

# ANNALES DE L'I. H. P., SECTION B

HILLEL FURSTENBERG

YUVAL PERES

BENJAMIN WEISS

## **Perfect filtering and double disjointness**

*Annales de l'I. H. P., section B*, tome 31, n° 3 (1995), p. 453-465

[http://www.numdam.org/item?id=AIHPB\\_1995\\_\\_31\\_3\\_453\\_0](http://www.numdam.org/item?id=AIHPB_1995__31_3_453_0)

© Gauthier-Villars, 1995, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## Perfect filtering and double disjointness

by

Hillel FURSTENBERG, Yuval PERES (\*) and Benjamin WEISS

Institute of Mathematics, the Hebrew University, Jerusalem.

---

ABSTRACT. – Suppose a stationary process  $\{U_n\}$  is used to select from several stationary processes, *i.e.*, if  $U_n = i$  then we observe  $Y_n$  which is the  $n$ 'th variable in the  $i$ 'th process. when can we recover the selecting sequence  $\{U_n\}$  from the output sequence  $\{Y_n\}$ ?

RÉSUMÉ. – Soit  $\{U_n\}$  un processus stationnaire utilisé pour la sélection de plusieurs processus stationnaires, c'est-à-dire si  $U_n = i$  alors on observe  $Y_n$  qui est le  $n$ -ième variable dans le  $i$ -ième processus. Quand peut-on reconstruire  $\{U_n\}$  à partir de  $\{Y_n\}$  ?

---

### 1. INTRODUCTION

Suppose a discrete-time stationary stochastic signal  $\{U_n\}$ , taking integer values, is transmitted over a noisy channel. If  $U_n = i$  then a random variable  $X_n^{(i)}$  is received at the end of the channel. In this note we give conditions for the “multiplexed” signal  $\{Y_n\} = \{X_n^{(U_n)}\}$  to uniquely determine the original signal  $\{U_n\}$  with probability 1, in a stationary setting. These conditions lead to some interesting questions in Ergodic theory, but leave open the algorithmic problem of explicitly recovering  $\{U_n\}$ .

---

(\*) Current address: Department of Statistics, University of California, Berkeley, CA 94720. Research partially sponsored by the Edmund Landau Center for research in Mathematical Analysis, supported by the Minerva Foundation (Germany).

DEFINITION. – Assume that  $X^{(i)} = \{X_n^{(i)}\}$  is a stationary stochastic process for each integer  $i \geq 1$ . Let  $U = \{U_n\}$  be a stationary process defined on the same probability space, where each variable  $U_n$  takes positive integer values, and the joint distribution of all processes is stationary. We say that the collection  $[U; X^{(1)}, X^{(2)}, X^{(3)}, \dots]$  admits a perfect filter if for every  $m \geq 1$ , the variable  $U_m$  may be expressed as a measurable function, defined almost everywhere, of the sequence  $\{Y_n\} = \{X_n^{(U_n)}\}$  (in other words,  $U_m$  is measurable with respect to the completion of the  $\sigma$ -algebra by the variables  $Y_1, Y_2, Y_3, \dots$  in the underlying probability space).

THEOREM 1. – Assume that for each integer  $i$ , the process  $X^{(i)} = \{X_n^{(i)}\}_{n \geq 1}$  consists of independent identically distributed random variables, where  $X^{(i)}$  and  $X^{(j)}$  have different distributions for  $i \neq j$ . If the integer valued stationary process  $\{U_n\}$  has zero entropy, then the collection  $[U; X^{(1)}, X^{(2)}, \dots]$  admits a perfect filter.

Perhaps the most general stationary filtering problem consists of recovering  $\{U_n\}_{n=1}^\infty$  from the data  $\{F(U_n, X_n)\}_{n=1}^\infty$ , where  $\{X_n\}_1^\infty$  and  $\{U_n\}_1^\infty$  are stationary processes; if each variable  $U_n$  takes only countable many values, this reduces to the filtering problem above. We return to the general case in §5 (cf. Theorem 7).

In  $[F]$  the case in which the binary operation  $F$  is addition was considered, and it was shown that disjointness of two integrable stationary processes  $\{V_n\}$  and  $\{Z_n\}$  is sufficient to determine  $\{V_n\}$  from the sum  $\{V_n + Z_n\}$ .

Recall from  $[F]$  that two stationary processes are disjoint if the only stationary coupling of them is the independent coupling (detailed definitions are given in the next section).

In the context of Theorem 1, disjointness of  $U$  from the  $X^{(i)}$  is not sufficient for perfect filtering; this may be seen by considering the case in which  $U$  is a  $\{1, 2\}$ -valued i.i.d. process and  $X^{(1)}, X^{(2)}$  are distinct zero entropy processes with identical finite state-spaces. The pertinent condition seems to be double disjointness of  $U$  from each  $X^{(i)}$ , which means that any stationary coupling (i.e., joining) of two copies of  $U$  must be disjoint from  $X^{(i)}$  (see §2 for details).

CONVENTION. – For a stationary process  $Z = \{Z_n\}_{n \geq 1}$ , we refer to the distribution of  $Z_1$  as the marginal distribution of the process  $Z$ .

We now state an extension of Theorem 1.

THEOREM 2. – Let  $X^{(1)}, X^{(2)}, X^{(3)}, \dots$  be stationary processes with distinct marginal distributions. If  $U = \{U_n\}$  is an  $\mathbf{N}$ -valued process, doubly

disjoint from each  $X^{(i)}$ , then the collection  $[U; X^{(1)}, X^{(2)}, \dots]$  admits a perfect filter.

The rest of the paper is organized as follows. Some examples and properties of double disjointness are discussed in §2. In particular, a system with positive entropy cannot be doubly disjoint from any nontrivial system. We also explain how Theorem 1, and similar filtering results, follow from Theorem 2. The latter is proved in §3. In §4 the notion of tightness for stationary processes, due to Ornstein and Weiss, is described. It is useful for obtaining versions of the filtering theorem valid for all generic sequences. The final section, §5, contains extensions of Theorem 2 to the case in which each  $U_n$  assumes a continuum of values and to the setting where the distributions of  $U$  and  $X^{(i)}$  are not known in advance.

## 2. DOUBLE DISJOINTNESS

DEFINITIONS. – Let  $Z = \langle \Omega_Z, \beta_Z, \mu_Z, T_Z \rangle$  and  $V = \langle \Omega_V, \beta_V, \mu_V, T_V \rangle$  be two measure preserving systems (all measure spaces we consider are Lebesgue spaces with probability measures).

(i) A joining of  $Z$  and  $V$  is a measure on  $\Omega_Z \times \Omega_V$  which is invariant under  $T_Z \times T_V$  and projects to  $\mu_Z$  and  $\mu_V$  respectively.

(ii) The systems  $Z$  and  $V$  are disjoint if their only joining is given by product measure.

(iii)  $Z$  is doubly disjoint from  $V$  if any joining of  $Z$  with an isomorphic system  $\tilde{Z}$  (i.e., a self-joining of  $Z$ ), is disjoint from  $V$ .

(iv) The definitions above, when applied to stationary stochastic processes, refer to the measure preserving systems given by the shift map.

Note the asymmetry in the definition of double disjointness. Indeed, in the following lemma, usually  $V$  is not doubly disjoint from  $U$ ; in particular, compare part (a) of the lemma with Proposition 4 (ii) below.

LEMMA 3. – In each of the following cases,  $U$  is doubly disjoint from  $V$ .

(a)  $U$  is a system with zero entropy,  $V$  is a  $K$ -system.

(b)  $U$  is a Kronecker system (i.e., a factor of an ergodic rotation on a compact abelian group) and  $V$  is weak-mixing.

(c)  $U$  is rigid and  $V$  is mild mixing (see [FW] for the definitions).

Remark. – In each of these cases, Theorem 2 gives a filtering result. In particular, by (a), Theorem 2 implies Theorem 1.

*Proof of the Lemma.* – (a) Any joining of two zero entropy systems also has zero entropy, and by [F] is disjoint from any  $K$ -system.

Parts (b), (c) are proved similarly. Note that in part (c), a joining of two rigid systems need not be rigid, but a joining of a rigid system with itself is easily seen to be rigid.  $\square$

In the converse direction we have

PROPOSITION 4. – *Let  $V$  and  $Z$  be two invertible nontrivial measure preserving systems (i.e.,  $T_V$  and  $T_Z$  are not the identity map). If  $V$  is doubly disjoint from  $Z$ , then*

- (i)  $Z$  is ergodic, and
- (ii)  $V$  has zero entropy.

*Proof.* – (i) Assume that  $A \subset \Omega_Z$  is an invariant set with  $0 < \mu(A) < 1$ . The hypothesis that  $V$  is nontrivial implies that the diagonal measure on  $\Omega_V \times \Omega_V$  is different from product measure  $\mu_V \times \mu_V$ . Therefore the measure  $\lambda$  on  $\Omega_V \times \Omega_V \times \Omega_Z$  defined by

$$\begin{aligned} & \int f(v, \tilde{v}, z) d\lambda \\ &= \int_{\Omega_V} \int_A f(v, v, z) d\mu_V(v) d\mu_Z(z) \\ & \quad + \int_{\Omega_V} \int_{\Omega_V} \int_{\Omega_Z \setminus A} f(v, \tilde{v}, z) d\mu_V(v) d\mu_V(\tilde{v}) d\mu_Z(z) \end{aligned}$$

yields a joining of two copies of  $V$  and  $Z$ , which shows that  $V$  is not doubly disjoint from  $Z$ .

(ii) A direct proof is possible (see the remark below), but it is instructive to see that this is an immediate consequence of the filtering result (Theorem 2). If  $V$  has positive entropy then by Sinai’s theorem [S] it has a Bernoulli factor, so there is a partition of  $V$  which defines a nonconstant i.i.d. process  $\{U_n\}$ , taking the values  $\{1, 2\}$ . Let  $\{X_n^{(1)}\}$  be a nonconstant  $\{0, 1\}$ -valued process defined by a partition of  $Z$ . Next, let  $\{\tilde{X}_n\}$  be an independent copy of  $\{X_n^{(1)}\}$ , and denote  $X_n^{(2)} = 1 - \tilde{X}_n$ . Since  $V$  is doubly disjoint from  $Z$ , Theorem 2 asserts that each  $U_m$  is measurable with respect to the complete  $\sigma$ -algebra spanned by the sequence  $\{X_n^{(U_n)}\}$ . However, by conditioning on the event  $[X_m^{(1)} = 1 = X_m^{(2)}]$ , which occurs with positive probability, we see that  $U_m$  is not determined by  $X_m^{(U_m)}$ . Since  $\{U_n\}$  is an i.i.d. process, this yields the desired contradiction.  $\square$

*Remark.* – Using Theorem 1 of [ST], it is easy to infer the truth of the following assertion, which implies part (ii) of the preceding proposition:

Let  $V$  be an invertible measure preserving system with positive entropy. For any nontrivial system  $Z$ , there is a joining of two copies of  $V$  which has a nontrivial common factor with  $Z$ .

*Examples.*

1. There exists a weak mixing system  $W$  with zero entropy, which has an ergodic self joining that is not weak mixing. We describe a variant of Chacon’s transformation (cf. [P], section 4.5) with these properties.

For  $n \geq 1$  and  $i, j \in \{0, 1\}$  define inductively words  $A_{ij}^n$  of length  $l_n$  as follows.

First, let  $l_0 = 2$  and  $A_{ij}^0 = ij$ . Next, assuming  $A_{ij}^n$  have been defined for all  $i, j \in \{0, 1\}$ , write

$$B_n = A_{00}^n A_{01}^n A_{10}^n A_{11}^n$$

and define

$$A_{00}^{n+1} = B_n B_n 1 B_n.$$

Finally, for every  $i, j \in \{0, 1\}$  let

$$A_{ij}^{n+1} = A_{00}^{n+1} \oplus (ijij \dots ji).$$

where both words on the right-hand side are of length

$$l_{n+1} = 12l_n + 1$$

and  $\oplus$  denotes coordinate-wise addition modulo 2. The words  $A_{00}^n$  converge, as  $n \rightarrow \infty$ , to an infinite word  $A_{00}^\infty$  and we take  $W = (\Omega_W, T_W)$  to be the orbit closure of  $A_{00}^\infty$  with respect to the shift map in  $\{0, 1\}^{\mathbb{N}}$ . Elementary block-counting shows that the topological entropy of  $W$  vanishes. One proves that  $W$  is strictly ergodic and weak-mixing with respect to the unique invariant measure  $\mu_W$ , exactly as is done for Chacon’s transformation ([P], section 4.5). For each  $i, j \in \{0, 1\}$ , the map  $f_{ij}$  defined by

$$f_{ij}(x) = x \oplus (ijij \dots)$$

acts on  $\Omega_W$  and satisfies  $T_W(f_{ij}(x)) = f_{ji}(T_W(x))$ . Therefore  $T_W \times T_W$  interchanges the graphs of  $f_{01}$  and  $f_{10}$ . Symmetry considerations and unique ergodicity imply that each  $f_{ij}$  preserves the measure  $\mu_W$ . The map  $x \rightarrow (x, f_{ij}(x))$  from  $W$  to the graph of  $f_{ij}$  sends  $\mu_W$  to a measure  $u_W^{ij}$  on this graph, and the projection of  $u_W^{ij}$  to each coordinate yields  $\mu_W$ . Thus the measure  $1/2 (\mu_W^{01} + \mu_W^{10})$  defines an ergodic self joining of  $W$  which has a factor of period 2. J. King and the referee both pointed out that a simple alternative construction with similar properties can be obtained by using two-point extensions.

2. As noted by J. King [personal communication], the construction described above immediately yields an example of two disjoint measure preserving systems, neither of which is doubly disjoint from the other. Indeed let  $R$  be the two point system and let  $B$  be any Bernoulli system. Then  $R \times B$  is disjoint from the system  $W$  of example 1 (pairwise disjointness of  $R$ ,  $B$  and  $W$  is clear; as  $B$  is disjoint from  $R \times W$  which has zero entropy, the three systems are mutually disjoint). However  $R \times B$  is not doubly disjoint from  $W$  by Proposition 4, and  $W$  is not doubly disjoint from  $R \times B$  by the previous example.

3. Rudolph's constructions of transformations with minimal self-joinings [R], show that there exist weak mixing systems which are doubly disjoint from all group rotations. Thus in this respect, the analogy with the positive entropy case breaks down.

*Remarks.*

a. In all the examples we know, when a system  $U$  is doubly disjoint from a system  $Z$  but not vice versa,  $Z$  is "more random" than  $U$ .

**Question;** If  $U$  is doubly disjoint from  $Z$ , it is necessarily *triply* disjoint from  $Z$  (i.e., is any joining of three copies of  $U$  disjoint from  $Z$ )?

Results of a similar flavor are proved in [K].

b. The analogy between parts (a) and (b) of Proposition 3 is incomplete, as there are many systems besides the Kronecker systems which are disjoint from all weak mixing systems (cf. [F] and [GW]).

**Question:** If a measure preserving system is disjoint from all weak mixing systems, is it doubly disjoint from each of them?

### 3. PROOF OF THE FILTERING THEOREM

First we establish a lemma.

LEMMA 5. – Let  $X^{(1)}, X^{(2)}, \dots$  be stationary stochastic processes with distinct marginal distributions, where  $X^{(i)} = \{X_n^{(i)}\}_{n \geq 1}$ . Let  $U = \{U_n\}$  be an  $\mathbf{N}$ -valued stationary process which is doubly disjoint from each  $X^{(i)}$ . Denote by  $\tilde{U} = \{\tilde{U}_n\}$  a process with the same distribution as  $U$ .

Suppose that we are given a joining of the processes  $X^{(1)}, X^{(2)}, \dots, U, \tilde{X}^{(1)}, \tilde{X}^{(2)}, \dots$  and  $\tilde{U}$ , such that for all  $n$  the identity

$$X_n^{(U_n)} = \tilde{X}_n^{(\tilde{U}_n)}$$

holds almost surely. Then in this joining

$$\forall n \quad U_n = \tilde{U}_n \quad \text{almost surely.}$$

*Remark.* – This lemma may be viewed as a “process level” version of Theorem 2.

*Proof of the Lemma.* – It suffices to show, for each  $i \neq j$ , that  $\mathbb{P} [U_n = i, \tilde{U}_n = j] = 0$ . Fix  $i \neq j$ . Since  $X_1^{(i)}$  and  $X_1^{(j)}$  have different laws, there exists a bounded function  $\varphi$  such that

$$E\varphi (X_1^{(i)}) \neq E\varphi (X_1^{(j)}).$$

By our hypothesis, for all  $n$

$$\varphi (X_n^{(U_n)}) = \varphi (\tilde{X}_n^{(\tilde{U}_n)}).$$

Multiplying this identity by  $1_{[U_n=i]} 1_{[\tilde{U}_n=j]}$  gives

$$\varphi (X_n^{(i)}) 1_{[U_n=i]} 1_{[\tilde{U}_n=j]} = \varphi (\tilde{X}_n^{(j)}) 1_{[U_n=i]} 1_{[\tilde{U}_n=j]}.$$

Taking expectations and using disjointness of the joining of  $U$  with  $\tilde{U}$  from  $X^{(i)}$  and from  $\tilde{X}^{(j)}$ , we find that

$$E [\varphi (X_n^{(i)})] \mathbb{P} [U_n = i, \tilde{U}_n = j] = E [\varphi (X_n^{(j)})] \mathbb{P} [U_n = i, \tilde{U}_n = j].$$

By our choice of  $\varphi$ , this forces the probability  $\mathbb{P} [U_n = i, \tilde{U}_n = j]$  to vanish.  $\square$

*Proof of Theorem 2.* – Rewriting our hypothesis, we have a measure preserving system  $X$  (the joining of all the  $X^{(i)}$ ), another system  $U$ , doubly disjoint from each  $X^{(i)}$ , and a factor map  $\psi : \Omega_X \times \Omega_U \rightarrow \Omega_Y$  from a joining  $X \vee U$  to a system  $Y = \langle \Omega_Y, \beta_Y, \mu_Y, T_Y \rangle$ , defined by  $\psi (x, u) = \{x_n^{(u_n)}\}_{n \geq 1}$ .

We must show that the projection map  $\pi_U : \Omega_X \times \Omega_U \rightarrow \Omega_Y$  is measurable with respect to the complete  $\sigma$ -algebra  $\psi^{-1} (\beta_Y)$ .

Let  $Z$  be the relatively independent joining of two copies of  $X \vee U$  over  $Y$ . Explicitly,

$$\Omega_Z = \{(x, u, \tilde{x}, \tilde{u}) \in \Omega_X \times \Omega_U \times \Omega_{\tilde{X}} \times \Omega_{\tilde{Y}} \mid \psi (x, u) = \psi (\tilde{x}, \tilde{u})\}$$

and  $\mu_Z$  is defined by

$$\begin{aligned} & \int_{\Omega_Z} g_1 (x, u) g_2 (\tilde{x}, \tilde{u}) d\mu_Z \\ &= \int_Y E [g_1 (x, u) \mid \psi (x, u) = y] E [g_2 (\tilde{x}, \tilde{u}) \mid \psi ((\tilde{x}, \tilde{u})) = y] d\mu_Y (y). \end{aligned}$$

Since  $Z$  is a joining of  $X, U, \tilde{X}, \tilde{U}$  in which  $\psi(x, u) = \psi(\tilde{x}, \tilde{u})$  holds almost surely, Lemma 5 implies that

$$\mu_Z [(x, u, \tilde{x}, \tilde{u}) \in \Omega_Z : u = \tilde{u}] = 1.$$

To complete the proof, it suffices to check that the conditional expectation operator  $E[\cdot|\psi(x, u)]$  preserves  $L^2$ -norm when applied to bounded functions  $h(u)$ . Let  $h : \Omega_U \rightarrow \mathbb{R}$  be a bounded,  $\beta_U$ -measurable function. Since  $\mu_Z$  projects to  $\mu_U$  on  $\Omega_U$  and  $u = \tilde{u}$  a.s., we have

$$\begin{aligned} \int_{\Omega_U} h(u)^2 d\mu_U &= \int_{\Omega_Z} h(u)^2 d\mu_Z(x, u, \tilde{x}, \tilde{u}) \\ &= \int_{\Omega_Z} h(u) h(\tilde{u}) d\mu_Z(x, u, \tilde{x}, \tilde{u}). \end{aligned}$$

Using the definition of  $\mu_Z$ , we conclude that

$$\begin{aligned} \int_{\Omega_U} h(u)^2 d\mu_U &= \int_{\Omega_Y} E[h(u)|\psi(x, u) = y] E[h(\tilde{u})|\psi(x, \tilde{u}) = y] d\mu_Y(y) \\ &= \int_{\Omega_Y} E[h(u)|\psi(x, u)]^2 d\mu_Y \end{aligned}$$

as claimed.  $\square$

*Remark.* – In the statement of Theorem 2, the hypothesis that the process  $X^{(i)}$  have distinct marginal distributions cannot be replaced by the assumptions that  $X^{(i)}$  have distinct distributions as processes. To see this let  $X^{(0)}$  and  $X^{(1)}$  be 0 – 1 valued Markov chains, with transition matrices  $\begin{pmatrix} p & q \\ q & p \end{pmatrix}$  and  $\begin{pmatrix} q & p \\ p & q \end{pmatrix}$  respectively, where  $p = 1 - q \neq 1/2$ . The stationary vector for both chains is  $(1/2, 1/2)$ . Let  $\{U_n\}$  be the two-point process given by

$$\mathbb{P}[U_1 = 0] = 1/2 = \mathbb{P}[U_1 = 1] \quad \text{and} \quad U_{n+1} = 1 - U_n \quad \text{for } n \geq 1.$$

Since  $\begin{pmatrix} p & q \\ q & p \end{pmatrix}^2 = \begin{pmatrix} q & p \\ p & q \end{pmatrix}$ , the two processes  $\{X_n^{(n \bmod 2)}\}_{n \geq 1}$  and  $\{X_n^{(n+1 \bmod 2)}\}_{n \geq 1}$  have the same distribution, so it is hopeless to recover  $\{U_n\}$  from  $\{X_n^{(U_n)}\}$ .

### 4. TIGHTNESS

The proof of Theorem 2 in the previous section is highly nonconstructive, and there is no control over the measure zero set which is discarded. Specifically, we would like a “Wiener-Wintner” version of Theorem 2, in which there is a set of full measure of  $U$ -sequences for which filtering is possible, that does not depend on the  $X^{(i)}$  processes.

A partial result in this direction is suggested by Lemma 5. The  $T$  be a continuous transformation on a compact metric space  $\Omega$ . A point  $\omega \in \Omega$  is called generic for a ( $T$ -invariant) measure  $\mu$ , if for all continuous real valued functions  $f$  on  $\Omega$ ,

$$\frac{1}{N} \sum_{n=1}^N f(T^n(\omega)) \rightarrow \int f d\mu \quad \text{as } N \rightarrow \infty.$$

If this holds only when  $N \rightarrow \infty$  along a subsequence  $\{N_j\}$  which does not depend on  $f$ , then  $\omega$  is called a quasi-generic point for  $\mu$ . Every point is quasi-generic for some invariant measure, as one sees by the diagonal method. When applying these definitions to a real-valued stationary process, we regard its distribution as a shift-invariant measure on the compact space  $(\mathbb{R} \cup \{\infty\})^{\mathbb{N}}$ .

**PROPOSITION 6.** – *Let  $\{u_n\}$  and  $\{v_n\}$  be two generic points for an integer valued stationary process  $U = \{U_n\}$ . Also, for each  $i \geq 1$ , let  $\{\xi_n^{(i)}\}_{n \geq 1} = \xi^{(i)}$  be a generic point for the stationary process  $X^{(i)}$ , where  $U$  is doubly disjoint from each of the processes  $X^{(i)}$ , and the  $X^{(i)}$  have distinct marginal distributions. If for all  $n \geq 1$*

$$\xi_n^{(u_n)} = \xi_n^{(v_n)},$$

*then the sequences  $\{u_n\}$  and  $\{v_n\}$  differ on a set of zero density, i.e.*

$$\frac{1}{N} \sum_{n=1}^N 1_{\{u_n \neq v_n\}} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

*Proof.* – Denote by  $\Omega_X$  the Cartesian product  $\prod_{i=1}^{\infty} \Omega_{X^{(i)}}$ , and by  $\xi$  the point  $(\xi^{(1)}, \xi^{(2)}, \xi^{(3)}, \dots)$  in  $\Omega_X$ .

Let  $\{N_j\}$  be an increasing sequence for which

$$\frac{1}{N_j} \sum_{n=1}^{N_j} 1_{\{u_n \neq v_n\}} \xrightarrow{j \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_{\{u_n \neq v_n\}}.$$

Passing to a subsequence, we may assume the point

$$(\xi, \{u_n\}, \xi, \{v_n\}) \in \Omega_X \times \Omega_U \times \Omega_X \times \Omega_U = \Omega$$

is quasi-generic for some measure  $\nu$  on  $\Omega$  with  $\{N_j\}$  the sequence in the definition of quasi-generic point. Now  $\nu$  defines a joining of  $X, U, \tilde{X}, \tilde{U}$  which satisfies the hypothesis of Lemma 5.

The conclusion of that lemma implies that the continuous function

$$f(x, \{u'_n\}, \tilde{x}, \{v'_n\}) = 1_{\{u'_1 \neq v'_1\}}$$

on  $\Omega$  has  $\int_{\Omega} f d\nu = 0$ .

Recalling the definition of quasi-generic points and our initial choice of  $\{N_j\}$ , we obtain

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_{\{u_n \neq v_n\}} = 0$$

as claimed.  $\square$

In order to sharpen the conclusion of the last proposition, further restrictions on the sequences  $\{u_n\}$  and  $\{v_n\}$  beyond genericity, are needed. This leads naturally to the following

DEFINITION. – *The discrete stationary process  $U = \{U_n\}$  is tight if there is a set of sequences  $\Lambda \subset \Omega_U$  such that  $\mathbb{P}[\{U_n\} \in \Lambda] = 1$  and for any two different sequences  $\{u_n\}, \{v_n\}$  in  $\Lambda$ , we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N 1_{\{u_n \neq v_n\}} > 0.$$

This notion was studied by Ornstein and Weiss [unpublished] who showed that a process with positive entropy is never tight, and there also exist zero entropy processes which are not tight.

On the other hand, it is easily verified that any Kronecker process (*i.e.*, a stationary process with discrete spectrum) is tight.

If the process  $\{U_n\}$  in Proposition 6 is tight, and the generic points  $\{u_n\}$  and  $\{v_n\}$  are chosen from the set  $\Lambda$  in the definition of tightness, then, of course, the conclusion of the proposition may be strengthened to  $u_n = v_n$  for all  $n$ .

If we seek a constructive filtering procedure, the most natural approach is to use a maximum-likelihood criterion.

**Question:** In the setting of Theorem 2, assume that the processes  $X^{(i)}$  have discrete marginal distributions, and define  $Y_n = X_n^{(U_n)}$ . Suppose that for every  $n \geq 1$ , there is a function  $\Gamma_n : \mathbb{R}^n \rightarrow \Omega_U$ , which assigns

to every vector  $(y_1, \dots, y_n)$  in the support of  $(Y_1, \dots, Y_n)$  a sequence  $\Gamma_n(y_1, y_2, \dots, y_n) = (v_1, v_2, v_3, \dots)$  in  $\Omega_U$ , for which the probability  $\mathbb{P}[X_1^{(v_1)} = y_1, \dots, X_n^{(v_n)} = y_n]$  is maximal. When can we ensure that the sequence of  $\Omega_U$ -valued random variables

$$\Gamma_n(X_1^{(U_1)}, X_2^{(U_2)}, X_n^{(U_n)})$$

converges almost surely (coordinate-wise) to  $(U_1, U_2, U_3, \dots)$  as  $n \rightarrow \infty$ ?

### 5. A CONTINUOUS VERSION AND UNKNOWN DISTRIBUTIONS

In [F], Theorem I.5, it is shown that if  $\{U_n\}$  and  $\{X_n\}$  are sequences of integrable real random variables which define two disjoint stationary processes, then the sum  $\{U_n + X_n\}$  determines  $\{U_n\}$ . It is noted there that the integrability assumption may be removed if we know that  $U_n$  is doubly disjoint from  $\{X_n\}$ . This remark, as well as our Theorem 2, are contained in the following.

**THEOREM 7.** – *Let  $U = \{U_n\}$  and  $X = \{X_n\}$  be real stationary stochastic processes, such that  $U$  is doubly disjoint from  $X$ . Denote by  $S$  the closed support of the distribution of  $U_1$ . Let  $\Psi : S \times \mathbb{R} \rightarrow \mathbb{R}$  be a measurable function, continuous in the first variable, such that for any two distinct points  $s \neq \tilde{s}$  in  $S$ , the distribution of the variables  $\Psi(s, X_1)$  and  $\Psi(\tilde{s}, X_1)$  are different. Define  $Y_n = \Psi(U_n, X_n)$ .*

*Then the sequence  $\{Y_n\}$  determines the sequence  $\{U_n\}$ , i.e., each  $U_m$  is measurable with respect to the complete  $\sigma$ -field spanned by the sequence of random variables  $\{Y_n\}$ .*

*Proof.* – As in the proof of Theorem 2, it suffices to establish the corresponding statement at process level. Namely, we may assume that  $\tilde{U}$  and  $\tilde{X}$  are processes with the same distributions as  $U$  and  $X$  respectively, and that a joining of  $U, X, \tilde{U}$  and  $\tilde{X}$  is given, such that

$$\forall n \quad \Psi(U_n, X_n) = \Psi(\tilde{U}_n, \tilde{X}_n) \quad \text{a.s.}$$

We must show that this implies

$$\forall n \quad U_n = \tilde{U}_n \quad \text{a.s.}$$

For  $\varepsilon > 0$ , points  $s_0, \tilde{s}_0$  in  $S$  and integer  $n$ , denote by  $A_\varepsilon = A_\varepsilon(s_0, \tilde{s}_0, n)$  the event

$$\{|U_n - s_0| < \varepsilon\} \cap \{|\tilde{U}_n - \tilde{s}_0| < \varepsilon\}$$

in the joining above. It is enough to show that for any two different points  $s_0 \neq \tilde{s}_0$  in  $S$ , the probability of  $A_\varepsilon$  vanishes for some  $\varepsilon > 0$ . As this probability does not depend on  $n$ , we take  $n = 1$ .

By the hypothesis on  $\Psi$ , there exists a continuous function of compact support  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$E [\varphi (\Psi (s_0, X_1))] > E [\varphi (\Psi (\tilde{s}_0, X_1))].$$

Denote by  $\Phi (s)$  the random variable  $\varphi (\Psi (s, X_1))$ . From our second hypothesis,

$$\Phi (U_1) = \Phi (\tilde{U}_1).$$

Rewrite this in the form

$$1_{A_\varepsilon} [\Phi (s_0) - \Phi (\tilde{s}_0)] = 1_{A_\varepsilon} [\Phi (s_0) - \Phi (U_1) - \Phi (\tilde{s}_0) + \Phi (\tilde{U}_1)].$$

Next, we take expectations, using the independence of  $\Phi (s)$  (for fixed  $s$ ) from the indicator  $1_{A_\varepsilon}$ , which is a function of the pair  $(U_1, \tilde{U}_1)$ . We get

$$\begin{aligned} & \mathbb{P} (A_\varepsilon) [E\Phi (s_0) - E\Phi (\tilde{s}_0)] \\ & \leq \mathbb{P} (A_\varepsilon) \left[ \sup_{|s-s_0|<\varepsilon} |E\Phi (s_0) - E\Phi (s)| + \sup_{|s-\tilde{s}_0|<\varepsilon} |E\Phi (\tilde{s}_0) - E\Phi (s)| \right] \end{aligned}$$

The left continuity of  $\Psi$  implies that  $E\Phi (s)$  depends continuously on  $s$ , and therefore the last inequality forces  $\mathbb{P} (A_\varepsilon) = 0$  for sufficiently small  $\varepsilon > 0$ .  $\square$

**Unknown distributions.**

Returning to the setting in which the process  $\{U_n\}$  takes integer values, we again consider the problem of determining  $\{U_n\}$  from  $\{X_n^{(U_n)}\} = \{Y_n\}$  as in the previous sections, but now we do not assume that the distributions of  $U$  and  $X^{(i)}$  are given *a priori*. Of course in this situation one can permute the processes  $X^{(i)}$ , and apply the same permutation to the values taken by  $\{U_n\}$ , without affecting the output sequence  $\{Y_n\}$ . Under an appropriate disjointness condition, this is the only remaining ambiguity, as is shown by the following analogue of Proposition 6.

PROPOSITION 8. – *Let  $U = \{U_n\}$  and  $V = \{V_n\}$  be  $\mathbf{N}$ -valued stationary processes. Suppose that for each  $i \geq 1$ ,  $X^{(i)}$  and  $Z^{(i)}$  are real stationary processes. We assume that*

(a) *For every  $i \neq j$  the processes  $X^{(i)}$  and  $X^{(j)}$  have different marginal distributions, and similarly for  $Z^{(i)}$  and  $Z^{(j)}$ .*

(b) *Every joining of  $U$  and  $V$  is disjoint from each  $X^{(i)}$  and also from each  $Z^{(i)}$ .*

(c)  $\{u_n\}, \{v_n\}, \{\xi_n^{(i)}\}_{n=1}^\infty$  and  $\{\zeta_n^{(i)}\}_{n=1}^\infty$  are generic sequences for  $U, V, X^{(i)}$  and  $Z^{(i)}$  respectively (for each  $i \geq 1$ ).

(d) The identity

$$\xi_n^{(u_n)} = \zeta_n^{(v_n)}$$

holds for all  $n \geq 1$ .

Then there exists a permutation  $\pi : \mathbf{N} \rightarrow \mathbf{N}$  such that for every  $i \geq 1$ , the distributions of  $X^{(i)}$  and  $Z^{(\pi(i))}$  are identical and also

$$\frac{1}{N} \sum_{n=1}^N 1_{\{\pi(u_n)=v_n\}} \rightarrow 1 \quad \text{as} \quad N \rightarrow \infty.$$

The proof proceeds along the lines of Proposition 6, i.e., one first establishes an analogue of Lemma 5 with  $\tilde{U}$  and  $\tilde{X}^{(i)}$  replaced by  $V$  and  $Z^{(i)}$  respectively and then deduces the assertion.

*Remark.* – Observe that hypothesis (b) in the last proposition is satisfied if  $U$  and  $V$  have zero entropy and all the processes  $X^{(i)}, Z^{(i)}$  are  $K$ -processes.

ACKNOWLEDGEMENTS

We are grateful to J. King, S. Mozes and J. P. Thouvenot for helpful conversations. The second author would like to thank J. F. Mela and F. Parreau for their hospitality at the University of Paris-Nord.

REFERENCES

[F] H. FURSTENBERG, Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation, *Math. systems theory*, **1**, 1967, pp. 1-49.  
 [FW] H. FURSTENBURG and B. WEISS, The finite multipliers of infinite ergodic transformations, *Lecture Notes in Math. #668*, Springer Verlag, 1977, pp. 127-132.  
 [GW] S. GLASNER and B. WEISS, Processes disjoint from weak mixing, *Trans. Amer. Math. Soc.*, **316**, 1989, pp. 689-703.  
 [K] J. L. KING, Ergodic properties where order 4 implies infinite order, preprint, 1991.  
 [P] K. PETERSEN, Ergodic theory, *Cambridge University Press*, 1983.  
 [R] D. J. RUDOLPH, An example of a measure preserving map with minimal self joinings and applications, *J. d'Analyse Math.*, **35**, 1979, pp. 98-122.  
 [S] Ja G. SINAI, A weak isomorphism of transformations with an invariant measure, *Dokl. Akad. Nauk., USSR*, **147**, 1962, pp. 797-800.  
 [ST] M. SMORODINSKI and J. P. THOUVENOT, Bernoulli factors that span a transformation, *Israel J. of Math.*, **32**, 1979, pp. 39-43.

(Manuscript received February 22, 1993;  
 revised September 30, 1993.)