ISRAEL JOURNAL OF MATHEMATICS, Vol. 32, No. 4, 1979

FINITE CODING FACTORS OF MARKOV GENERATORS

BY
M. RAHE

ABSTRACT

Using Thouvenot's relativized isomorphism theory, the author develops a conditionalized version of the Friedman–Ornstein result on Markov processes. This relativized statement is used to study the way in which a factor generated by a finite length stationary coding sits in a Markov process. All such factors split off if they are maximal in entropy. Moreover, one can show that if a finite coding factor fails to split off, it is relatively finite in a larger factor which either generates or itself splits off.

1. Introduction

In the following we shall be concerned with an invertible measure preserving transformation T acting on a probability space $(X, \mathcal{B}, \mu)$ which we take to be isomorphic to the unit interval with Lebesgue sets and Lebesgue measure. P and H are two measurable finite partitions of X, while (T, P) and (T, H) are the associated stationary stochastic processes. We shall assume that $P \vee H$ generates, i.e. that $\bigvee_{-\infty}^{\infty} T^i(P \vee H) = (P \vee H)_{-\infty}^{\infty} = \mathcal{B}$.

Associated with a finite partition H, we shall consider two sub-σ-algebras of $\mathcal{B}$. We use the symbol $\sigma(H)$ to indicate the possibly non-T-invariant σ-algebra consisting of finite disjoint unions of sets in H. By the factor $\mathcal{H} = H_{-\infty}^{\infty}$ generated by H we mean the smallest T-invariant sub-σ-algebra of $\mathcal{B}$ containing H and all subsets of sets of measure zero. A fiber in $\mathcal{H}$ is the collection of all points in X which share the same doubly infinite H-name. For a fixed version of probability conditioned on $\mathcal{H}$ and a fixed fiber h of $\mathcal{H}$, we study the behavior of the sequence $\{T^iP\}$, $i = \cdots, -1, 0, 1, \cdots$ on h, viewed as a non-stationary process that the partition P and the transformation T induce on it.

Received January 5, 1978 and in revised form November 12, 1978

It is known from the work of Friedman and Ornstein [1] that if (T, P) is a finite state n-step Markov process which is mixing, then (T, P) satisfies a condition called weak Bernoulli, which in turn implies that (T, P) is isomorphic to a Bernoulli shift. Moreover, if such a process fails to be mixing, the structure is of a simple type: (T, P) is isomorphic to the direct product of a Bernoulli shift with a rotation on a finite number of points.

In [4] J.-P. Thouvenot showed that much of the Ornstein theory continues to hold when one examines the behavior of processes relative to an invariant sub-σ-algebra $\mathscr{H}$. The result in the Thouvenot theory that corresponds to the Friedman–Ornstein theorem is that a process which is n-step mixing Markov relative to $\mathscr{H}$ is isomorphic to the direct product of (T, H) and a Bernoulli shift.

In the first part of this note we shall quickly sketch the conditionalized version of the Friedman–Ornstein result. Following that development, we apply the result to study the factors generated by finite codings of finite state n-step Markov generators. Here, we say a partition P is an n-step Markov generator if $\mathscr{B} = P_{-\infty}^{\infty}$ and $\mu(P \mid P_{-\infty}^{-1}) = \mu(P \mid P_{-n}^{-1})$. By a finite coding of length l we mean a partition $H \in P_0^{l-1}$ for some non-negative integer l. When l is one, we shall refer to H as a clumping of P.

One of the motivations for studying such objects is the observation by J.-P. Thouvenot that a clumping factor of an independent generator always splits off, i.e., that there exists a finite partition B of X such that the sequence $\{T^iB\}$, $i = \cdots, -1, 0, 1, \cdots$, is independent, $B_{-\infty}^{\infty} \perp H_{-\infty}^{\infty}$, and $(B \vee H)_{-\infty}^{\infty} = \mathscr{B}$. (This result was communicated to the author by D. S. Ornstein.)

We remark in passing that in certain situations one can actually see the "orthogonal complement" to $\mathscr{H}$. For instance if the independent process is given by rolling a six sided die and if we choose $H = \{A, B, C\}$ with $A = \{1, 2\}$, $B = \{3, 4\}$, and $C = \{5, 6\}$, then the factor given by the letter partition has as complement the factor generated by the two set partition which chooses the even or odd number of the clumping. On the other hand, the general result is non-constructive and depends on the relativized isomorphism theory applied to a situation where the non-stationary processes which (T, P) induces on the fibers of $\mathscr{H}$ are independent ones.

This note is an attempt to extend Thouvenot's result on clumping factors of independent generators, both by considering more general finite codings and by considering a more general kind of Markov generator than the independent ones. In one direction, we are able to state that every clumping factor of a finite state Markov process will split off, provided only that all transition probabilities in the process are strictly positive. In the converse direction, an example is given of a clumping factor of a finite state mixing Markov process which fails to split off when certain transitions are impossible.

For an arbitrary factor $\mathcal{F}$ which fails to split off, it is not known whether there exists a factor $\mathcal{G}$ with $\mathcal{F} \subseteq \mathcal{G}$ where $\mathcal{G}$ splits off. However in the example presented, where the finite coding factor fails to split off, we observe the following behavior: each fiber in the coding factor $\mathcal{H}$ contains at most finitely many fibers of another factor $\hat{\mathcal{H}}$. Further, the factor $\hat{\mathcal{H}}$ generates. (We shall say that $\mathcal{H}$ is relatively finite in $\hat{\mathcal{H}}$.) We show that this case is the general one: if a finite coding factor of a Markov generator fails to split off, the structure is quite simple. First, each such coding factor is relatively finite in another factor. Second, in this case we are able to show that the larger factor either generates or itself splits off.

2. The relativized Friedman–Ornstein result

In the following we use $\mu_{\mathcal{H}}$ to indicate a fixed version of conditional probability with respect to the sub-σ-algebra $\mathcal{H}$.

DEFINITION. We say that (T, P) is n-step Markov conditioned on $\mathcal{H}$ if $\mu_{\mathcal{H}}(P \mid P_{-\infty}^{-1}) = \mu_{\mathcal{H}}(P \mid P_{-n}^{-1})$.

DEFINITION. We say that (T, P) is mixing conditioned on $\mathcal{H}$ if given any two finite cylinder sets A, B in the (T, P) process we have that $\lim_{n\to\infty} |\mu_{\mathcal{H}}(T^n A \cap B) - \mu_{\mathcal{H}}(T^n A)\mu_{\mathcal{H}}(B)| = 0$ for all (mod 0) $\mathcal{H}$-fibers.

DEFINITION. We say that two partitions $\mathcal{P}$ and $\mathcal{Q}$ are ε-independent conditioned on $\mathcal{H}$ if $\sum_{A\in\mathcal{P}, B\in\mathcal{Q}} |\mu_{\mathcal{H}}(A \cap B) - \mu_{\mathcal{H}}(A)\mu_{\mathcal{H}}(B)| \leqq \varepsilon$. (Notation: $(\mathcal{P} \perp_{\varepsilon} \mathcal{Q})_{\mathcal{H}}$.)

DEFINITION. We say that (T, P) is weak Bernoulli conditioned on $\mathcal{H}$ if given $\varepsilon > 0$ there exists a positive integer n and a set G with $\mu(G) > 1 - \varepsilon$ of $\mathcal{H}$-fibers for which $(P_{-m}^{0} \perp_{\varepsilon} P_{n}^{n+m})_{\mathcal{H}}$ for all positive integers m.

DEFINITION. Given a finite sequence $\{P_i,\ i = 1, 2, \cdots, n\}$ of partitions on a probability space (Z, ρ) and a finite sequence $\{R_i,\ i = 1, 2, \cdots, n\}$ of partitions from a probability space (Y, σ) we define

$$\bar{d}_n(\{P_i\}_1^n, \{R_i\}_1^n) = \inf_{\phi\in\mathcal{M}} n^{-1} \sum_{i=1}^{n} \rho[(P_i \setminus \phi^{-1}R_i) \cup (\phi^{-1}R_i \setminus P_i)],$$

where $\mathcal{M}$ is the class of all measure preserving transformations from Z to Y and where we sum over corresponding sets in partitions.

DEFINITION. We say that (T, P) is very weak Bernoulli conditioned on $\mathcal{H}$ if given $\varepsilon > 0$ there exists a positive integer n and a set G with $\mu(G) > 1 - \varepsilon$ of $\mathcal{H}$-fibers for which $\bar{d}_n(\{T^iP \mid C\}_1^n, \{T^iP\}_1^n) < \varepsilon$ on a set $\mathcal{C}_h$ of atoms $C \in P_{-m}^0 \cap h$ with $\mu_{\mathcal{H}}(\mathcal{C}_h) > 1 - \varepsilon$. (The $\bar{d}_n$ is measured using $\mu_{\mathcal{H}}$.)

LEMMA 1. *If (T, P) is n-step mixing Markov conditioned on $\mathcal{H}$, then (T, P) is weak Bernoulli conditioned on $\mathcal{H}$.*

PROOF. For arbitrary integer t let $A' \in P_{-m}^{-n}$, $A \in P_{-n+1}^0$, $B \in P_t^{t+n-1}$, $B' \in P_{t+n}^{t+m}$. Then by the Markov property we have that

$$\sum_{A,B,A',B'} |\mu_{\mathcal{H}}(A' \cap A \cap B \cap B') - \mu_{\mathcal{H}}(A' \cap A)\mu_{\mathcal{H}}(B \cap B')|$$

$$\leqq \sum_{A,B,A',B'} \mu_{\mathcal{H}}(A' \mid A)\mu_{\mathcal{H}}(B' \mid B)|\mu_{\mathcal{H}}(A \cap B) - \mu_{\mathcal{H}}(A)\mu_{\mathcal{H}}(B)|$$

$$\leqq \sum_{A,B} |\mu_{\mathcal{H}}(A \cap B) - \mu_{\mathcal{H}}(A)\mu_{\mathcal{H}}(B)|.$$

Since this sum is finite, we may use the assumption of mixing to choose t so large that the sum is less than ε on a set G with $\mu(G) > 1 - \varepsilon$ of $\mathcal{H}$-fibers.

LEMMA 2. *If (T, P) is weak Bernoulli conditioned on $\mathcal{H}$, then (T, P) is very weak Bernoulli conditioned on $\mathcal{H}$.*

PROOF. By the weak Bernoulli property applied for $\eta = \varepsilon^2/16$ there is an n and a set G with $\mu(G) > 1 - \eta$ for which $(P_{-m}^0 \perp_\varepsilon P_n^{n+m})_{\mathcal{H}}$ for all m. We take $m > 4n/\varepsilon$ and get $\bar{d}_m(\{T^iP \mid C\}_1^m, \{T^iP\}_1^m) < \varepsilon$ on a set $\mathcal{C}_h$ of atoms $C \in P_{-m}^0 \cap h$ with $\mu_{\mathcal{H}}(\mathcal{C}_h) > 1 - \varepsilon$ for all $\mathcal{H}$-fibers h in G.

The proof of the following lemma may be found in [5].

LEMMA 3. *If (T, P) is very weak Bernoulli conditioned on $\mathcal{H}$, then (T, P) is finitely determined conditioned on $\mathcal{H}$.*

By Thouvenot's relativized isomorphism theorem [4], if (T, P) is finitely determined conditioned on $\mathcal{H}$ we have that $\mathcal{H}$ splits off with an orthogonal Bernoulli complement. This result, along with the three previous lemmas, constitutes the proof of the following theorem, which is a relativized version of the Friedman–Ornstein result.

THEOREM 1. *If (T, P) is n-step mixing Markov conditioned on $\mathcal{H}$, then $\mathcal{H}$ splits off.*

3. Applications of the relativized Friedman–Ornstein result

In this section we examine the way that certain factors sit in an n-step Markov process. (For more on the concept of how a factor "sits", see [2].) The factors $\mathcal{H}$ that we consider are those generated by finite coding partitions $H \in P_0^{l-1}$. The way that the factors sit will be determined by whether or not they split off in the full process, and also by whether or not they are relatively finite in a larger factor.

We begin by observing that we can reduce the apparently more general case of n-step Markov processes and l length codings to a consideration of 1-step Markov processes and clumpings. For suppose that (T, P) is an n-step Markov process and that $\mathcal{H} = \hat{H}_{-\infty}^{\infty}$, with $\hat{H} \in P_0^{j-1}$ where $j < n$. We can then take $H = \hat{H}_0^{n-j}$ and use the fact that $\mathcal{H} = H_{-\infty}^{\infty}$. In general, suppose that l is the length of the coding where $l \geqq n$. Then since (T, P) is also l-step Markov, we may choose instead to work with the 1-step Markov process (T, P_0^{l-1}). The partition H generating $\mathcal{H}$ is a clumping of P_0^{l-1}.

Next we examine the non-stationary processes on fibers of a clumping factor. Let I denote the set of integers $\{i \mid u \leqq i < v\}$; and J, the set of integers $\{i \mid v < i \leqq w\}$. We get the following conditional independence statement directly from the Markov property: whenever $E \in T^vP$, $F \in \sigma(\bigvee_{i\in I} T^iP)$, $G \in \sigma(\bigvee_{i\in J} T^iP)$ then $(\bigvee_{i\in I} T^iP \perp \bigvee_{i\in J} T^iP)_{E\cap F\cap G}$. This easily proved fact may be restated in the following form.

Lemma 4. *If (T, P) is Markov and $\mathcal{H}$ is a clumping factor, then (T, P) is Markov conditioned on $\mathcal{H}$.*

Next, as in ([3], [4]) we define the conditional Pinsker algebra.

Definition. The Pinsker algebra $\pi(T \mid \mathcal{H})$ of T relative to a factor $\mathcal{H}$ is given by $\pi(T \mid \mathcal{H}) = \bigcap_{n>0}(\bigvee_{-\infty}^{-n} T^iP \vee \mathcal{H}) = \bigcap_{n>0}(\bigvee_{n}^{\infty} T^iP \vee \mathcal{H})$.

One of the properties of the relative Pinsker algebra in this case is:

Lemma 5. *If (P, T) is Markov conditioned on $\mathcal{H}$, then $\pi(T \mid \mathcal{H})$ is measurable with respect to $\mathcal{H} \vee P$.*

Proof. The Markov property of Lemma 4 shows that we have the following conditional independence: $(\bigvee_{-\infty}^{-1} T^iP \vee \mathcal{H} \perp \bigvee_1^{\infty} TP \vee \mathcal{H})_{P\vee\mathcal{H}}$. Hence $(\pi(T \mid \mathcal{H}) \perp \pi(T \mid \mathcal{H}))_{P\vee\mathcal{H}}$. Thus $\pi(T \mid \mathcal{H}) \subseteq P \vee \mathcal{H}$.

Finally, we define one more property.

Definition. We say that a factor $H_{-\infty}^{\infty}$ is maximal in entropy if whenever $H_{-\infty}^{\infty} \subseteq F_{-\infty}^{\infty}$ and $h(H, T) = h(F, T)$ for a partition F then $H_{-\infty}^{\infty} = F_{-\infty}^{\infty}$.

As shown in the appendix to [3], for a factor $\mathscr{H}$ to be maximal in entropy is equivalent to having an $\mathscr{H}$-conditionally trivial Pinsker algebra of T. Moreover, as in [3], either of these conditions implies that the main process is mixing relative to $\mathscr{H}$. Hence we have the following result:

THEOREM 2. *If (T, P) is n-step Markov and $\mathscr{H}$ is a finite coding factor which is maximal in entropy, then $\mathscr{H}$ splits off.*

PROOF. Direct from Lemma 4 and Theorem 1.

As a corollary we get the following generalization of Thouvenot's result on clumping factors of independent processes.

THEOREM 3. *Every clumping factor in a Markov process with strictly positive transition probabilities will split off.*

PROOF. By Lemma 5 we have that $\pi(T \mid \mathscr{H}) \subseteq P \vee \mathscr{H}$. Since $\pi(T \mid \mathscr{H})$ is invariant under T, and since each P state goes to each P state under T, $\pi(T \mid \mathscr{H})$ must have only one atom relative to $\mathscr{H}$.

On the other hand, in the converse direction, observe that even if the original process (T, P) is independent, where we certainly have non-zero transitions, the reduction to the case of (T, P_0^{l-1}) may introduce zero probability transitions. In particular, suppose that we consider a coding of length two on an independent process with $\mu(P_0) = \mu(P_1) = 0.5$. Let the coding output be A if the sum of the last two P-outputs is even; B, if it is odd. The factor generated by this coding partition does not split off since each fiber in $\{A, B\}_{-\infty}^{\infty}$ contains exactly two points of the (T, P) process. Hence the factor is of full entropy, but does not generate. It is therefore not maximal in its entropy class; and by the corollary to theorem 5 in [4], it cannot split off.

What we now show is that the pathology in the preceding example is the only kind possible for a finite coding factor of a Markov generator. If a finite coding factor fails to split off, then it is relatively finite in another factor which either generates or which itself splits off in the main process. (For an example of a very weak Bernoulli generator which does not have this nice property, see [2].)

THEOREM 4. *If (T, P) is n-step Markov and $\mathscr{H}$ is a finite coding factor, then if $\mathscr{H}$ does not split off, it must be relatively finite in a larger factor $\hat{\mathscr{H}}$ which either generates or itself splits off.*

PROOF. If the factor $\mathscr{H}$ is not maximal in entropy, it is easy to see that there exists a unique factor $\hat{\mathscr{H}}$ containing $\mathscr{H}$ and having the same entropy as $\mathscr{H}$ which is maximal in entropy. Lemma 5 and the proof of Lemma 4 now show that the non-stationary process induced by (T, P) on a fiber of $\hat{\mathscr{H}}$ is again non-homogeneous Markov. We apply Theorem 1 to conclude that $\hat{H}$ either generates or itself splits off.

REFERENCES

1. N. A. Friedman and D. S. Ornstein, *On isomorphism of weak Bernoulli transformations*, Advances in Math. **5** (1970), 365–394.

2. D. S. Ornstein, *Factors of Bernoulli shifts*, Israel J. Math. **21** (1975), 145–153.

3. M. Rahe, *Relatively finitely determined implies relatively very weak Bernoulli*, Canad. J. Math. **30** (1978), 531–548.

4. J.-P. Thouvenot, *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schèma de Bernoulli*, Israel J. Math. **21** (1975), 177–207.

5. J.-P. Thouvenot, *Remarques sur les systèmes dynamiques donnes avec plusiers facteurs*, Israel J. Math. **21** (1975), 215–232.

DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843 USA