A mixing Markov chain with exponentially decaying return times is finitarily Bernoulli

DANIEL J. RUDOLPH

Department of Mathematics, University of Maryland, College Park, Maryland 20742

Abstract. We here show that the necessary condition of exponentially decaying first return times to a cylinder set is also sufficient for a countable state Markov chain to be almost continuously isomorphic to a Bernoulli shift.

In [5] Smorodinsky and Keane showed that a finite state mixing (non-cyclic) Markov chain is finitarily Bernoulli, i.e. conjugate to a finite state Bernoulli shift by a conjugating map continuous almost everywhere in the topologies generated by cylinder sets of the two processes. If the Markov chain has countably many states, even assuming it has an invariant probability measure (is 'ergodic' in the sense of [1]) and is mixing, this need not be the case, as Smorodinsky has shown that to be finitarily Bernoulli, the return time probabilities to a state must decay at least exponentially. This is in contrast to the fact that such a Markov chain will be measurably isomorphic to a Bernoulli shift.

What we will demonstrate here is that if the chain has finite entropy such exponentially decaying return times are also a sufficient condition for the chain to be finitarily Bernoulli. We do this by verifying that the chain satisfies the 'finitarily Bernoulli' condition of [3] which is there shown to be equivalent to the existence of such an almost continuous conjugacy. The question of infinite entropy remains open, although the work of B. Petit [2] showing that all infinite entropy Bernoulli shifts are almost continuously conjugate, and our work here, indicates that they should also be finitarily Bernoulli.

We begin our discussion by stating the finitarily Bernoulli condition which must be verified. The reader is referred to [3] for notation and further discussion. Let (T, P) be a countable state stationary stochastic process on the Lebesgue probability space $(\Omega, \mathcal{F}, \mu)$, with the process topology generated by cylinder sets.

We say (T, P) is finitarily Bernoulli (f.B.) if the following structures with the given properties exist.

For $\varepsilon > 0$ there is a countable partition

$$A(\varepsilon) = \{A_i(\varepsilon), \bar{A}_i(\varepsilon)\}$$

of Ω into almost open, hence almost closed, sets with

$$E(\varepsilon) = \bigcup_i A_i(\varepsilon)$$

nesting down in ε to a null set. There also is, for each i, a subset $\hat{A}_i(\varepsilon)$ with

$$\bar{A}_i(\varepsilon) \subseteq \hat{A}_i(\varepsilon) \subseteq A(\varepsilon).$$

For any $\omega \in \Omega$ and $n \in \mathbb{N}$ we define

$$t_0(\omega, \varepsilon, n) = \{\omega' | i \in \{0, \ldots, n-1\}, T^{-i}(\omega) \text{ and } T^{-i}(\omega') \text{ are both in the same } A_j(\varepsilon),$$
$$\text{or } P_i(\omega') = P_i(\omega) \text{ and } T^{-i}(\omega'), T^{-i}(\omega) \in \hat{A}_j(\varepsilon)\}.$$

This 't_0-thickening' is a restricted $\bar{d}$-neighbourhood of the P, n-name of ω, the $\bar{d}$-errors required to lie in the $A_i(\varepsilon)$. We define for $E \in \mathcal{F}$,

$$t_0(E, \varepsilon, n) = \bigcup_E t_0(\omega, \varepsilon, n),$$

an almost clopen set.

We also have, for each $\varepsilon > 0$, $n \in \mathbb{N}$, a 'nice partition', $C(\varepsilon, n)$ of Ω into almost clopen sets where $C(\varepsilon', n')$ refines $C(\varepsilon, n)$ if $\varepsilon' \leq \varepsilon$ and $n' \geq n$. This must be a partition into small sets in that for any atom $c \in C(\varepsilon, n)$, for any $\omega \in c$, $c \subset t_0(\omega, \varepsilon, n)$. We define $t_1(\omega, \varepsilon, n)$ to be that $c \in C(\varepsilon, n)$ with $\omega \in C$.

The t_0-thickening must also be small in an asymptotic sense in that for any $\varepsilon_0 > 0$ there are $\alpha_1(\varepsilon_0), \alpha_2(\varepsilon_0), \ldots$ so that if $\varepsilon_i \leq \alpha_i(\varepsilon_0)$ and we define $E_1 = t_0(\omega, \varepsilon_1, n)$, and inductively $E_{i+1} = t_0(E_i, \varepsilon_i, n)$, then for all $i, E_i \subset t_0(\omega, \varepsilon_0, n)$, i.e. if we t_0-thicken slowly enough we always remain inside a t_0, ε_0-thickening.

Next, for each $\varepsilon > 0$ we have $N(\varepsilon) \nearrow \infty$ and almost clopen 'spacer' sets $S(\varepsilon)$ with $\mu(S(\varepsilon)) > 0$, satisfying:

(i) for all $\varepsilon', n, m, t_1(T^n(S(\varepsilon)), \varepsilon', n + N(\varepsilon) + m) = T^n(S(\varepsilon))$, i.e. t_1-thickening over any block of indices containing a spacer does not increase the spacer,

(ii) there is an N_0 so that for any ε, if ε' is small enough, then for any n with

$$-N(\varepsilon') + N_0 \leq n \leq N(\varepsilon) - N_0,$$
$$T^n(S(\varepsilon')) \cap S(\varepsilon) = \varnothing,$$

i.e. spacers can overlap by at most N_0-digits,

(iii) for any ε_0, there is an $\bar{N}(\varepsilon_0)$ so that for any $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{k_1}, \ldots, \varepsilon_k$ and $\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_{k_1'}', \ldots, \varepsilon_{k'}'$, if $n_i < -N(\varepsilon_i) - \bar{N}(\varepsilon_0)$ and $m_i > 0$ then

$$\mu\Big(\bigcap_{i=1}^{k_1} T^{n_i}(S(\varepsilon_i)) \bigcap_{i=k_1+1}^{k} T^{n_i}(S(\varepsilon_i)^c) \cap \bigcap_{i=1}^{k_1'} T^{m_i}(S(\varepsilon_i')) \bigcap_{i=k_1'+1}^{k'} T^{m_i}(S(\varepsilon_i')^c)\Big)$$

$$= \mu\Big(\bigcap_{i=1}^{k} T^{n_i}(S(\varepsilon_i)) \bigcap_{i=k_1+1}^{k} T^{n_i}(S(\varepsilon_i)^c)\Big) \cdot \mu\Big(\bigcap_{i=1}^{k_1'} T^{m_i}(S(\varepsilon_i')) \bigcap_{i=k_1'+1}^{k'} T^{m_i}(S(\varepsilon_i')^c)\Big) 2^{\pm \varepsilon_0},$$

i.e. there is a uniform rate of mixing on cylinder sets in the spacer process (this 'USM' condition is precisely where the exponential decay of return times in our Markov process is needed).

As in [3], for ease of expression, we introduce the objects (S, n), $S \in \mathcal{F}$, $n \in Z^+$ with the product rule

$$(S, n) \circ (S', n') = (S \cap T^n(S'), n + n').$$

When we refer to (S, n) as a point set, we are referring to S.

(iv) there is a $\sigma(\varepsilon) > 0$ so that for any $\varepsilon' \leq \varepsilon$, if $c' \subset C(\varepsilon, n')$ with $c' \subset T^{n'-N(\varepsilon'')}(S(\varepsilon''))$, $c' \subset c \subset C(\varepsilon, n)$, $n \leq n'$ with $c \subset T^{n-N(\varepsilon''')}(S(\varepsilon'''))$ then if $\mu((S(\varepsilon'), N(\varepsilon'))(c', n')) \neq 0$, we must have

$$\mu((S(\varepsilon'), N(\varepsilon'))(c', n') | (S(\varepsilon'), N(\varepsilon'))(c, n)) > \sigma(\varepsilon)^{n'-n},$$

i.e. the sizes of sets in $C(\varepsilon, n)$ which end in a spacer decrease in n with at most an exponential rate, the bound depending on ε,

(v) for any ε, $i = 1, \ldots, k_1, \ldots, k$, $\varepsilon_i \geq \varepsilon$, $0 \leq j(i) \leq J$,

$$\mu \left(\bigcap_{i=1}^{k_1} T^{j(i)}(S(\varepsilon_i)) \bigcap_{i=k_1+1}^{k} T^{j(i)}(S(\varepsilon_i)^c) \right)$$

is either 0 or $> \sigma(\varepsilon)^J$, i.e. spacer cylinders for $\varepsilon_i \geq \varepsilon$ which have positive mass have sizes decaying at most exponentially.

Thickened blocks separated by spacers now must satisfy the following 'conditional block independence' (CBI). For any $\varepsilon_1, \ldots, \varepsilon_k$, $n_1, \ldots, n_k$, sets $E_1, \ldots, E_k$ and $\bar{\varepsilon}_i \geq \max(\varepsilon_i, \varepsilon_{i+1})$, setting $\bar{E}_i = (t_1(E_i, \bar{\varepsilon}_i, n_i), n_i)$ and $\bar{S}_i = (S(\varepsilon_i), N(\varepsilon_i))$, then

$$\mu(\bar{E}_1 \circ \bar{S}_1 \circ \bar{E}_2 \circ \bar{S}_2 \circ \cdots \circ \bar{S}_{k-1} \circ \bar{E}_k \circ \bar{S}_k | \bar{E}_1 \circ \bar{S}_1 \circ \cdots \circ \bar{E}_{k-1} \circ \bar{S}_{k-1})$$
$$= \mu(\bar{S}_{k-1} \circ \bar{E}_k \circ \bar{S}_k | \bar{S}_{k-1}) 2^{\pm \varepsilon_{k-1}}.$$

See [3] for a proof of the following result.

THEOREM 1. *If (T, P) is a countable state stationary process, $h(P) < \infty$, then (T, P), with its process topology, is almost continuously conjugate to a finite state Bernoulli shift iff (T, P) is finitarily Bernoulli.*

Let (T, P) be a countable state mixing Markov process,

$$p_{i,j} = \mu(T^{-1}(\omega) \in P_j | \omega \in P_i) = p_{i,j}^{(1)},$$
$$p_{i,j}^{(k)} = \mu(T^{-k}(\omega) \in P_j | \omega \in P_i).$$

Setting $\Pi = (p_{i,j}^{(1)})$, the Markov condition now says

$$(p_{i,j}^{(k)}) = \Pi^k.$$

That T is a mixing process is equivalent to
(i) T is recurrent, i.e. for all j there is a k with

$$p_{1,j}^{(k)} > 0 \text{ and}$$

(ii) T is non-cyclic, i.e.

$$\text{GCD}(\{k | p_{1,1}^{(k)} > 0\}) = 1.$$

A mixing Markov process with invariant vector

$$\vec{p} = \begin{pmatrix} \mu(1) \\ \mu(2) \\ \vdots \end{pmatrix}$$

is completely determined by its transition matrix Π as the following compilation of results from [1] shows.

LEMMA. 2. (a) *If Π is the matrix of a recurrent chain, then the sequence of powers of P is Cesaro summable to a limiting matrix L with $L \geq 0$,*

$$L\Pi = L = \Pi L = L^2$$

[1: Theorem 6-1, p. 130].

Further one says Π is a 'null chain' if $L = 0$. If $L \neq 0$, Π is said to be an 'ergodic chain' and the matrix $L = 1\vec{\alpha}$ where $\vec{\alpha}$ is an invariant probability vector. If Π has an invariant probability vector it must be an ergodic chain.

(b) If Π is non-cyclic and recurrent, then $\lim_{n \to \infty} \Pi^n$ exists. If Π is ergodic, $\lim \Pi^n = 1\vec{p}$ and $\lim_n \|\vec{\alpha}\Pi^n - \vec{p}\| = 0$ for every probability vector $\vec{\alpha}$ [1: Theorem 6-38, p. 153].

COROLLARY 3. *If $a_1, a_2, \ldots \geq 0$, $\sum_i a_i = 1$, $\sum_i i a_i < \infty$ and*

$$GCD (\{i | a_i > 0\}) = 1,$$

and if we set $x_0 = 1$, $x_{-1}, x_{-2}, \ldots = 0$, and for $i > 0$,

$$x_i = \sum_{j=1}^{\infty} a_j x_{i-j}$$

then $x_i \xrightarrow{i} \bar{x} = (\sum_i i a_i)^{-1} > 0$.

Proof. Let $p_{1,j} = a_j$ and $p_{i,j} = \delta_{j,i-1}$ if $i \neq 1$. As $p_{1,1}^{(k)} \geq a_k$,

$$GCD (\{k | p_{1,1}^{(k)} > 0\}) = 1.$$

If only finitely many $a_j > 0$, then let a_{J} be the largest, and otherwise $J = \infty$.

Thus for $j < J + 1$, there is a k with $p_{1,j}^{(k)} > 0$. Notice $(\Pi(\vec{y}))_i = y_{i+1} + a_i y_0$ and so if

$$\vec{y} = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}$$

then $x_i = (\Pi^i(\vec{y}))_0$. The stationary probability vector for Π is

$$p_i = \sum_{j=i}^{\infty} a_j \Big/ \sum_{i=1}^{\infty} i a_i$$

hence by lemma 2

$$x_i \xrightarrow{i} p_o = \bar{x}. \qquad \square$$

We are now ready to discuss exponentially decaying return times. For any set S let

$$a_i(S) = \mu(\{T(\omega), \ldots, T^{i-1}(\omega) \notin S \text{ but } T^i(\omega) \in S | \omega \in S\})$$

be the probability that the first return of S to itself is after i steps. Clearly

$$\sum_{i=1}^{\infty} i a_i(S) = \frac{1}{\mu(S)}.$$

We say S has 'exponentially decaying return time' (e.d.r.t.) if for all i

$$a_i(S) \le ac^i \quad \text{for some } a \ge 0, \qquad 0 \le c < 1.$$

THEOREM (Smorodinsky). *If (T, P) is finitarily Bernoulli (not necessarily Markov), and S is an almost open set, then S has e.d.r.t.*

Proof. It is enough to show

$$b_i = \mu(\{T(\omega), \ldots, T^{i-1}(\omega) \notin S\}) < c' a'^i \quad \text{as } a_i(S) \le \frac{b_i}{\mu(S)}.$$

As (T, P) is finitarily Bernoulli there is a finite partition B of Ω with (T, B) a Bernoulli shift where S is a.s. a union of T, B-cylinders. Let $b \in B_{-N}^{N}, b \subset S$.
Now

$$b_i \le \mu(\{T(\omega), \ldots, T^{i-1}(\omega) \notin b\})$$

$$\le \mu(\{T(\omega), T^{2N+1}(\omega), \ldots, T^{2Nk+1}(\omega) \notin b\}), k = \left[\frac{i-1}{2N}\right].$$

But this is

$$(1 - \mu(b))^{k+1} \le \frac{((1-\mu(b))^{1/2N})^{i-1}}{1-\mu(b)} = c' a'^i. \qquad \square$$

Smorodinsky has used this fact to construct countable state mixing Markov shifts not finitarily Bernoulli by selecting $a_i(1) > 0$ with $\sum i a_i(1) < \infty$ but $a_i(1)$ not decaying exponentially.

What we will now show is that if a single state (here we will use 1) has e.d.r.t., then (T, P) is finitarily Bernoulli.

Our first step is to find a certain nice set.

We say a set S decays calmly if for every $\varepsilon > 0$ there is an N so that for all k

$$\varepsilon \sum_{j=k}^{k+N-1} a_j(S) > \sum_{j=k+N}^{\infty} a_j(S).$$

It is easy to see that a set decaying calmly must have e.d.r.t.

We now must show that such a set exists.

THEOREM 4. *If (T, P) is a countable state mixing Markov process and state 1 has e.d.r.t., then for some t, the set*

$$A_t = 1 \cap T^t(1) \bigcap_{k=1}^{t-1} T^k(1^c)$$

decays calmly.

Proof. If only finitely many $a_i(1)$ are positive then $A_{-1} = 1$ decays calmly, so we may assume infinitely many $a_i(1) > 0$. We know $\sum_i a_i(1) = 1$ and $a_i(1) < Ca^i$ for $0 \le a < 1$, so for any $1 < z < a^{-1}$,

$$\sum_{i=1}^{\infty} a_i(1) z^i < \infty.$$

Select $a_t(1) > 0$ so small that for some such $z_0 > 1$,

$$\sum_{i \neq t} a_i(1) z_0^i > 1.$$

Hence for some $z < z_0$, $\sum_{i \neq t} a_i(1) z^i = 1$, and as $z < a^{-1}$

$$\sum_{i \neq t} i a_i(1) z^i < \infty.$$

Be sure t is so large that

$$\text{GCD}\ (\{i \,|\, i \neq t,\, a_i(1) > 0\}) = 1.$$

Letting .

$$a_i = \begin{cases} a_i(1) z^i & \text{if } i \neq t \\ 0 & \text{if } i = t, \end{cases}$$

$x_0 = 1$, $x_{-k} = 0$,

$$x_i = \sum_{j=1}^{\infty} a_j x_{i-j},$$

by corollary 3, $x_i \overset{i}{\to} \bar{x} > 0$.

Thus if we define $y_0 = 1$, $y_{-k} = 0$, $y_i = \sum_{j \neq t} a_j(1) y_{i-j}$ one easily checks that $y_i = x_i / z^i$, hence $z^i y_i \overset{i}{\to} \bar{x}$ and asymptotically y_i is $\bar{x}/z^i$. Letting $S = A_t$ it is easy to check that for $i > 0$

$$y_i = \mu(\{T(\omega), \ldots, T^{i+t-1}(\omega) \notin S \,|\, \omega \in S\}),$$

so for $i \geq t$

$$a_i(S) = y_{i-t} - y_{i-t+1}.$$

For any $\varepsilon > 0$, once i is sufficiently large we know

$$y_i = \frac{\bar{x}(1 \pm \varepsilon)}{z^i},$$

hence

$$a_i(S) = \frac{\bar{x}}{z^{i-t}}\left(1 - \frac{1}{z}\right)(1 \pm \bar{\varepsilon})$$

where $\bar{\varepsilon} = \varepsilon/(1 - 1/z)$.

Thus for $\bar{a} = \bar{x} z^t (1 - 1/z)$, for any ε, once i is sufficiently large,

$$a_i(S) = \frac{\bar{a}}{z^i}(1 \pm \varepsilon).$$

Thus for $\bar{\varepsilon} = \tfrac{1}{2}$, there is a K so that for $k \geq K$, for any N,

$$\sum_{i=k}^{k+N-1} a_i(S) > \frac{(1 - \bar{\varepsilon})\bar{a}}{z^k}\left(\frac{1 - 1/z^N}{1 - 1/z}\right)$$

and

$$\sum_{i=k+N}^{\infty} a_i(S) < \frac{(1 + \bar{\varepsilon})\bar{a}}{z^{k+N}}\left(\frac{1}{1 - 1/z}\right),$$

and hence

$$\frac{\sum_{j=k+N}^{\infty} a_j(S)}{\sum_{j=k}^{k+N-1} a_j(S)} < \frac{3}{z^N - 1}.$$

Thus for any ε, choose N_0 so large that $3/(z^{N_0} - 1) < \varepsilon$. Letting $N = K + N_0$, we easily check that S decays calmly. $\square$

Now to see what we gain from such a set. Let $S = A_t$ be this calmly decaying set. Let Q be the partition of Ω into almost open sets defined by $\omega \in \bar{j}$ if $j - 1$ is the smallest integer ≥ 0 with $T^j(\omega) \in S$. Thus $S = \bar{1}$ and one computes

$$\mu(\bar{j}) = \mu(S) \sum_{i=j}^{\infty} a_i(S).$$

As S is of the form A_t, (T, Q) is a mixing Markov process with transition matrix $\bar{\Pi}$ given by

$$\bar{p}_{\bar{1},\bar{j}}^{(1)} = a_j(S)$$

and for $i \neq 1$

$$\bar{p}_{\bar{i},\bar{j}} = \delta_{i-1,j}.$$

Let $C_m = T^{-m}(\bar{1}) \bigcap_0^{j=m-1} T^{-j}(\bar{1}^c)$, not cylinders of (T, Q) but of the process $(T, \{\bar{1}, \bar{1}^c\})$. The following result gives us the needed uniform mixing for spacers.

LEMMA 5. *For any $\varepsilon > 0$, there is an N so that for all $n \geq N$ and m and $\bar{j}$,*

$$\mu(C_m \cap T^n(\bar{j})) = \mu(C_m)\mu(\bar{j})2^{\pm \varepsilon}.$$

Proof. Fix m and let

$$x_j^{(i)} = \mu(T^i(\bar{j})|C_m), \quad \vec{x}^{(i)} = \begin{pmatrix} x_0^{(i)} \\ \vdots \\ x_j^{(i)} \\ \vdots \end{pmatrix}.$$

For $0 \leq i \leq m$, $a_i = a_i(S)$

$$x_j^{(i)} = \begin{cases} 0 & \text{if } i + j - 1 \leq m \\ a_{i+j} \Big/ \sum_{k=m}^{\infty} a_k = C a_{i+j} & \text{if } i + j - 1 > m, \end{cases}$$

and for $i > m$

$$\vec{x}^{(i)} = \bar{\Pi}^{i-m}(\vec{x}^{(m)}).$$

From the form of $\bar{\Pi}$, for $i > m$ we can write

$$\vec{x}^{(i)} = \bar{\Pi}^{i-m-1}\begin{pmatrix} C a_{m+1} \\ 0 \\ \vdots \end{pmatrix} + \bar{\Pi}^{i-m-2}\begin{pmatrix} C a_{m+2} \\ 0 \\ \vdots \end{pmatrix} + \cdots + \bar{\Pi}\begin{pmatrix} C a_{i-1} \\ 0 \\ \vdots \end{pmatrix} + \begin{pmatrix} C a_i \\ 0 \\ \vdots \end{pmatrix}.$$

As (T, Q) is mixing,

$$\bar{\Pi}^N\begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix} \xrightarrow{N} \vec{x} = \begin{pmatrix} \mu(\bar{1}) \\ \mu(\bar{2}) \\ \vdots \end{pmatrix}.$$

For $\bar{\varepsilon} < \frac{1}{3}$, choose N_0 so large that for all m,

$$\sum_{k=N_0+m+1}^{\infty} Ca_k < \frac{\bar{\varepsilon}\mu(\bar{1})}{2} \sum_{k=m+1}^{N_0+m} a_k$$

and for $i > N_0$,

$$\left(\bar{\Pi}^i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) = \mu(\bar{1})(1 \pm \bar{\varepsilon}),$$

and now if $i > N_0 + m + 1$,

$$x_1^{(i)} = \mu(\bar{1})(1 \pm \bar{\varepsilon}) \sum_{k=m+1}^{N_0+m} Ca_k \pm \sum_{k=N_0+1+m}^{\infty} Ca_k$$

$$= \mu(\bar{1})(1 \pm z\bar{\varepsilon}),$$

where N_0 does not depend on m.

This verifies

$$\mu(C_m \cap T^n(\bar{1})) \to \mu(C_m)\mu(\bar{1})$$

uniformly in m.

Now for $j \neq 0$, $i \geq m+1$

$$x_j^{(i)} = \sum_{k=m+1}^{i} x_1^{(k)} a_{j+i-k} = \sum_{k=j}^{j+i-m-1} x_1^{(j+i-k)} a_k.$$

Choose N_0 so large that for $n \geq N_0$, for all j

$$\frac{\mu(\bar{1})\bar{\varepsilon}}{2} \sum_{k=j+1}^{j+N_0} a_k > \sum_{k=j+N_0+1}^{\infty} a_k$$

and make sure i is so large that for $n \geq i - N_0$,

$$x_1^{(n)} = \mu(\bar{1})(1 + \bar{\varepsilon})$$

and now

$$x_j^{(i)} = \sum_{k=j}^{j+N_0} x_1^{j+i-k} a_k + \sum_{k=j+N_0+1}^{j+i-m-1} x_1^{(j+i-k)} a_k$$

$$= \left(\sum_{k=j+1}^{j+N_0} a_k \right) \mu(\bar{1})(1 \pm \bar{\varepsilon}) \pm \sum_{k=j+N_0+1}^{\infty} a_k$$

$$= \mu(\bar{1})(1 \pm 2\bar{\varepsilon}) \sum_{k=j}^{\infty} a_k = \mu(\bar{j})(1 \pm 2\bar{\varepsilon})$$

and we have the result for $C_m \cap T^n(\bar{j})$. $\square$

COROLLARY 6. *For any $\varepsilon > 0$ there is an N so that for any set*

$$A \subset \bigvee_{i=0}^{-\infty} T^i(\{\bar{1}, \bar{1}^c\}),$$

and any $\bar{j}, n \geq N$

$$\mu(A \cap T^n(\bar{j})) = \mu(A)\mu(\bar{j})2^{\pm \varepsilon}.$$

Proof. Write $A = \bigcup_m (C_m \cap A_m)$ where $A_m \subset \bigvee_{i=-m-1}^{-\infty} T^i(\{\bar{1}, \bar{1}^c\})$, a disjoint union, and now by the Markov property

$$\mu(A \cap T^n(\bar{j})) = \sum_m \mu(A_m | T^{-m}(\bar{1})) \cdot \mu(C_m \cap T^n(\bar{j})).$$

The result follows from the uniform convergence of lemma 5. $\square$

COROLLARY 7. *For any $\varepsilon > 0$ there is an N so that for any*

$$A \subset \bigvee_{i=0}^{-\infty} T^i(\{\bar{1}, \bar{1}^c\})$$

and

$$B \subset \bigvee_{i=1}^{\infty} T^i(\{\bar{1}, \bar{1}^c\}),$$

$$\mu(A \cap T^n(B)) = \mu(A)\mu(B)2^{\pm\varepsilon}.$$

Proof. As in the previous corollary, the result is uniformly true for all sets $B \cap \{\bar{j}\}$, hence for their union. $\square$

We are now prepared to prove our main result.

THEOREM 8. *If (T, P) is a countable state mixing Markov process and $1 \in P$ has e.d.r.t., then (T, P) is finitarily Bernoulli.*

Proof. What we must do is demonstrate that all the necessary structures of the definition exist in (T, P).

Let $S = A_t$ be the calmly decaying set of theorem 4. Partition Ω into almost clopen sets $F_{i,j} = \{\omega | \omega \in T^{-i}(S), \omega \in T^{j-i}(S),$ but for no $0 < i' < i$ is $\omega \in T^{i-i'}(S)$ where $0 \le j, 0 < i\}$.

Thus $F_{i,j}$ consists of those points which lie between two consecutive occurrences of S, j indices apart in their orbits and which, in this block between occurrences, are i indices from the left end.

Let $k_1, \ldots, k_t$ be a finite set of integers ≥ 1, not all 1 with

$$\text{GCD}(\{k_1, \ldots, k_t\}) = 1$$

and $a_{k_i} > 0$. Let

$$S_i = \left(S \bigcap_{j=1}^{k_i-1} T^j(S^c), k_i \right).$$

Thus any set of the form $S_{i(1)} \circ S_{i(2)} \circ \cdots \circ S_{i(p)} \circ (S, 1)$ has positive probability $\mu(S) \cdot a_{k_{i(1)}} \cdot a_{k_{i(2)}} \cdot \cdots \cdot a_{k_{i(p)}}$.

A spacer will have the form

$$S(j) = S_1 \circ S_2 \circ S_1^{j+2} \circ S_2 \circ S_1 \circ (S, 1),$$

and

$$N(j) = k_1(j+4) + 2k_2 + t + 1,$$

is the 'length' of this set, as it is a union of cylinders in $P_0^{N(j)-1}$.

To define $A(j)$ and $C(n, j)$ we work as follows. There is an $\tilde{N}$ so large that for $n \geq \tilde{N}$,

$$n = k_1 + 6k_2 + \sum_{i=1}^{t} a_i k_i + 1$$

where $a_i > 1$ are integers. Let

$$G_n = S_1 \circ S_2 \circ S_2 \circ S_2 \circ \prod_{i=1}^{t} S_i^{a_i} \circ S_2 \circ S_2 \circ S_2 \circ (S, 1)$$

which is a union of cylinders in P_0^{n+t-1} which, when expressed as a single name in $\{S, S^c\}$ begins with an S and ends with an S. Notice that

$$\mu(G_n) = \mu(S) a_{k_1} a_{k_2}^6 \prod_{i=1}^{t} a_{k_i}^{a_i} > \bar{\sigma}^n \quad \text{for some } \bar{\sigma} > 0.$$

We will use such a cylinder to 'thicken' pieces of a name which are too unlikely to give us the bound of condition (iv) (p. 87).

From their form we easily check that if $T^s(G_m) \cap G_n \neq \varnothing$ then either $s = 0$ and $n = m$ or $s \geq n - 1$. Thus all sets $T^{-s}(G_m)$ for all m and $0 \leq s < m - 1$ are disjoint. Also if

$$T^s(S(j)) \cap G_n \neq \varnothing$$

then $s \leq -N(j) + t + 1$ or $s \geq n - 1$ and if $T^s(S(j)) \cap S(j') \neq \varnothing$ then either $s = 0$ and $j = j'$ or $s \leq -N(j') + k + t$ or $j \geq N(j) - k - t$.

Returning to the proof, let

$$A(j) = \{\bar{A}_i(j), A_{i',j'}(j), A_0(j)\}$$

($A(j)$ need only be a countable partition, it does not matter how we index it), where

$$\bar{A}_i(j) = \begin{cases} i \cap \left(\bigcup_{j' > j+k} F_{i',j'} \cup \bigcup_{\substack{j' > j+k \\ i' < j'}} T^{-i'}(G_{j'}) \right)^c & \text{if } i \leq j \\ \\ \varnothing & \text{if } i > j \end{cases}$$

$$A_{i',j'}(j) = \begin{cases} F_{i',j'} \cup T^{-i'}(G_{j'}) & \text{if } j' > j+k, \, i' < j' - 1 \\ \varnothing & \text{if } j' \leq j+k, \, i' < j - 1, \end{cases}$$

and

$$A_0(j) = \bigcup_{\substack{j' \leq j+k \\ i'}} \left(F_{i',j'} \cap \bigcup_{i>i} i \right),$$

and $k = \max(k_1, \ldots, k_t)$.

The $\bar{A}_i(j)$ are clearly disjoint and disjoint from the $A_{i',j'}(j)$ and $A_0(j)$, and $A_0(j)$ is clearly disjoint from the $A_{i',j'}(j)$. The $A_{i',j'}(j)$ are disjoint as the $T^{-i'}(G_j)$ for $i' < j' - 1$ are. Thus $A(j)$ is a partition.

The $A_{i',j'}(j)$ are clearly almost clopen and the $\bar{A}_i(j)$ are clearly almost closed. That they are almost open follows from the fact that

$$\left(\bigcup_{\substack{j' > j+k \\ i'}} F_{i',j'} \right)^c = \bigcup_{\substack{j' < j+k \\ i'}} F_{i',j'}$$

is open and the set $(S_1 \circ S_2 \circ S_2 \circ S_1)$ lies in no set of the form $T^{-i'}(G_{j'})$, $i' < j'$. Hence reading the P-name of a point ω to the left and right to the first occurrences of this set, from only this finite piece of name we can decide whether ω is in $\bigcup_{i' > i+k} T^{-i'}(G_{j'})$ or its compliment, and so both are almost open, as now is $\bar{A}_i(j)$.

Clearly $\bigcup_{i',j'} A_{i',j'}(j) \cup A_0(j) \searrow \varnothing$, as $\bigcup_{i',j'} F_{i',j'} = \Omega$ and $T^{-i'}(G_{j'})$ are disjoint.

We set $\hat{A}_i(j) = \bar{A}_i(j)$. This now defines $t_0(\omega, j, n)$. It can be described as follows. Read the P-name to the left of ω to the first occurrence of S, and to the right to the first occurrence of S beyond index n. This name now breaks into blocks between consecutive S's. If a string of such blocks forms the name of some G_n, lump them together as such a G_n. We know two such cannot overlap. We now have broken the name into blocks $I_1, I_2, \ldots, I_k$ which are either between two consecutive S's or are a copy of some G_n. To build $t_0(\omega, j, n)$, we allow any block I_i of length more than $j + k$ to be replaced by any other such block, including the appropriate G_n. Given any block of length $\leq j + k$, we allow the P-name across this block to be modified a little, allowing any symbol $> j$ to be replaced by any other $> j$. Lastly, the name preceding 0 or after n can be modified in any way as long as the first and last occurrences of S remain the first before 0 and last after n. This constructs a collection of P-names the union of which is $t_0(\omega, j, n)$.

The following lemma is easily seen from this discussion.

LEMMA 9. *If $t_0(\omega, j, n) \cap t_0(\omega', j, n) \neq \varnothing$ then $t_0(\omega, j, n) = t_0(\omega', j, n)$.*

As $\hat{A}_i(j) = \bar{A}_i(j)$ we know $t_0(\omega, j, n)$ is that element of $A(j)_0^{n-1}$ containing ω. Let $C(j, n) = A(j)_0^{n-1}$, and as $A(j+1)$ refines $A(j)$ the needed conditions on $C(j, n)$ follow.

We have already defined $S(j)$ and $N(j)$ so what remains undemonstrated are conditions (i) through (v) and CBI.

Condition (i) follows as any $S(j) \subset \{1, 1^c\}_0^{N(j)-1}$ is made of blocks between consecutive S's at most k apart and cannot intersect any $T^{-i'}(G_{j'})$, $0 \leq i' \leq j' - 1$.

We have already verified (ii) with $N_0 = k + t$.

A set of the form

$$\bigcap_{i=1}^{k_1} T^{n_i}(S(j_i)) \bigcap_{i=k_1+1}^{k} T^{n_i}(S(j_i)^c)$$

where $n_i < -N(j_i)$ is in $\bigvee_{i=0}^{-\infty} T^i(\{S, S\})$ and if $m_i > 0$,

$$\bigcap_{i=1}^{k_1} T^{m_i}(S(j_i')) \bigcap_{i=k_1+1}^{k'} T^{m_i}(S(j_i')^c)$$

is in $\bigvee_{i=1}^{\infty} T^i(\{S, S^c\})$ and so condition (iii) follows from corollary 7.

For (iv), if $c' \in C(j, n')$ ends in a spacer $S(j'')$ and

$$c' \subset c \in C(j, n), \quad n \leq n',$$

ends in a spacer $S(j''')$, then if

$$\omega \in (S(j'), N(j'))(c', n'),$$

the P-name of ω from 0 to $N(j') + n'$ begins and ends in an occurrence of S. Thus

when we thicken ω across this piece, getting of course

$$(S(j'), N(j'))(c', n'),$$

we need not read either to the left or right of this section. Let $I_1, .., I_p$ be the blocks into which we break the name. If the length l_i of I_i is $< j+k$, then the thickened block across this section can be one of only a finite collection of sets, as it is given by a name of length l_i in the symbols

$$\{1, \ldots, j, A_0(j)\}.$$

Each such name defines a set, those points which have that name from 0 to l_i. Choose σ_1 so small that for any set

$$A \in \{1, \ldots, j, A_0(j)\}^l, \quad l \le j+k$$

of positive probability, $\mu(A) > \sigma_1^l$.

For any I_i with length $l_i > j+k$, this block in $(S(j'), N(j'))(c', n')$ contains G_{l_i}, and we know $\mu(G_{l_i}) > \bar{\sigma}^{l_i}$, between S's with a G_n is needed, and shows why, even though the $a_i(S)$ must decay at least exponentially, there is no harm in their decaying faster.

For condition (v), as any set of the form

$$\bigcap_{i=1}^{k_1} T^{j(i)}(S(j_i)) \bigcap_{i=k_1+1}^{k} T^{j(i)}(S(j_i)^c),$$

$0 \le j(i) \le J, j_i \le j$, is a union of cylinders in $\{S, S^c\}_0^{J+N(j)}$, all we need verify is that such cylinders have sizes decaying at most exponentially fast in their length. By the Markov property on (T, Q), it is enough to note the fact about $S = A_t$ from theorem 4 that either once k is sufficiently large $a_k(S) = 0$, or $a_k(S) > \bar{\sigma}^k$ for some $\bar{\sigma} > 0$. This now forces the size of a cylinder of length n to be 0 or at least

$$\left(\frac{\bar{\sigma}}{\mu(S)}\right)^n = \sigma^n.$$

The proof of theorem 8 is now finished by noting that CBI follows without an error term from the Markov property. $\square$

As one application of this result we see that β-automorphisms, which were shown in [4] to be measurably isomorphic to Bernoulli shifts, as they have c.d.r.t., are finitarily isomorphic.

This now settles the question of when a mixing Markov shift is almost continuously isomorphic to a Bernoulli shift.

Interesting questions, though, present themselves concerning a possibly cyclic Markov shift, even on a finite state space. It is known [6] that such are of the form Bernoulli × finite rotation, and that the finite rotation factor is finitely coded from the Markov partition. It should be true that the isomorphism to Bernoulli × finite rotation can be achieved almost continuously. This should be true even for countable state processes. For the finite state case, the argument is probably just a variant

of that for the mixing case [5]. The countable state argument, although it could be done directly, as could the mixing case we prove here, might be best and most usefully approached by proving a relativized version, as Thouvenot [7] has done for the isomorphism theory, of the f.B. condition.

REFERENCES

[1] J. G. Kemeny, J. L. Snell & A. W. Knapp. *Denumerable Markov Chains.* Springer Verlag: Berlin, 1976.

[2] B. Petit, Deux Schémas de Bernoulli d'alphabet dénombrable et de même entropie sont finitairement isomorphe. Preprint.

[3] D. Rudolph. A characterization of those processes finitarily isomorphic to a Bernoulli shift. *Proceedings of the Special year in Ergodic Theory and Dynamic Systems.* University of Maryland, 1980.

[4] M. Smorodinsky. β-automorphisms are Bernoulli shifts. *Acta. Math. Acad. Sci. Hungar.* 24 (1973), 273–278.

[5] M. Smorodinsky & M. Keane. Finitary isomorphisms of irreducible Markov shifts. *Israel J. Math.* 34 (1979), 281–286.

[6] M. Smorodinsky, P. Shields & R. Adler. Irreducible Markov shifts. *Ann. Math. Statist.* 43 (1972), 1027–1029.

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