The Annals of Probability
1984, Vol. 12, No. 1, 204–211

A SIMPLE DEVELOPMENT OF THE THOUVENOT RELATIVE ISOMORPHISM THEORY

BY JOHN C. KIEFFER[1]

University of Missouri-Rolla

A simple proof is given of the Thouvenot relative isomorphism theorem for conditional very weak Bernoulli processes. As a special case, one obtains a simple proof of the Ornstein isomorphism theorem for very weak Bernoulli processes.

I. Introduction. For the purposes of this paper, a *process* is a bilateral sequence $X = \{X_i\}_{i=-\infty}^{\infty}$ of random variables defined on some common probability space $((\Omega, \mathscr{F}, P)$, say) and taking their values in a common *finite* set (A, say), called the *state space* of the process. Regarding X as a random sequence mapping Ω into A^∞, the measurable space of bilateral sequences from A, we let $\operatorname{dist} X$ denote the probability measure on A^∞ into which P is carried by X. A *pair process* (X, Y) is a pair of processes X, Y defined on the same probability space $((\Omega, \mathscr{F}, P)$, say) with possibly different state spaces (A, B, respectively, say). Regarding the pair process (X, Y) as a map from Ω to $A^\infty \times B^\infty$, we let $\operatorname{dist}(X, Y)$ denote the probability measure on $A^\infty \times B^\infty$ into which P is carried by (X, Y). The pair process (X, Y) is said to be *stationary* (*ergodic*) if X, Y are jointly stationary (jointly ergodic) processes. Similar comments apply to processes consisting of more than two component processes (triple processes, quadruple processes, etc.).

We say the stationary pair processes (X, Y) and (U, V) are *relatively isomorphic* if there is process Z such that

(i) $\operatorname{dist}(X, Z) = \operatorname{dist}(U, V)$, and

(ii) (X, Y) and (X, Z) are almost surely stationary codings of each other.

(By (ii), we mean that, almost surely, $(X, Y) = \phi(X, Z)$ and $(X, Z) = \psi(X, Y)$ for some pair of measurable maps (ϕ, ψ) which commute with the shifts on the respective sequence spaces.)

Consider a pair process (X, Y) where X is a stationary and ergodic process and Y is a stationary independent process statistically independent of X. We state here the solution of Thouvenot ([12], Proposition 3 and [13], Lemma 6) to the problem of determining which pair processes are relatively isomorphic to (X, Y). Let (U, V) be a pair process; we say V is *U-conditionally very weak Bernoulli* (VWB) if (U, V) is stationary and ergodic and

$$(1) \qquad \lim_{m\to\infty} E\bar{d}_m(\operatorname{dist}(V_1^m \mid U), \operatorname{dist}(V_1^m \mid V_{-\infty}^0, U)) = 0,$$

where $\bar{d}_m$ denotes the $\bar{d}$-distance between pairs of m-dimensional probability

---

Received July 1982.
[1]Research of author supported by NSF Grant ECS-7821335.
AMS 1980 *subject classifications*. Primary 28D20, 60G10; secondary 94A24.
Key words and phrases. Relatively isomorphic pair processes, relativized $\bar{d}$-distance, conditional very weak Bernoulli processes, conditional finitely determined processes.

distributions [6] and the arguments of $\bar{d}_m$ in (1) are the random measures such that when the bilateral random sequence $U = u$ and the unilateral random sequence $V_{-\infty}^0 = (\cdots, V_{-1}, V_0)$ takes as a value the particular sequence of symbols $(\cdots, v_{-1}, v_0)$, then $\text{dist}(V_1^m \mid U)$ equals the conditional distribution of the random vector $V_1^m = (V_1, \cdots, V_m)$ given $U = u$ and $\text{dist}(V_1^m \mid V_{-\infty}^0, U)$ equals the conditional distribution of V_1^m given $V_{-\infty}^0 = (\cdots, v_{-1}, v_0)$ and $U = u$. Thouvenot showed that (U, V) is relatively isomorphic to (X, Y) if and only if

(a) dist $U =$ dist X:
(b) V is U-conditionally VWB; and
(c) The entropy rates $H(X, Y)$, $H(U, V)$ of the pair processes (X, Y), (U, V) are equal,

where, as usual, the entropy rate $H(X, Y)$ of the pair process (X, Y) is defined to be

$$\lim_{n\to\infty} n^{-1}[\text{entropy of } (X_1, X_2, \cdots, X_n, Y_1, Y_2, \cdots, Y_n)].$$

(Similarly, one defines the entropy rate $H(X)$ of a single process X, the entropy rate $H(X, Y, Z)$ of a triple process (X, Y, Z), etc.)

Note that if in Thouvenot's result one takes X to be a degenerate process, one obtains as a special case the solution of Ornstein [6] to the problem of determining which processes are isomorphic to independent processes. Thouvenot's proof is somewhat complex. In this paper we give a simple proof of Thouvenot's result by means of a lemma (Lemma 1 of Section III) whose use makes unnecessary the involved arguments involving the marriage lemma appearing in previous proofs.

II. Conditionally finitely determined processes. If (X, Y), (U, V) are two stationary pair processes such that dist $X =$ dist U and Y, V have the same state space, then $\bar{d}_r((X, Y), (U, V))$, the relativized $\bar{d}$-distance between the two pair processes [12], is the infimum of $\text{Prob}[\tilde{Y}_0 \neq \tilde{V}_0]$ over all stationary triple processes $(\tilde{X}, \tilde{Y}, \tilde{V})$ such that $\text{dist}(\tilde{X}, \tilde{Y}) = \text{dist}(X, Y)$ and $\text{dist}(\tilde{X}, \tilde{V}) = \text{dist}(U, V)$. Following [12], if (U, V) is a pair process, we say V is *U-conditionally finitely determined* (FD) if (U, V) is stationary and ergodic and $\bar{d}_r((U^{(n)}, V^{(n)}), (U, V)) \to 0$ for every sequence of stationary and ergodic pair processes $\{(U^{(n)}, V^{(n)})\}_{n=1}^{\infty}$ such that

(i) For all n, dist $U^{(n)} =$ dist U and $V^{(n)}$ and V have the same state space;
(ii) $(U^{(n)}, V^{(n)}) \to_d (U, V)$, where the "$d$" denotes convergence in distribution; and
(iii) $H(U^{(n)}, V^{(n)}) \to H(U, V)$.

Our proof of Thouvenot's theorem will use the result that V is U-conditionally VWB if and only if V is U-conditionally FD. It is not hard to show that the first condition implies the second (see [13], proof of Lemma 6). The original proof [7] that the second condition implies the first is hard but an easy proof is now known [4].

III. A relativized Sinai theorem. Sinai [10] showed that if Y is a stationary and ergodic process and if U is a stationary independent process whose entropy rate is no greater than that of Y, then U is a factor of Y in the sense that there is a stationary coding $\hat{U}$ of Y for which dist $\hat{U}$ = dist U. Ornstein [6] generalized this result to the case of a very weak Bernoulli process U. Ornstein [5] also proved a relativized version of Sinai's theorem. In this section we generalize Ornstein's relativized result to show that if (X, Y) is a stationary, ergodic pair process and (U, V) is a pair process such that V is U-conditionally VWB and dist U = dist X and $H(U, V) \leq H(X, Y)$, then there is a process Z which is a stationary coding of (X, Y) and for which $\mathrm{dist}(X, Z) = \mathrm{dist}(U, V)$. For use in our proof of Thouvenot's theorem (see Section V), we prove a bit more, namely the following.

THEOREM 1. *Let (X, Y) be a stationary and ergodic pair process. Let (U, V) be a pair process such that V is U-conditionally* VWB *and* dist U = dist X *and* $H(X, Y) \geq H(U, V)$. *Let the process Z be a stationary coding of (X, Y) for which $\bar{d}_r((X, Z), (U, V)) < \varepsilon$. Then there is a stationary coding $\hat{Z}$ of (X, Y) for which*

(i) $\mathrm{dist}(X, \hat{Z}) = \mathrm{dist}(U, V)$, *and*
(ii) $\mathrm{Prob}[\hat{Z}_0 \neq Z_0] < \varepsilon$.

Our main tool for proving Theorem 1 is the following lemma, whose proof is given in the next section. This lemma is of interest not only in the development of the Thouvenot theory, but in other contexts as well. For example, a variant of this lemma was used in [3] to prove some multiterminal coding theorems of information theory.

LEMMA 1. *Let (X, Y, U) be a stationary and ergodic triple process. Suppose also that the pair process (X, Y) is aperiodic (meaning that* $\mathrm{Prob}[X = x, Y = y] = 0$ *for all x, y) and that $H(X, Y) \geq H(X, U)$. Then there is a sequence of stationary processes $\{U^{(n)}\}_{n=1}^{\infty}$, having the same state space as U, such that*

(i) *each $U^{(n)}$ is a stationary coding of (X, Y);*
(ii) $(X, Y, U^{(n)}) \to_d (X, Y, U)$; *and*
(iii) $H(X, U^{(n)}) \to H(X, U)$.

PROOF OF THEOREM 1.

CASE 1. $H(X, Y) = 0$. Then $H(U, V) = 0$ and so by Lemma A1 of the appendix, V is a stationary coding of U almost surely. Because of this and the fact that $\bar{d}_r((X, Z), (U, V)) < \varepsilon$, it is not hard to show there must be a coding $\hat{V}$ of X satisfying $\mathrm{Prob}[Z_0 \neq \hat{V}_0] < \varepsilon$ and $\mathrm{dist}(X, \hat{V}) = \mathrm{dist}(U, V)$.

CASE 2. $H(X, Y) > 0$. Then (X, Y) is aperiodic. Find positive numbers $\varepsilon_1, \varepsilon_2, \cdots$ so that $\sum_{i=1}^{\infty} \varepsilon_i < \varepsilon$ and $\bar{d}_r((X, Z), (U, V)) < \varepsilon_1$. Redefining (X, Y) on a new probability space if necessary, find a process $\hat{V}$ such that $(X, Y, \hat{V})$ is a stationary and ergodic triple process for which $\mathrm{dist}(X, \hat{V}) = \mathrm{dist}(U, V)$ and $\mathrm{Prob}[Z_0 \neq \hat{V}_0]$

$< \varepsilon_1$. (The existence of such a triple process which is stationary follows from the definition of the metric $\bar{d}_r$. If the triple process is not ergodic, it can be replaced by an ergodic one by appealing to the ergodic decomposition theorem.) Applying Lemma 1 to the triple process $(X, (Y, Z), \hat{V})$, and using the fact that $\hat{V}$ is X-conditionally FD, we may find a stationary coding $Z^{(1)}$ of (X, Y) such that $(X, (Y, Z), Z^{(1)})$ is so close in distribution to $(X, (Y, Z), \hat{V})$ and $H(X, Z^{(1)})$ is so close to $H(X, \hat{V})$ that $\mathrm{Prob}[Z_0 \neq Z_0^{(1)}] < \varepsilon_1$ and $\bar{d}_r((X, Z^{(1)}), (U, V)) < \varepsilon_2$. Applying this argument repeatedly, obtain stationary codings $Z^{(2)}, Z^{(3)}, \cdots$ of (X, Y) so that for each $i = 2, 3, \cdots, \bar{d}_r((X, Z^{(i)}), (U, V)) < \varepsilon_{i+1}$ and $\mathrm{Prob}[Z_0^{(i-1)} \neq Z_0^{(i)}] < \varepsilon_i$ hold. Then the process Z which is the almost sure limit of the processes $\{Z^{(i)}\}_{i=1}^{\infty}$ is the process we seek.

Combining Lemma 1 and Theorem 1, one can obtain the following modification of Lemma 1 which will be of use to us later.

LEMMA 2. *Let (X, Y, U) be a stationary and ergodic triple process such that (X, Y) is aperiodic, U is X-conditionally* VWB, *and $H(X, Y) \geq H(X, U)$. Then there is a sequence $\{U^{(n)}\}$ of stationary codings of (X, Y) such that*

(i) $(X, Y, U^{(n)}) \rightarrow_d (X, Y, U)$
(ii) $\mathrm{dist}(X, U^{(n)}) = \mathrm{dist}(X, U)$ *for all n.*

IV. Proof of basic lemma. In this section we prove Lemma 1, which was basic to our approach to the Thouvenot theory. (The reader who is not interested in how Lemma 1 is proved may skip to Section V.) First, we introduce some notation and terminology to be used in this section.

In the following, if Z is a process let Z^n denote the random vector $(Z_0, \cdots, Z_n)$. If $m > N \geq 1$ and $\delta > 0$, we say an m-sequence $\underset{\sim}{z}$ is (N, δ) typical of Z if the frequency with which each N-sequence b appears in $\underset{\sim}{z}$ is within δ of $\mathrm{Prob}[Z^N = b]$. Similarly, if (X, Y) is a pair process, we can define what it means for a pair of m-sequences $(\underset{\sim}{x}, \underset{\sim}{y})$ to be jointly (N, δ) typical of (X, Y), and if (X, Y, U) is a triple process, we can define what it means for a triple of m-sequences $(\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{u})$ to be (N, δ) typical of (X, Y, U). If S is a set, $|S|$ is used to denote the number of elements in S. Finally, if X, Y are processes, $H(X \mid Y)$ denotes the conditional entropy rate for the process X given the process Y, which is equal to $H(X, Y) - H(Y)$.

Lemma 1 will follow from Lemmas 3 and 4 below. We omit the proof of Lemma 3, as it is a simple consequence of the Shannon-McMillan Theorem.

LEMMA 3. *Let (X, Y, U) be a stationary and ergodic pair process, where X has state space A, Y has state space B, and U has state space C.*

(i) *Suppose for $n = 1, 2, \cdots, E_n$ is a subset of $A^n \times C^n$ for which*

$$\limsup_{n\to\infty} n^{-1} \log |\{u \in C^n : (X^n, u) \in E_n\}| < H(U \mid X) \quad \text{a.s.}$$

Then $\mathrm{Prob}[(X^n, U^n) \in E_n] \rightarrow 0$.

(ii) *Suppose for $n = 1, 2, \cdots, F_n$ is a subset of $A^n \times B^n \times C^n$ for which* $\limsup_{n\to\infty} \mathrm{Prob}[(X^n, Y^n, U^n) \in F_n] > 0$. *Then there exists for each n a* *map ϕ_n: $A^n \to C^n$ such that the following inequality holds with positive probability*:

$$\limsup_{n\to\infty} n^{-1}\log |\{y \in B^n : (X^n, y, \phi_n(X^n)) \in F_n\}| \geq H(Y \mid U, X).$$

LEMMA 4. *Assume in addition to the hypotheses of Lemma 3 that α, β are positive numbers for which $H(Y \mid X) - H(U \mid X) < \alpha < \beta < H(Y \mid U, X)$. For each $n = 1, 2, \cdots$, and each $x \in A^n$, let $\{(B_i^n(x), u_i^n(x)) : i = 1, \cdots, k_n(x)\}$ be a largest possible set of pairs such that*:

(a) *The coordinates $B_i^n(x)$ ($i = 1, \cdots, k_n(x)$) are disjoint subsets of B^n each having between $2^{n\alpha}$ and $2^{n\beta}$ elements*;
(b) *The coordinates $u_i^n(x)$ ($i = 1, \cdots, k_n(x)$) are distinct elements of C^n*;
(c) *$(x, y, u_i^n(x))$ is (N, δ) typical of (X, Y, U), $y \in B_i^n(x)$, $i = 1, \cdots, k_n(x)$.*

Then if N is large enough and δ is small enough,

$$\operatorname{Prob}[Y^n \in \cup_{i=1}^{k_n(X^n)} B_i^n(X^n)] \to 1.$$

PROOF. Fix $\tau > 0$ so that $H(Y \mid X) - H(U \mid X) + \tau < \alpha$. Take N large enough and δ small enough so that

$$\limsup_{n\to\infty} n^{-1}\log |\{y \in B^n : (X^n, y) \text{ is } (N, \delta) \text{ typical of } (X, Y)\}| \leq H(Y \mid X) + \tau \quad \text{a.s.} \tag{2}$$

Then since the argument of the logarithm in (2) is at least $2^{n\alpha}k_n(X^n)$,

$$\limsup_{n\to\infty} n^{-1}\log k_n(X^n) \leq H(Y \mid X) + \tau - \alpha < H(U \mid X) \quad \text{a.s.}$$

By part (i) of Lemma A3,

$$\operatorname{Prob}[U^n \in \{u_i^n(X^n) : i = 1, \cdots, k_n(X^n)\}] \to 0. \tag{3}$$

Take F_n to be the set of all triples (x, y, u) from $A^n \times B^n \times C^n$ which are (N, δ) typical of (X, Y, U) and satisfy $y \notin \cup_{i=1}^{k_n(x)} B_i^n(x)$, $u \notin \{u_i^n(x) : i = 1, \cdots, k_n(x)\}$. By part (ii) of Lemma 3 and the maximality (in terms of cardinality) of the set of pairs $\{(B_i^n(x), u_i^n(x)\}$ relative to the conditions (a)–(c), we must have

$$\operatorname{Prob}[(X^n, Y^n, U^n) \in F_n] \to 0. \tag{4}$$

The conclusion of the lemma follows from (3), (4) and the fact that $\operatorname{Prob}[(X^n, Y^n, U^n)$ is (N, δ) typical$] \to 1$.

PROOF OF LEMMA 1. We can assume $H(Y \mid X, U) > H(Y \mid X) - H(U \mid X)$. (For, if $H(U) = 0$ the lemma is trivially true because any sequence $\{U^{(n)}\}$ satisfying (ii) of Lemma 1 also satisfies (iii). If $H(U) > 0$, it is not hard to show that there exists a sequence of processes $\{\hat{U}^{(n)}\}$ jointly ergodic with (X, Y) for which $(X, Y, \hat{U}^{(n)}) \to (X, Y, U)$, $H(X, \hat{U}^{(n)}) \to H(X, U)$, and $H(Y \mid X, \hat{U}^{(n)}) > H(Y \mid X) - H(U^{(n)} \mid X) \geq 0$ for all n.) Employ Lemma 4 to get a certain sequence of block encodings of (X, Y), and then by the standard technique from information theory (see, for example, [2], page 960) via the strong form [9, page 22] of the Rokhlin-Kakutani theorem (valid since (X, Y) is aperiodic), replace the block encodings by sliding-block encodings (finite stationary encodings) $\{U^{(n)}\}$ of (X, Y) so that (ii) of Lemma 1 holds and $\limsup_{n\to\infty} H(Y \mid X, U^{(n)}) \le H(Y \mid X) - H(U \mid X)$, whence $\liminf_{n\to\infty} H(X, U^{(n)}) \ge H(X, U)$. By uppersemicontinuity of the entropy rate with respect to convergence in distribution, one has automatically that $\limsup_{n\to\infty} H(X, U^{(n)}) \le H(X, U)$. Hence, (iii) of Lemma 1 holds.

V. Proof of the Thouvenot Theorem. Here is the result of Thouvenot we wish to prove.

THEOREM 2. *Let (X, Y), (U, V) be pair processes such that Y is X-conditionally* VWB, *V is U-conditionally* VWB, *$H(X, Y) = H(U, V)$ and* dist $X =$ dist U. *Then (X, Y), (U, V) are relatively isomorphic. Furthermore, if Z is a stationary coding of (X, Y) for which $\bar{d}_r((X, Z), (U, V)) < \varepsilon$, there is a process $\hat{Z}$ such that*

(i) *$(X, \hat{Z})$ and (X, Y) are almost surely stationary codings of each other;*
(ii) dist$(X, \hat{Z}) =$ dist(U, V), *and*
(iii) Prob$[\hat{Z}_0 \ne Z_0] < \varepsilon$.

For the proof of Theorem 2, we will need Theorem 1 plus the following two lemmas.

LEMMA 5. *Let the pair process (X, Y) be stationary and ergodic. Let Z be a stationary coding of (X, Y) for which $H(X, Z) = H(X, Y)$. Then for any $\varepsilon > 0$, there is a stationary coding $\hat{Z}$ of (X, Y), whose state space contains that of Z, and which satisfies*

(i) Prob$[\hat{Z}_0 \ne Z_0] < \varepsilon$
(ii) *$H(Y_0 \mid X, \hat{Z})$, the conditional uncertainty for Y_0 given $(X, \hat{Z})$, is smaller than ε.*

NOTE. Roughly speaking, Lemma 5 says that any stationary coding (X, Z) of (X, Y) of full entropy (i.e., $H(X, Z) = H(X, Y)$) is "almost" relatively isomorphic to (X, Y).

PROOF OF LEMMA 5. Since $H(Y \mid X, Z)$, the conditional entropy rate of the process Y given the pair process (X, Z), is zero, the Slepian-Wolf theorem [11] [1] implies the existence of a binary process U which is a stationary coding of Y, and for which $H(Y_0 \mid U, X, Z) < \varepsilon$ and Prob$[U_0 = 1] < \varepsilon$ both hold. Define $\hat{Z}$ to be the process

$$\hat{Z}_i = Z_i \quad \text{if} \quad U_i = 0 \quad \hat{Z}_i = (Z_i, U_i) \quad \text{if} \quad U_i = 1.$$

LEMMA 6. *Let (X, Y) be a pair process such that Y is X-conditionally* VWB. *Let Z be a stationary coding of (X, Y) such that $H(X, Z) = H(X, Y)$. Then for any $\varepsilon > 0$ there is a process $\hat{Z}$ such that*

(a) *(X, Y) and $(X, \hat{Z})$ are almost surely stationary codings of each other:*
(b) dist$(X, \hat{Z}) =$ dist(X, Z), *and*
(c) Prob$[\hat{Z}_0 \ne Z_0] < \varepsilon$.

PROOF. We can assume (X, Y) is aperiodic. (Otherwise, by Lemma A1 of the Appendix, Y is a stationary coding of X and Lemma 6 is trivially true with $\hat{Z} = Z$.) It suffices to prove the weaker result that there exists a stationary coding $\hat{Z}$ of (X, Y) satisfying (b), (c), and

(d) $H(Y_0 \mid X, \hat{Z}) < \varepsilon$.

(Just apply this weaker result repeatedly to get Lemma 6.) By Lemma 5, find a stationary coding W of (X, Y) such that (i) $\mathrm{Prob}[W_0 \neq Z_0] < \varepsilon/2$ and (ii) $H(Y_0 \mid X, W) < \varepsilon$. By Lemma A2 of the Appendix, W is X-conditionally VWB; also $H(X, W) \leq H(X, Z)$ and $\bar{d}_r((X, W), (X, Z)) < \varepsilon/2$ hold and so by Theorem 1, redefining X, Y on a new probability space if necessary, there is a process $\tilde{Z}$ such that the triple process $(X, Y, \tilde{Z})$ is stationary and ergodic, (iii) $\mathrm{dist}(X, \tilde{Z}) = \mathrm{dist}(X, Z)$, (iv) $\mathrm{Prob}[\tilde{Z}_0 \neq W_0] < \varepsilon/2$, and (v) W is a stationary coding of $(X, \tilde{Z})$.

Then from (ii) and (v) we have (vi) $H(Y_0 \mid X, \tilde{Z}) < \varepsilon$. Applying Lemma 2 to the triple process $(X, (Y, W), \tilde{Z})$, we can just as well assume that (iii), (iv) and (vi) hold when $\tilde{Z}$ is replaced by an appropriate process $\hat{Z}$ which is stationary coding of (X, Y). Thus we have (b) and (d). Also, (c) follows from (iv) with $\tilde{Z}$ replaced by $\hat{Z}$ and (i).

PROOF OF THEOREM 2. By Theorem 1, find a stationary coding $\tilde{Z}$ of (X, Y) such that $\mathrm{dist}(X, \tilde{Z}) = \mathrm{dist}(U, V)$ and $\mathrm{Prob}[\tilde{Z}_0 \neq Z_0] < \varepsilon$. Then apply Lemma 6.

APPENDIX

LEMMA A1. *Let (U, V) be a pair process such that V is U-conditionally* VWB *and $H(V \mid U) = 0$. Then V is a stationary coding of U.*

PROOF. The left side of (1) is no less than

$$\lim_{m\to\infty} m^{-1} \textstyle\sum_{i=1}^{m} E\bar{d}_1(\mathrm{dist}(V_i \mid U), \mathrm{dist}(V_i \mid V_{-\infty}^{0}, U))$$
$$= \lim_{m\to\infty} E\bar{d}_1(\mathrm{dist}(V_0 \mid U), \mathrm{dist}(V_0 \mid V_{-\infty}^{-m}, U)).$$

This latter quantity then being zero, we have $\lim_{m\to\infty} H(V_0 \mid V_{\infty}^{-m}, U) = H(V_0 \mid U)$. The left side of this equality is zero, which we see by applying Pinsker's formula for $H(V \mid U)$ [8, page 24, equation (28)]. Therefore the right side is zero.

LEMMA A2. *Let (X, Y) be a pair process such that Y is X-conditionally* VWB. *Let Z be a stationary coding of (X, Y). Then Z is also X-conditionally* VWB.

PROOF. Clearly (X, Z) is a $\bar{d}_r$-limit of pair processes of form $(X, \hat{Z})$, where $\hat{Z}$ is a finite stationary coding (sliding-block coding) of (X, Y). It is easy to see directly from the definition of the conditional VWB concept that $\hat{Z}$ is X-conditionally VWB for such a pair process $(X, \hat{Z})$. Now apply the result that the conditional VWB property is stable under the taking of $\bar{d}_r$-limits. (An easy proof of the fact may be found in [4].)

REFERENCES

[1] KIEFFER, J. C. (1980a). Some universal multiterminal source coding theorems. *Inform. and Control* **46** 93–107.

[2] KIEFFER, J. C. (1980b). On the transmission of Bernoulli sources over stationary channels. *Ann. Probab.* **8** 942–961.

[3] KIEFFER, J. C. (1981). A method for proving multiterminal source coding theorems. *IEEE Trans. Inform. Theory* **27** 565–570.

[4] KIEFFER, J. C. (1982). A direct proof that VWB processes are closed in the $\bar{d}$-metric. *Israel J. Math.* **41** 154–160.

[5] ORNSTEIN, D. S. (1971). Two Bernoulli shifts with infinite entropy are isomorphic. *Adv. in Math.* **5** 339–348.

[6] ORNSTEIN, D. S. (1974). *Ergodic Theory, Randomness, and Dynamical Systems*, Yale Univ. Press, New Haven, Conn.

[7] RAHE, M. (1978). Relatively finitely determined implies relatively VWB. *Canadian J. Math.* **30** 531–548.

[8] ROKHLIN, V. A. (1967). Lectures on the entropy theory of measure preserving transformations. *Russian Math. Surveys* **22** 1–52.

[9] SHIELDS, P. C. (1973). *The Theory of Bernoulli Shifts*. Univ. Chicago Press.

[10] SINAI, Y. G. (1962). Weak isomorphism of transformations with invariant measure. *Dokl. Akad. Nauk SSSR* **147** 797–800.

[11] SLEPIAN, D. and WOLF, J. (1973). Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory* **19** 417–480.

[12] THOUVENOT, J.-P. (1975a). Quelques proprietes des systems dynamiques qui se decomposent en un produit de deux systems dont l'un est un schema de Bernoulli, *Israel J. Math.* **21** 177–207.

[13] THOUVENOT, J.-P. (1975b). Remarques sur les systems dynamiques donnes avec plusieurs facteurs. *Israel J. Math.* **21** 215–232.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MISSOURI-ROLLA
ROLLA, MISSOURI 65401