequently, for every $\lambda > 0$ there exists $h \in I^+$ such that

(5.7) $$\gamma T^{-i}(E_h^i) > \lambda/p$$

$$\gamma T^{-j}(E_h^i) < \lambda/p \quad \text{for all} \quad i, j = 0, 1, \ldots, p - 1 ; \quad i > j .$$

If we define a decision function ϱ to be equal 0 for $\mathfrak{b} \in E_h^0$, to be equal $j = 1, 2, \ldots$ $\ldots, p - 2$ for

$$\mathfrak{b} \in E_h^j - \bigcup_{i=0}^{j-1} E_h^i ,$$

and to be equal $p - 1$ for

$$\mathfrak{b} \in \mathfrak{B} - \bigcup_{i=0}^{p-2} E_h^i \subset E_h^{p-1} ,$$

then $\varrho \in \mathscr{R}_h$. Moreover, in view of (5.3) and Lemma 3, it follows that $r_{\mu\nu}^h < \lambda$ (cf. (4.12)). Since λ may be taken arbitrarily small, the desired statement holds.

Theorem 2. *If μ is a stationary source and ν a stationary channel and if φ is a synchronizing (n, p)-encoder with respect to μ and ν, then there is a decision function $\varrho \in \mathscr{R}$ such that $r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) = 0$ for all $\pi \in \mathscr{P}$.*

Proof. Theorem 2 is an immediate consequence of Theorem 1 (cf. property (III) of synchronizing encoders).

6. APPLICATION OF SOME DECISION THEORY RESULTS

It is easy to see that the rate of convergence of $r_{\mu\gamma}^h$ in (3.14) to zero is an important economical characteristics of every synchronizing (n, p)-encoder. That is why we shall give some results concerning it. Our results are based on the ideas and results concerning data reduction problems of statistical decision theory, latterly developed by A. Perez [3], [5].

Let φ be an (n, p)-encoder, μ a source, ν a channel, and π a distribution on I. Denote by $\mathscr{I}$ the σ-algebra of all subsets of the set I and by P_π the probability measure defined on $\mathscr{I} \otimes \mathscr{B}$ by

(6.1) $$P_\pi(E) = \sum_{i\in I} \pi(i)\gamma T^{i}(\{\mathfrak{b} : (i, \mathfrak{b}) \in E\}) \quad \text{for every} \quad E \in \mathscr{I} \otimes \mathscr{B} ,$$

where γ is the probability measure on $\mathscr{B}$, defined for the given φ by (2.6) and (3.12). Denote by $\tilde{P}_\pi$ the probability measure defined on $\mathscr{I} \otimes \mathscr{B}$ by

(6.2) $$\tilde{P}_\pi = \pi \otimes P'_\pi ,$$

where P'_π is the marginal probability measure of the measure P_π, defined on $\mathscr{B}$, i.e.

(6.3) $$P'_\pi(E) = \sum_{i \in I} \pi(i) \gamma T^i(E) \quad \text{for every} \quad E \in \mathscr{B} ,$$

(6.4) $$\tilde{P}_\pi(E) = \sum_{i,j \in I} \pi(i)\, \pi(j)\, \gamma T^i(\{\mathfrak{b} : (j, \mathfrak{b}) \in E\}) \quad \text{for every} \quad E \in \mathscr{I} \otimes \mathscr{B} .$$

It follows from (6.1) and (6.4) that $P_\pi \ll \tilde{P}_\pi$ and, consequently, the Radon - Nikodym density function

$$f_\pi = \frac{\mathrm{d}P_\pi}{\mathrm{d}\tilde{P}_\pi}$$

exists and is $\mathscr{I} \otimes \mathscr{B}$-measurable. Denoting by P^h_π, $\tilde{P}^h_\pi$ the restrictions of the measures P_π and $\tilde{P}_\pi$ on the σ-algebra $\mathscr{I} = \mathscr{B}^h_1$, the Radon - Nikodym density function

$$f^h_\pi = \frac{\mathrm{d}P^h_\pi}{\mathrm{d}\tilde{P}^h_\pi}$$

also exists and is $\mathscr{I} \otimes \mathscr{B}^h_1$-measurable. Hence we may define informations (cf. [4]):

$$J(P_\pi) = \int_{I \otimes \mathfrak{B}} \log f_\pi(i, \mathfrak{b})\, \mathrm{d}P_\pi(i, \mathfrak{b}) ,$$

$$J(P^h_\pi) = \int_{I \otimes \mathfrak{B}} \log f^h_\pi(i, \mathfrak{b})\, \mathrm{d}P^h_\pi(i, \mathfrak{b}) ,$$

where, according to Theorem 6 of [4], $J(P_\pi) \geqq J(P^h_\pi)$ for every $h \in I^+$.

In case that φ is a synchronizing (n, p)-encoder with respect to μ and ν, the Bayes risk of the decision problem with parameter measure space $(I, \mathscr{I}, \pi)$, sample space $(B, \mathscr{B})$, decision space I, probability measure P_π on $\mathscr{I} \otimes \mathscr{B}$, and weight function $\tilde{w}_p$ is zero (cf. (3.13), (4.11)). Hence, according to Theorem 4.3 of [3] Corollary 4.1 of [5].

(6.5) $$\inf_{\varrho \in \mathscr{R}_h} r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) \leqq 2(J(P_\pi) - J(P^h_\pi)) \quad \text{for every} \quad \pi \in \mathscr{P} ,$$

where the number on the left side is the Bayes risk of the decision problem differing from the above decision problem only in what concerns the sample space: in the place of $(\mathfrak{B}, \mathscr{B})$ we have now $(\mathfrak{B}, \mathscr{B}^h_1)$.

Defining

(6.6) $$\tilde{\pi}(i) = 1/p \quad \text{for every} \quad i = 0, 1, \ldots, p - 1 ,$$
$$\tilde{\pi}(i) = 0 \quad \text{for} \quad i \notin \{0, 1, \ldots, p - 1\}$$

we obtain:

Theorem 3. *For every stationary source μ, stationary channel ν, and for every* *synchronizing (n, p)-encoder φ with respect to μ and ν,*

$$r^h_{\mu\nu} \leqq 2p(J(P_{\tilde{\pi}}) - J(P^h_{\tilde{\pi}})) \tag{6.7}$$

(*cf.* (4.12)).

Proof. Using (5.3) we obtain the inequality

$$r_{\mu\nu}(\varrho, \pi, \tilde{w}_p) \leqq p r_{\mu\nu}(\varrho, \tilde{\pi}, \tilde{w}_p)$$

for $\tilde{\pi}$ given by (6.6). In view of (6.5), it follows that

$$r^h_{\mu\nu} \leqq p \inf_{\varrho \in \mathscr{R}_h} r_{\mu\nu}(\varrho, \tilde{\pi}, \tilde{w}_p)$$

and hence, to obtain (6.7) it suffice to use (6.5).

The source μ is said to be independent if

$$\mu = \bigotimes_{i=-\infty}^{+\infty} \mu_i ,$$

where μ_i is a probability measure defined on $\mathscr{C}^1$ by

$$\mu_i(E) = \mu(\{\mathfrak{c} : (\mathfrak{c})_1 \in E\}) \quad \text{for every} \quad E \in \mathscr{C}^1 , \quad i \in I .$$

Remark. It is easily verified that

$$\mathscr{C} = \bigotimes_{i=-\infty}^{+\infty} \mathscr{C}_i ,$$

where $\mathscr{C}_i = \mathscr{C}^1$ for all $i \in I$ (cf. Sec 2).

The channel ν is said to be memoryless if the zero-past-history condition

$$\nu(\{\mathfrak{b} : (\mathfrak{b})_1 \in E\} \mid \mathfrak{a}) = \nu(\{\mathfrak{b} : (\mathfrak{b})_1 \in E \mid \mathfrak{a}') \tag{6.8}$$

is satisfied for every $E \in \mathscr{B}^1$, $\mathfrak{a}, \mathfrak{a}' \in \mathfrak{A}$, $(\mathfrak{a})_1 = (\mathfrak{a}')_1$ and if, for every $\mathfrak{a} \in \mathfrak{A}$,

$$\nu(. \mid \mathfrak{a}) = \bigotimes_{i=-\infty}^{+\infty} \nu_1(. \mid (\mathfrak{a})_i) ,$$

where $\nu_1(. \mid a)$, defined by

$$\nu_1(E \mid a) = \nu(\{\mathfrak{b} : (\mathfrak{b})_1 \in E\} \mid \mathfrak{a}) , \quad \mathfrak{a} \in \mathfrak{A} , \quad (\mathfrak{a})_1 = a$$

(cf. (6.8)), for every $E \in \mathscr{B}^1$, is for every $a \in A$ a probability measure on $\mathscr{B}^1$.

Let μ be an independent source, ν a memoryless channel and let φ be an (n, p)-encoder. Let $\tilde{\pi}$ be defined by (6.6) and let $j \in \{0, 1, \ldots, p - 1\}$ be a realized value of the unknown parameter. In this case the measure on the sample space is γT^{-j} and the sequence $\{(\mathfrak{b})^{p(i+1)}_{pi+1}\}$, $i \in I$, is, for every j taken into the consideration, a sequence of independent equally distributed random variables. This implies that in case the alphabet B is finite we can use the result of Rényi [6] (cf. also § 5 of [5]) to obtain the following assertion: $J(P_{\tilde{\pi}}) = \log p$ and there is $\Delta_0 \in I^+$ and $0 < \varepsilon_0 < 1$ such that

$$J(P_{\tilde{\pi}}) - J(P_{\tilde{\pi}}^h) < \Delta_0 \varepsilon_0^{[h/p]} \quad \text{for every} \quad h \in I^+ , \tag{6.9}$$

where $[h/p]$ is non-negative integer defined by the inequality $[h/p] \leqq h/p < [h/p] + 1$. Hence, in view of Theorem 3, we have obtained the following Theorem (cf. also inequality (5.8) of [5]).

Theorem 4. *If μ is an independent source and ν a memoryless channel with finite output alphabet B, then, for every synchronizing (n, p)-encoder φ with respect to μ and ν, there is $\Delta \in I^+$ and $0 < \varepsilon < 1$ such that*

$$r_{\mu\nu}^h < \Delta \varepsilon^h \tag{6.10}$$

for every $h \in I^+$.

(Received January 14th, 1966.)

REFERENCES

[1] A. Я. Хинчин: Об основных теоремах теории информации, Усп. мат. наук *11* (1956), № 1 (67).

[2] J. Nedoma: The Synchronization for Ergodic Channels. In "Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes 1962". Praha 1964.

[3] A. Perez: Information, ε-Sufficiency and Data Reduction Problems. Kybernetika *1* (1965), No. 4.

[4] A. Perez: Notions généralisées d'incertitude, d'entropie et d'information du point de vue de la théorie de martingales. In "Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes 1956", Praha 1957.

[5] A. Perez: Information theory methods in reducing complex decision problems. In "Transaction of the Fourth Prague Conference on Information Theory, Statistical Desicion Functions, Random Processes 1965" (to be published).

[6] A. Rényi: On the amount of information in a sample concerning a parameter. Publications of the Math. Inst. of the Hungarian Acad. Sci. (in print).

[7] I. Vajda: A Synchronization Method for Totally Ergodic Channels. In "Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes 1965" (to be published).

Problém synchronizace v teorii informace

IGOR VAJDA

Abychom při sdělování zpráv sdělovacími kanály při použití blokových kódů mohli použít optimálních dekódovacích procedur je nutné nejdříve rozdělit přijatou zprávu v bloky, které by „časově" odpovídali vyslaným blokům. V případech kdy uvažovaný sdělovací kanál je jediným sdělovacím prostředkem mezi vysílatelem a příjemcem tedy vzniká jakýsi problém synchronizace. Tento problém byl v posledních létech hluboce studován z hlediska algebraické teorie kódů za předpokladu, že uvažovaný kanál je bezšumový. Zdá se, že v případě šumových kanálů je nutné tento problém studovat z hlediska obecné teorie statistického rozhodování o což se pokouší předkládaná práce. Hlavní pozornost je věnována definici synchronizačního kódu. Za jistých předpokladů zdá se rozumným nazvat synchronizačním takový kód, který umožňuje rozdělení libovolné výstupní posloupnosti sestávající z h písmen v časově odpovídající bloky se střední pravděpodobností chyby r^h, kde $\lim r^h = 0$ pro $h \to \infty$. V práci jsou ukázány některé nutné a postačující podmínky proto, aby blokový kód byl synchronizační (Theorem 1) a je nalezen jistý obecný odhad čísla r^h. V posledním paragrafu práce je dokázáno, že v případě kanálu bez paměti a nezávislého zdroje r^h konverguje k nule exponenciálně pro jakýkoliv synchronizační blokový kód.

Ing. Igor Vajda, Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.