Z. Wahrscheinlichkeitstheorie verw. Gebiete
69, 177–186 (1985)

Zeitschrift für
**Wahrscheinlichkeitstheorie**
und verwandte Gebiete


# The Exchangeable Sigma-Field of Markov Chains

José Luis Palacios

Universidad Simón Bolívar, Departamento de Matemáticas, Apartado 80659, Caracas, Venezuela

In memoriam Julio Cortázar (1914–1984)

**Summary.** Blackwell and Freedman [2] proved that the exchangeable $\sigma$-field $\mathscr{E}$ of a homogeneous recurrent Markov chain is atomic. If the chain is finite, the atoms can be found explicitly by means of an algorithm given below. The approach in [2] cannot be extended to a non-homogeneous chain, but a description of $\mathscr{E}$ can be obtained in this case by using coupling methods, provided the chain is finite and satisfies certain conditions.

## Introduction

Let $(S^\infty, \mathscr{s}^\infty)$ be the usual product space and product $\sigma$-field of countably many factors $(S, \mathscr{s})$. Let $\Sigma$ be the class of permutations $\pi$ on the nonnegative integers $\mathbb{N}$ that leave all but finitely many integers unchanged. A set $A \in \mathscr{s}^\infty$ is called *exchangeable* if $\pi A = A$ for all $\pi \in \Sigma$; the collection of all exchangeable sets is the *exchangeable $\sigma$-field $\mathscr{E}$*. Given a sequence of random variables $X_i$: $(\Omega, \mathscr{F}) \to (S, \mathscr{s})$, one defines *the exchangeable $\sigma$-field of the process* $\mathbf{X} = \{X_i\}$, $i \in \mathbb{N}$, as the collection $\mathscr{E}(\mathbf{X})$ of events of the form $\{\mathbf{X} \in A\}$ with $A \in \mathscr{E}$.

The earliest result concerning $\mathscr{E}(\mathbf{X})$ is the Hewitt-Savage 0–1 law (see [3]), which states that if $\mathbf{X}$ is a sequence of i.i.d.'s the $\sigma$-field $\mathscr{E}(\mathbf{X})$ is trivial. Aldous and Pitman gave in [1] a full account of this $\sigma$-field when $\mathbf{X}$ is a sequence of independent, not necessarily identically distributed variables. The study of $\mathscr{E}(\mathbf{X})$ in the case when $\mathbf{X}$ is a recurrent homogeneous countable Markov chain was carried out by Blackwell and Freedman in [2]; they proved that in this context $\mathscr{E}(\mathbf{X})$ is atomic, the atoms being sets of the form $\{X_0 \in I\}$, where $I$ is a certain subset of $S$. Grigorenko ([5]), independently of the results in [2], gave a description of $\mathscr{E}(\mathbf{X})$ in the same context of homogeneous recurrent chains using a graph-theoretical approach, whose only advantage is that it enables one to find by hand the subsets of states $I$ in the atoms mentioned above; in fact, in the first section of this work we give an algorithm which performs this task.

The proofs in [2] and [5] depend basically on this fact: in a homogeneous recurrent chain, the sequences of states between consecutive visits to a fixed state are i.i.d.'s. This property is obviously not shared by general non-homogeneous chains, and thus the approach in [2] cannot be extended to these latter chains. In Sect. 2 we give a description of $\mathscr{E}(\mathbf{X})$ for a non-homogeneous chain whose state space $S$ is finite; the main tool here is a coupling argument which is derived from that of Aldous and Pitman in [1].

It is well known that $\mathscr{E}(\mathbf{X})$ contains the tail $\sigma$-field $\mathscr{T}(\mathbf{X})$. It is also known (see [4]) that for a finite Markov chain, $\mathscr{T}(\mathbf{X})$ is atomic and the number of atoms is bounded by the number of states; the example given at the end of this introduction shows that nothing of this sort can be expected to hold for $\mathscr{E}(\mathbf{X})$, which can be rather complicated. However, the conditions we impose on the transition probabilities in our theorems are sufficient for $\mathscr{E}(\mathbf{X})$ to be trivial or atomic.

*Example.* Let $X_1, X_2, X_3, \ldots$ be a sequence of independent 0–1 valued random variables on some probability space and $\mathscr{F} = \sigma(X_1, X_2, X_3, \ldots)$. Consider the sequence $Y_1, Y_2, Y_3, \ldots$, where $Y_k = X_n$ for $2^{n-1} \leq k < 2^n$, $n = 1, 2, 3, \ldots$. Then $\{Y_n\}$ is a Markov Chain for which $\mathscr{T}(\mathbf{Y})$ is trivial and $\mathscr{E}(\mathbf{Y})$ is the full $\sigma$-field $\mathscr{F}$.

To see this, $\sigma(Y_{2^n}, Y_{2^n+1}, \ldots) = \sigma(X_{n+1}, X_{n+2}, \ldots)$ and $\bigcap_{n=0}^{\infty} \sigma(Y_{2n}, Y_{2n+1}, \ldots) = \bigcap_{n=0}^{\infty} \sigma(X_{n+1}, X_{n+2}, \ldots) =$ trivial by Kolmogorov's 0–1 law. Hence, $\mathscr{T}(\mathbf{Y}) =$ trivial.

Let us now remark that in general, if $Z_1, Z_2, Z_3, \ldots$ is a sequence of random variables, the exchangeable $\sigma$-field $\mathscr{E}(\mathbf{Z})$ generated by these variables can be described as the $\sigma$-field of those events that, for any $n$, depend on the knowledge of the first $n$ variables up to a permutation and the knowledge of the rest of the variables $Z_{n+1}, Z_{n+2}, \ldots$ in order.

Since our $Y_i$'s are 0–1 valued, if we write $S_n = \sum_{i=1}^{n} Y_i$, knowing $S_n$ is the same as knowing $Y_1, \ldots, Y_n$ up to a permutation, and we can write:

$$\mathscr{E}(\mathbf{Y}) = \bigcap_{n=1}^{\infty} \sigma(S_n, Y_{n+1}, Y_{n+2}, \ldots). \tag{1}$$

Furthermore, in this particular example we have that $Y_{2^n-1} = X_n$ implying $S_{2^n-1} = \sum_{i=0}^{n-1} 2^i X_{i+1}$, so that we can determine $X_1, \ldots, X_n$ from the binary expansion of $S_{2^n-1}$, implying:

$$\sigma(S_{2^n-1}, Y_{2^n}, Y_{2^n+1}, \ldots) = \sigma(X_1, X_2, \ldots, X_n, Y_{2^n}, Y_{2^n+1}, \ldots) = \mathscr{F} \tag{2}$$

and since

$$\bigcap_{n=1}^{\infty} \sigma(S_n, Y_{n+1}, Y_{n+2}, \ldots) = \bigcap_{n=1}^{\infty} \sigma(S_{2^n-1}, Y_{2^n}, Y_{2^n+1}, \ldots)$$

then $\mathscr{E}(\mathbf{Y}) = \mathscr{F}$ follows from (1) and (2).

## 1. Homogeneous Case

Let $\mathbf{X} = \{X_i\}$, $i \in \mathbb{N}$ be a homogeneous Markov chain and let $P = \{p(i,j)\}$, $i,j \in S$ be its transition matrix. A sequence of states $\gamma = (i_1, i_2, \ldots, i_n)$ is called a *path* if $\prod_{h=1}^{n-1} p(i_h, i_{h+1}) > 0$.

Let $i,j\in S$; say $i\sim j$ if there are paths $\gamma$, $\delta$ and $k\in S$ such that $i\gamma$ is a permutation of $j\delta$ and $i\gamma k$ and $j\delta k$ are paths. The relation $\sim$ is an equivalence relation and $S$ can be split into equivalence classes, which are called *exchangeable classes.*

For the statement of Theorem 1.2, take $\Omega=S^\infty$, $\mathscr{F}=\mathscr{s}^\infty$ and consider the *shift* $T$: $S^\infty\to S^\infty$ defined by $T(i_1,i_2,i_3,\ldots)=(i_2,i_3,i_4,\ldots)$. The results of Blackwell and Freedman, and Grigorenko can now be given, respectively, as follows:

**Theorem 1.1** (see [2]). If $\mathbf{X}$ *is a recurrent homogeneous Markov chain then* $\mathscr{E}(\mathbf{X})=\sigma(\{X_0\in I_e\})$, *where* $I_e$ *is an exchangeable class.*

**Theorem 1.2** (see [5]). *If* $\mathbf{X}$ *is a recurrent homogeneous Markov chain then* $\mathscr{E}(\mathbf{X})=T(\sigma(X_0)\cap\sigma(X_1))$.

The connection between Theorems 1.1 and 1.2 is this: $\{X_1\in A\}$ is an atom of $\sigma(X_0)\cap\sigma(X_1)$ if and only if $A$ is an exchangeable class, i.e., if and only if $\{X_0\in A\}$ is an atom of $\mathscr{E}(\mathbf{X})$. Once the connection is made, we can give the procedure to find the exchangeable classes:

**Algorithm 1.3.** *Begin by constructing a symbolic transition matrix, whose entry* $(i,j)$ *is* 0 *unless* $p(i,j)>0$ *in which case we write* $X$ *or any other sign.* For illustration consider a $6\times 6$ matrix with $X$'s in the entries (1,3), (1,6), (2,1), (3,1), (3,2), (3,4), (4,3), (4,6), (5,2), (5,4), (6,5) and 0's elsewhere.

*Step 1: Choose an arbitrary entry* $(i_0,j_0)$ *with an* $X$ *and circle it.*
*Step 2: Circle all entries* $(i_s,j_0)$ *and* $(i_0,j_t)$, $0\leqq s\leqq n$, $0\leqq t\leqq n$, *in row* $i_0$ *and column* $j_0$ *bearing an* $X$.
*Step 3: Circle all* $X$*-entries in the row and column of each of the entries in step 2.*
*Step 4. Continue the circling procedure until there are no new* $X$*-entries to circle. When this happens, we will have gotten:*

(i) *One atom of* $\sigma(X_0)\cap\sigma(X_1)$. *Namely, if we consider*
$A=\{i$: $(i,j)$ was selected in steps 1–4$\}$ *and*
$B=\{j$: $(i,j)$ was selected in steps 1–4$\}$, *then* $\{X_0\in A\}=\{X_1\in B\}$.

(ii) *One exchangeable class, namely the set* $B$, *and ergo, one atom of* $\mathscr{E}(\mathbf{X})$: $\{X_0\in B\}$.

In our example, suppose we choose (2,1); after step 2 we get (3,1); after step 3 we get (3,2), (3,4); after step 4 we get (5,2), (5,4). So, $\{X_0\in\{2,3,5\}\}=\{X_1\in\{1,2,4\}\}$ is an atom of $\sigma(X_0)\cap\sigma(X_1)$, $\{1,2,4\}$ is an exchangeable class, and $\{X_0\in\{1,2,4\}\}$ is an atom of $\mathscr{E}(\mathbf{X})$.

*Step 5: If there are* $X$*-entries which have not been circled yet, choose one such entry and repeat steps 1–4 to get another atom of* $\mathscr{E}(\mathbf{X})$. *Clearly this algorithm terminates in a finite number of steps.*

In our example we restart step 1 with, say, entry (1,3), that we mark with a square, and at the end of step 4 we have marked with squares the entries (1,3), (1,6), (4,6), (4,3), so another atom of $\mathscr{E}(\mathbf{X})$ is $\{X_0\in\{3,6\}\}$.

Finally we are left with entry (6,5) so the other atom of $\mathscr{E}(\mathbf{X})$ is $\{X_0=5\}$, and that's it.

*Remarks.* A sufficient condition for $\mathscr{E}(\mathbf{X})$ to be trivial is that the transition matrix be irreducible, to ensure the recurrence of the chain, and have one row with all entries positive, so that all states belong to the same exchangeable class. Another simple sufficient condition is that the matrix be irreducible and for all $i$, $p(i,i)>0$. In fact, we can do better than that if we define, for any subset $I$ of the state space $S=\{1,\ldots M\}$, $F(I)$ to be the set of indices $j\in S$ such that $p(i,j)>0$ for some $i\in I$ (set $F(\emptyset)=\emptyset$), and define a nonnegative $M\times M$ matrix $P$ to be *connected* in case (i) it is irreducible and (ii) for any proper nonempty subset $I$ of $S$, $F(I)\cap F(I^c)\neq\emptyset$. This definition of a connected matrix differs slightly from Seneta's (see [7], p. 70). There, (i) is replaced by the condition that $P$ have no zero row or column. Connected matrices are used by Seneta in the estimation of nonnegative matrices from marginal totals, and with our definition it is not hard to prove that $\mathscr{E}(\mathbf{X})$ is trivial if and only if the transition matrix $P$ is connected. It is also readily seen that (i) a connected matrix $P$ is also primitive, i.e., there is a positive integer $t$ such that $P^t>0$, (ii) if $P$ is connected then all powers of $P$ are connected, implying that if $\mathscr{E}(\mathbf{X})$ is trivial, so is the exchangeable $\sigma$-field generated by the sequence $X_0$, $X_n$, $X_{2n}$, $X_{3n},\ldots$ for all $n$ (compare with Corollary 2.1.2. in [4]), and (iii) if $P$ is connected then its transpose is also connected, implying that if $\mathscr{E}(\mathbf{X})$ is trivial, so is the exchangeable $\sigma$-field associated to the "reversed" chain $\hat{X}_n$ with transitions satisfying $\hat{p}(i,j)>0$ iff $p(j,i)>0$. For all these details see [6].

## 2. Nonhomogeneous Case

Let $\mathbf{X}=\{X_n\}$ be a nonhomogeneous Markov chain with finite state space $S=\{1,2,\ldots,M\}$, transition probability matrices (t.p.m.) $P_n=\{p_n(i,j)\}$, $i,j\in S$ and initial distribution $\pi=\{\pi(i)\}$, $i\in S$ determining the Markov measure $P$ on $(\Omega,\mathscr{F})$. Let $P_i$ be the Markov measure determined by the same t.p.m. and initial distribution concentrated on the state $i$.

Consider a bivariate process $\tilde{\mathbf{X}}=\{\tilde{X}_n\}=\{(X_n^1,X_n^2)\}$ on the space $(\tilde{\Omega},\tilde{\mathscr{F}})$, the usual product of $(\Omega,\mathscr{F})$ by itself. The new state space is $\tilde{S}=S\times S$. The diagonal of $\tilde{S}$ is $D=\{(s,s)\colon s\in S\}$. Let $T_D$ be a random time on $(\tilde{\Omega},\tilde{\mathscr{F}})$ and define $\tilde{\Omega}^*=\{\tilde{\omega}\in\tilde{\Omega}\colon T_D(\tilde{\omega})<\infty$ and $\tilde{X}_n(\tilde{\omega})\in D$ for all $n>T_D\}$, $\tilde{\Omega}^\infty=\{\tilde{\omega}\in\tilde{\Omega}\colon T_D(\tilde{\omega})=\infty\}$.

Let $i,j\in S$ and let $\tilde{P}_{ij}$ be a distribution for $\tilde{\mathbf{X}}$ on $(\tilde{\Omega},\tilde{\mathscr{F}})$. We say that the process $\tilde{\mathbf{X}}$ is a *coupling* for $\mathbf{X}$ (or $\tilde{P}_{ij}$ is a coupling for $P_i$ and $P_j$) in case:

(i) $\tilde{P}_{ij}(\cdot\times\Omega)=P_i(\cdot)$ and $\tilde{P}_{ij}(\Omega\times\cdot)=P_{ij}(\cdot)$

(ii) $\tilde{P}_{ij}(\tilde{\Omega}^*+\tilde{\Omega}^\infty)=1$

for any two starting points $i,j$. We say the coupling is *successful* if $\tilde{P}_{ij}(\tilde{\Omega}^*)=1$. Condition (i) states that the marginal processes $\{X_n^1\}$, $\{X_n^2\}$ are copies of the given chain started at $i$ and $j$ respectively; condition (ii) requires that $\{\tilde{X}_n\}$ remain on the diagonal after $T_D$; we do not rule out the case that $T_D$ might be infinite (i.e., the coupling might not be successful). $T_D$ is not necessarily the first time that the marginal processes meet.

Define the modified measure $P_j^m$, $m\in\mathbb{N}$ on $(\Omega,\mathscr{F})$ as the Markov measure started at $j$ with t.p.m. $\{Q_n\}$, where $Q_n=P_{n+m}$. That is, $P_j^m$ is the distribution of

$(X_m, X_{m+1}, X_{m+2}, \ldots)$ given that $X_m = j$. Now we can give the statement of the following lemma whose proof can be found in [4].

**Lemma 2.1.** *If for every m, i, j there is a successful coupling* $\tilde{P}_{ij}^m$ *for* $P_i^m$ *and* $P_j^m$ *(i.e.,* $\tilde{P}_{ij}^m(T_D < \infty) = 1$*), then the tail* $\sigma$*-field* $\mathcal{T}(\mathbf{X})$ *is trivial.*

To state another important lemma we need to introduce some more notation. If $\mathbf{X}$ is the given Markov chain, the *augmented chain* $\mathbf{Y} = \{Y_n\}$, $n \in \mathbb{N}$ is defined by $Y_n(\omega) = (m_n(\omega), X_n(\omega))$, where $m_n$ is the *occupation vector of* $\mathbf{X}$ *until time n* which keeps track of the states visited until time $n$ and is defined as a vector in $\mathbb{N}^M$ whose $i$-th coordinate is the integer $q$ iff the state $i$ appears exactly $q$ times in the list $\{X_0(\omega), X_1(\omega), \ldots, X_n(\omega)\}$. The state space of the augmented chain is $S' = \mathbb{N}^M \times S$, and we can think of $\{Y_n\}$ as the coordinate process on $\Omega'$, the product space of countably many identical copies of $S'$. Call $\mathcal{F}'$ the corresponding product $\sigma$-field. In this setup we have the following lemma ([1] Lemma 4.6):

**Lemma 2.2.** *The tail* $\sigma$*-field of the process* $\mathbf{Y}$, $\mathcal{T}(\mathbf{Y})$, *is equal to the exchangeable* $\sigma$*-field of the process* $X$, $\mathcal{E}(\mathbf{X})$.

We will try to apply Lemma 2.1 to the augmented chain $\{Y_n\}$ next. We will need, then, to construct a successful coupling $\tilde{P}^m_{(v,i),(w,j)}$ for the bivariate process $\{\tilde{Y}_n\} = \{(Y_n^1, Y_n^2)\}$ started at an arbitrary time $m$ at the places $(v, i)$ and $(w, j)$ where $i, j \in S$, and $v = (v_1, v_2, \ldots, v_M)$ and $w = (w_1, w_2, \ldots, w_M)$ are the occupation vectors of two paths of length $m$ with $v_i > 0$ and $w_j > 0$. We will call such occupation vectors $v$ and $w$, *permissible for m, i and j.*

With this notation we have the following:

**Lemma 2.3.** $\mathcal{E}(\mathbf{X})$ *is trivial if conditions* (a) *and* (b) *are met:*

a) *For all* $i, j \in S$ *and* $m \in \mathbb{N}$ *there is a successful coupling* $\tilde{P}_{ij}^m$ *of the ordinary chain.*

b) *For all* $i \in S$, $m \in \mathbb{N}$ *and* $v, w$ *occupation vectors permissible for i and m, there is a successful coupling* $\tilde{P}_{ii}^m$ *of the ordinary chain with both marginals started at i, such that the occupation vectors of the marginal processes satisfy:*

$$m_n(X^1) - m_n(X^2) = w - v$$

*for all n sufficiently large.*

*Proof.* By Lemma 2.2 we reduce the problem of proving $\mathcal{E}(\mathbf{X}) = $ trivial, to that of proving $\mathcal{T}(\mathbf{Y}) = $ trivial, and this latter problem in turn is reduced by Lemma 2.1 to finding couplings for the augmented chain.

Let us start with arbitrary $i, j \in S$, $m \in \mathbb{N}$ and $v, w$, permissible for $i, j, m$. First we will define the distribution $\hat{P}_{ij}^m$ of the process $\tilde{\mathbf{X}}$ as follows: Allow $\tilde{X}_n$ to evolve according to $\tilde{P}_{ij}^m$ until the marginal processes meet for the first time, i.e., define:

$$\hat{P}_{ij}^m(\tilde{X}_{n+1} | \tilde{X}_n, \ldots, \tilde{X}_0;\ \tilde{X}_i \notin D, 1 \leq i \leq n) := \tilde{P}_{ij}^m(\tilde{W}_{n+1} | \tilde{W}_n, \ldots, \tilde{W}_0),$$

where $\tilde{\mathbf{W}}$ is the coupling for $\mathbf{X}$ under the measure $\tilde{P}_{ij}^m$. By (a), $\tilde{X}_n$ reaches the diagonal $D$ in a finite time. Assume that the marginals meet at time $N$ at the state $k$; then $m_N(X^1)=m_N(W^1)=v'$, say, and $m_N(X^2)=m_N(W^2)=w'$, and $v+v'$ and $w+w'$ are permissible for $m+N$ and $k$. Now let the process $\tilde{X}_n$ evolve according to the coupling $\tilde{P}^{m+N}_{kk}$, i.e., define

$$\hat{P}^m_{ij}\left(\tilde{X}_{N+r}|\tilde{X}_{N+r-1},\ldots,\tilde{X}_0;\ \tilde{X}_i\notin D, 1\leqq i\leqq N-1;\right.$$
$$\left.\tilde{X}_n=\begin{pmatrix}k\\k\end{pmatrix}\right):=\tilde{P}^{m+N}_{kk}(\tilde{Z}_r|\tilde{Z}_{r-1},\ldots,\tilde{Z}_0),$$

where $\tilde{\mathbf{Z}}$ is the coupling for $\mathbf{X}$ under the measure $\tilde{P}^{m+n}_{kk}$ mentioned in (b).

By (b), the occupation vectors of the marginals will satisfy: $m_r(Z^1)-m_r(Z^2)=w+w'-(v+v')$ for all $r$ sufficiently large.

Once the distribution $\hat{P}^m_{ij}$ of $\tilde{\mathbf{X}}$ is defined in the manner described above, we define on $(\tilde{\Omega}',\tilde{\mathscr{F}}')$, the product of $(\Omega',\mathscr{F}')$ by itself, the distribution $\tilde{P}^m_{(v,i),(w,j)}$ of $\tilde{\mathbf{Y}}$ as follows:

Define $\tilde{P}^m_{(v,i),(w,j)}(\tilde{Y}_0=((v,i),(w,j)))=1$ and

$$\tilde{P}^m_{(v,i),(w,j)}(\tilde{Y}_{n+1}=((v_{n+1},i_{n+1}),(w_{n+1},j_{n+1}))|\tilde{Y}_n=((v_n,i_n),(w_n,j_n)))$$
$$:=\begin{cases}\hat{P}^m_{ij}(\tilde{X}_{n+1}=(i_{n+1},j_{n+1})|\tilde{X}_n=(i_n,j_n)) & \text{in case: }\begin{cases}v_{n+1}=v_n+e_{i_{n+1}}\\ w_{n+1}=w_n+e_{j_{n+1}}\end{cases}\\ 0\quad\text{otherwise}\end{cases}$$
$$\text{for all } n\geqq 0$$

($e_k$ is the vector in $\mathbb{N}^M$ whose entries are 0 except the $k$-th, which is 1).

Then, by construction:

i) The marginals of $\tilde{P}^m_{(v,i),(w,j)}$ are $P^m_{(v,i)}$, and $P^m_{(w,j)}$, where the latter probabilities are the distributions of $(Y_m, Y_{m+1}, Y_{m+2},\ldots)$ conditional on $Y_m$ equal to $(v,i)$ and $(w,j)$ respectively. In other words, $\tilde{P}^m_{(v,i),(w,j)}$ is a coupling for the augmented chain.

ii) $\tilde{P}^m_{(v,i),(w,j)}$ is successful because $\hat{P}^m_{ij}$ is successful (i.e., $(X^1_n, X^2_n)$ eventually stays in the diagonal $D$), and also we have under $\tilde{P}^m_{(v,i),(w,j)}$ that $m_N(X^1)-m_N(X^2)=w-v$ for all $n$ sufficiently large, or $m_N(X^1)+v=m_N(X^2)+w:=v_n$ for all sufficiently large $n$; or, in other words, $\tilde{Y}_n$ stays eventually in the diagonal $D'=\{(s',s'): s'\in S'\}$.

The next simple result will be essential when constructing couplings:

**Lemma 2.4.** *If $\mu$ is a probability on a countable space $S$ and $\mu(x_1)\geqq\delta$, $\mu(x_2)\geqq\delta$, then there exists a pair of $S$-valued random variables $(Y,Z)$ such that both $Y$ and $Z$ have $\mu$ as distribution and:*

$$P(Y=x_1,Z=x_2)=\delta;\quad P(Y=x_2,Z=x_1)=\delta;\quad P(Y=Z)=1-2\delta.$$

*Proof.* Define the joint distribution as follows:

$$P(Y=x_1,Z=x_2)=\delta;\quad P(Y=x_2,Z=x_1)=\delta;\quad P(Y=Z=x_1)=\mu(x_1)-\delta;$$
$$P(Y=Z=x_2)=\mu(x_2)-\delta;\quad P(Y=Z=x_i)=\mu(x_i)\quad\text{for } i\geqq 3.$$

Lemma 2.5 and Theorem 2.6 below yield conditions for $\mathcal{T}(\mathbf{X})$ and $\mathcal{E}(\mathbf{X})$ to be trivial; these conditions are generalizations of this simpler condition: all entries of all matrices are greater than a fixed positive number. The proof of Lemma 2.5 can be found in [4]; here the idea is to use what Griffeath calls "the classical coupling", due in its original form to W. Doeblin.

**Lemma 2.5.** *If $\{X_n\}$ is a Markov Chain with $\delta_n = \min_{i,j}\{p_{n+1}(i,j)\}$ satisfying $\sum_{n=0}^{\infty} \delta_n = \infty$, then Lemma 2.1 applies, and $\mathcal{T}(\mathbf{X})$ is trivial.*

**Theorem 2.6.** *If $\{X_n\}$ is a Markov Chain with $\delta_n = \min_{i,j,s}\{p_{n+1}(i,j)\,p_{n+2}(j,s)\}$, $n = 0, 1, 2, \ldots$ satisfying $\sum_n \delta_n = \infty$, then Lemma 2.3 applies and $\mathcal{E}(\mathbf{X})$ is trivial.*

*Proof.* Since $\min_{i,j}\{p_n(i,j)\} \geqq \min_{i,j,s}\{p_n(i,j)\,p_{n+1}(j,s)\}$, $n = 1, 2, 3, \ldots$, Lemma 2.5 applies and condition (a) of 2.3 is met.

Let $k \in S$, $m \in \mathbb{N}$, and $v, w$ permissible for $k, m$. Let $u_n = m_n(X^1) - m_n(X^2) - v + w$; $u_n = (u_n^1, u_n^2, \ldots, u_n^M)$. We are going to show that (b) of 2.3 is also satisfied.

For notational convenience, we will denote $\tilde{X}_n = \begin{pmatrix} X_n^1 \\ X_n^2 \end{pmatrix}$, so that $(\tilde{X}_n, \tilde{X}_{n+1})$ is to be read as $\begin{pmatrix} X_n^1 & X_{n+1}^1 \\ X_n^2 & X_{n+1}^2 \end{pmatrix}$, etc.

*Step I:* Look at the smallest $t$ for which $u_0^t \neq 0$. Look at the states $t$ and $M$ (the choice of this latter is arbitrary; any other state would work) and define the joint distribution of $(\tilde{X}_{2n+1}, \tilde{X}_{2n+2})$ given that $\tilde{X}_{2n} = \begin{pmatrix} s \\ s \end{pmatrix}$ as follows for $n = 0, 1, 2, \ldots$ and $s \in S$ arbitrary: Apply Lemma 2.4 with $\mu$ being the distribution of the pair $(X_{m+2n+1}, X_{m+2n+2})$ conditional on $X_{m+2n} = s$, and $x_1$ and $x_2$ being, respectively $(t, M)$ and $(M, M)$, i.e.,

$$\tilde{P}_{kk}^m\left((\tilde{X}_{2n+1}, \tilde{X}_{2n+2}) = \begin{pmatrix} t & M \\ M & M \end{pmatrix} \middle| \tilde{X}_{2n} = \begin{pmatrix} s \\ s \end{pmatrix}\right) = \delta_{m+2n},$$

$$\tilde{P}_{kk}^m\left((\tilde{X}_{2n+1}, \tilde{X}_{2n+2}) = \begin{pmatrix} M & M \\ t & M \end{pmatrix} \middle| \tilde{X}_{2n} = \begin{pmatrix} s \\ s \end{pmatrix}\right) = \delta_{m+2n},$$

$$\tilde{P}_{kk}^m\left((\tilde{X}_{2n+1}, \tilde{X}_{2n+2}) \in D \times D \middle| \tilde{X}_{2n} = \begin{pmatrix} s \\ s \end{pmatrix}\right) = 1 - 2\delta_{m+2n}.$$

Notice that the probabilities thus defined do not depend on $s$. Now we take

$$\tilde{P}_{kk}^m(\tilde{X}_{2n+1}, \tilde{X}_{2n+2} \mid \tilde{X}_{2n}, \tilde{X}_{2n-1}, \ldots, \tilde{X}_1, \tilde{X}_0) = \tilde{P}_{kk}^m(\tilde{X}_{2n+1}, \tilde{X}_{2n+2} \mid \tilde{X}_{2n}).$$

In other words, in step I, the behavior of $\tilde{\mathbf{X}}$ under $\tilde{P}_{kk}^m$ is this:

If $(X_{2n+1}^1, X_{2n+2}^1) = (t, M)$, then $(X_{2n+1}^2, X_{2n+2}^2)$ is forced to be either $(t, M)$ or $(M, M)$, and if $(X_{2n+1}^1, X_{2n+2}^1)$ takes a value $(i, j)$ different from either $(t, M)$ or $(M, M)$, then $(X_{2n+1}^2, X_{2n+2}^2)$ takes the same value $(i, j)$.

Define now for $n=1,2,3,\ldots$

$$Z_n := \begin{cases} 1 & \text{if } (\tilde{X}_{2n+1}, \tilde{X}_{2n+2}) = \begin{pmatrix} t & M \\ M & M \end{pmatrix} \\ -1 & \text{if } (\tilde{X}_{2n+1}, \tilde{X}_{2n+2}) = \begin{pmatrix} M & M \\ t & M \end{pmatrix} \\ 0 & \text{otherwise} \end{cases}$$

and consider the vectors $u_n$ every two steps, that is: define $d_n = u_{2n}$ for $n = 0,1,2,\ldots$ and look at their $t$-th coordinates $d_n^t$ for $n=0,1,2,\ldots$. It is clear that $d_N^t = d_0^t + Z_1 + Z_2 + Z_3 + \ldots + Z_n$ and the $Z_n$'s are independent r.v.'s on $(\tilde{\Omega}, \tilde{\mathscr{F}})$ with distribution:

$$\tilde{P}_{kk}^m(Z_n = 1) = \tilde{P}_{kk}^m(Z_n = -1) = \delta_{m+2n}, \qquad \tilde{P}_{kk}^m(Z_n = 0) = 1 - 2\delta_{m+2n}.$$

By Borel-Cantelli, $\tilde{P}_{kk}^m(Z_n = 1 \text{ i.o.}) = \tilde{P}_{kk}^m(Z_n = -1 \text{ i.o.}) = 1$, so that if we look at the sequence $d_n^t$, $n=0,1,2,\ldots$ only when the $Z_n$'s jump to $Z_n+1$ or $Z_n-1$ we have a symmetric simple random walk that hits zero in a finite time with probability 1, so that the $t$-th coordinate of $u_n$ equals 0 in a finite time $T_1$. In fact, $u_{T_1}^1 = u_{T_1}^2 = \ldots = u_{T_1}^{t-1} = u_{T_1}^t = 0$.

*Step II.* Now look at the smallest $r > t$ for which $u_{T_1}^r \neq 0$, and define the joint distribution of the next pairs after time $T_1$ as in step I, substituting the state $t$ by the state $r$; this will make the $r$-th coordinate and all the preceding coordinates of $u_n$ equal to zero in a finite time $T_2$. Repeat this procedure until all coordinates of $u_n$ are 0 at the finite time $T_D$.

*Step III.* Once $u_n = (0,0,\ldots,0)$, force both marginal processes to evolve together. Define:

$$\tilde{P}_{kk}^m \left( \tilde{X}_{n+1} = \begin{pmatrix} i \\ i \end{pmatrix} \middle| \tilde{X}_n = \begin{pmatrix} j \\ j \end{pmatrix} \right) = p_{n+m}(j,i) \qquad \text{for all } n \geqq T_D.$$

*Remark.* As mentioned before, Theorem 2.6 applies when $p_n(i,j) \geqq \delta > 0$ for all $i$, $j$, $n$ but does not work for instance when $p_n(i,j) = 0$ for all $n$ for a fixed entry $(i,j)$. However, we can extend the ideas in the proof of 2.6 to the case when all transition matrices have the same symbolic matrix, which is connected (see remark after Algorithm 2.3) and for all entries $(i,j)$ either $p_n(i,j) = 0$ for all $n$ or $p_n(i,j) \geqq \delta > 0$ for all $n$. In fact we can go a bit farther as the next theorem shows.

**Theorem 2.7.** *Assume all transition matrices have the same symbolic connected matrix. Let* $L = 4M^2 - 2M - 1$ *and* $Q_n = P_n P_{n+1} \ldots P_{n+L}$. *Let* $\delta_n = \min_{i,j} \{q_n(i,j)\}$. *Assume that* $\sum_n \delta_n = \infty$.

*Then* $\mathscr{E}(\mathbf{X}) =$ *trivial.*

*Sketch of Proof.* Since the symbolic matrix is connected, it is primitive (see remarks in Sect. 1), and according to Wielandt's theorem (see [7] p. 58) all the matrices $Q_n$ are positive; that means that it is possible to go from any state to any other state in $L$ steps, and thus the same idea in 2.6 works here mutatis mutandis, i.e.: the classical coupling is successful when we look at the chain $X_0, X_L, X_{2L}, \ldots$, and therefore (a) of 2.3 is satisfied.

To prove that (b) of 2.3 is also satisfied, we modify the technique in 2.6 to construct a successful coupling as follows: we force the two copies of the chain to move in $L$-step jumps either along the same path or different paths that amount to an eventual equalization of the coordinates of $u_n$ to zero by the simple random walk argument. The choice of these $L$-long paths goes like this in case we want to make $u_n^t$, the $t$-th coordinate of $u_n$, equal to zero: because the symbolic matrix is connected (and so there is only one exchangeable class), given the states $t$ and $M$, there are paths $\alpha, \beta$ and a state $k$ such that $t\alpha k$ and $M\beta k$ are paths that are permutations of one another and by a simple combinatorial argument (see [6]) can be shown to have length $2M^2-M$. Because the transpose of the matrix is also connected, for the same $t$ and $M$, there are paths $\hat{\delta}$ and $\hat{\varepsilon}$ and a state $l$ such that $t\hat{\delta}l$ and $M\hat{\varepsilon}l$ are paths of length $2M^2-M$ which are permutations of one another. If we reverse the hatted paths we get that $\gamma_1 = l\delta t\alpha k$ and $\gamma_2 = l\varepsilon M\beta k$ are paths of length $L$ which are permutations of one another *except that* $t$ *appears once more in* $\gamma_1$ *than it does in* $\gamma_2$ *and* $M$ *appears once more in* $\gamma_2$ *than it does in* $\gamma_1$ and thus, if we force the marginals of the chain to move along these paths as in step I of 2.6 we will get $u_n^t = 0$ in a finite time.

The argument used in the proof of 2.7 is valid only if the symbolic matrix is connected. If we weaken that hypothesis and require only that the symbolic matrix be irreducible (this would imply recurrence in the homogeneous case), we need to strengthen the assumptions on the transition probabilities in order to be able to use the same coupling ideas. In this connection we have the following

**Theorem 2.8.** *Assume all transition matrices have the same symbolic irreducible matrix and for every entry* $(i,j)$ *either* $p_n(i,j)=0$ *or* $p_n(i,j) \geqq \delta > 0$ *for all* $n$. *Then* $\mathscr{E}(\mathbf{X})$ *is trivial under the probability* $P_i$ *for all* $i$.

*Sketch of Proof.* Fix $i \in S$. We decompose the state space into communicating classes that in turn are decomposed into cyclically moving subclasses. (These concepts depend only on the existence of appropriate paths and can be used in our nonhomogeneous setup.) Without loss of generality, we assume that there is only one communicating class partitioned into $d$ subclasses. Then the chain $X_0 = i, X_d, X_{2d}, \ldots$ is a chain for which 2.5 applies, i.e., the classical coupling is successful and thus (a) of 2.3 holds.

To prove that (b) of 2.3 also holds, suppose that we start both marginals in some state $k$. We force the marginals to evolve together until they hit the state $i$ at some time $n$. Then, since the underlying measure is $P_i$, we find two paths $\gamma_1, \gamma_2$ of length, say, $L$ whose endpoints are the state $i$ and such that they are composed of the same states visited by the respective marginals up to time $n$. After that we force both copies of the chain to move in $L$-long jumps either together or in such a way that if one of them goes along $\gamma_1$ the other goes along $\gamma_2$ and vice versa. A random walk argument like that of 2.6 and 2.7 finishes the proof.

**Corollary.** *With the hypothesis of* 2.6 $\mathscr{E}(\mathbf{X})$ is atomic, namely $\mathscr{E}(\mathbf{X})=\sigma(\{X_0 \in I_e\})$ where $I_e$ *represents the exchangeable classes of the symbolic matrix.*

*Proof.* Adapt the argument on p. 1294 of [2] to this nonhomogeneous setup.

## References

1. Aldous, D., Pitman, J.: On the zero-one law for exchangeable events. Ann. Probability **7**, 704–723 (1979)
2. Blackwell, D., Freedman, D.: The tail $\sigma$-field of a Markov Chain and a theorem of Orey. Ann. Math. Statist. **35**, 1291–1295 (1964)
3. Breiman, L.: Probability. Reading, Massachusetts: Addison-Wesley 1968
4. Griffeath, D.: Coupling methods for Markov processes. In: Studies in Probability and Ergodic Theory. New York: Academic Press 1978
5. Grigorenko, L.A.: On the $\sigma$-algebra of symmetric events for a countable Markov Chain. Theory Probability Appl. **24**, 199–204 (1979)
6. Palacios, J.L.: The exchangeable sigma-field of Markov Chains. Ph.D. thesis. University of California. Berkeley (1982)
7. Seneta, E.: Nonnegative matrices and Markov Chains. New York: Springer 1981

Received January 18, 1983; in revised form March 26, 1984