Z. Wahrscheinlichkeitstheorie verw. Gebiete
50, 273-285 (1979)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Tail Structure of Markov Chains on Infinite Product Spaces

H. Föllmer

Mathematikdepartement, ETH-Zentrum, CH-8092 Zürich

Dedicated to L. Schmetterer

1. Introduction

Consider a Markov chain given by a transition probability P on some measurable state space $(E, \mathscr{E})$. For an initial probability distribution μ on $\mathscr{E}$ let μP^n denote the resulting distribution in period n. If P is an aperiodic Harris chain then Orey's theorem implies

$$(1.1) \quad \lim_n \|\mu P^n - \nu P^n\|_{\mathscr{E}} = 0$$

for all initial distributions μ, ν, where $\| - \|_{\mathscr{E}}$ denotes the variational distance for probability measures on $\mathscr{E}$. By well known results of Blackwell, Orey et al., this *asymptotic loss of memory* can be characterized in various equivalent ways. In potential theoretic terms it means that all bounded space-time harmonic functions are constant. In the canonical model $(\Omega, (X_n), (P_\mu))$ where Ω is the space of all trajectories on E and P_μ is the measure on Ω induced by the initial distribution μ, it is equivalent to the mixing condition

$$(1.2) \quad P_\mu = 0 - 1 \quad \text{on } \mathscr{A}$$

for each initial μ, where $\mathscr{A} = \bigcap_n \bigvee_{k \geqq n} X_k^{-1}(\mathscr{E})$ is the σ-field of asymptotic events.

The purpose of this paper is to adapt these results to models for the time evolution of infinite particle systems where the state space is an infinite product space; see for example [2, 3]. In these models, the chain is typically not a Harris chain, and property (1.2) is too restrictive. We shall therefore look at suitable modifications of (1.2), prove the corresponding versions of the Blackwell-Orey equivalence, and give conditions of Dobrushin-Wasserstein type which guarantee that these modified properties do hold.

To this end, we first consider a general Markov chain and prove the following extension of the Blackwell-Orey equivalence. Let $\mathscr{A}_1$ be a sub-σ-field of $\mathscr{A}$ and denote by $\mathscr{H}_1$ the class of bounded space-time harmonic functions $h = (h_n)_{n \geqq 0}$ associated to $\mathscr{A}_1$; see (2.10). For each event $A \in \mathscr{A}_1$ there is a function h

0044-3719/79/0050/0273/$02.60

$=(h_n)\in\mathscr{H}_1$ such that

(1.3) $P_\mu[A\,|\,X_0,\ldots,X_n]=h_n(X_n)$.

It is therefore of interest to clarify the measurability properties of the functions in $\mathscr{H}_1$: They show which information is needed in order to make predictions on $\mathscr{A}_1$. Let $\mathscr{E}_0$ be a sub-σ-field of $\mathscr{E}$, and let us say that $h=(h_n)$ is $\mathscr{E}_0$-measurable if each h_n is $\mathscr{E}_0$-measurable. Denote by $\mathscr{A}_0=\bigcap_n \bigvee_{k\geqq n} X_k^{-1}(\mathscr{E}_0)$ the sub-σ-field of $\mathscr{A}$ corresponding to $\mathscr{E}_0$, and consider the following conditions:

(1.4) $\mu=\nu$ on $\mathscr{E}_0 \Rightarrow P_\mu=P_\nu$ on $\mathscr{A}_1$,

(1.5) $\mathscr{A}_1=\mathscr{A}_0 \bmod P_\mu$ for all initial μ,

(1.6) Each $h\in\mathscr{H}_1$ is $\mathscr{E}_0$-measurable.

Theorem (2.11) shows that these conditions are essentially equivalent. For $\mathscr{E}_0=\{\emptyset,\Omega\}$ and $\mathscr{A}_1=\mathscr{A}$, condition (1.4) is equivalent to (1.1), and the theorem reduces to the classical case.

Sections 3 and 4 contain applications to the infinite particle case. Here the state space is of the form $(E,\mathscr{E})=\prod_{i\in I}(E_i,\mathscr{E}_i)$, where I is a countable set of sites and $(E_i,\mathscr{E}_i)$ is an individual state space for site $i\in I$. We denote by $\mathscr{V}$ the class of finite subsets of I, by $\mathscr{E}_V$ the σ-field on E generated by the coordinates in $V\in\mathscr{V}$, and by $\mathscr{A}_V=\bigcap_n \bigvee_{k\geqq n} X_k^{-1}(\mathscr{E}_V)$ the σ-field on Ω which describes the asymptotic behavior of the sites in V. The first application consists in localizing the mixing condition (1.2). We obtain different characterizations of the condition that

(1.7) $P_\mu=0-1$ on $\mathscr{A}_{\text{loc}}=\bigvee_{V\in\mathscr{V}}\mathscr{A}_V$

for each initial μ, taking $\mathscr{E}_0=\{\emptyset,\Omega\}$ and $\mathscr{A}_1=\mathscr{A}_{\text{loc}}$ in (1.4); see Theorem (3.2). *Local mixing* in the sense of (1.7) is stronger than *local convergence*

(1.8) $\lim_n \|\mu P^n-\nu P^n\|_{\mathscr{E}_V}=0 \quad (V\in\mathscr{V})$

where $\|-\|_{\mathscr{E}_V}$ denotes the variational distance between probability measures on $\mathscr{E}_V$. Let a_{ik} measure the influence of site i on site k; see (3.6) below. Wasserstein [8] has shown that the Dobrushin condition

(1.9) $\sup_k \sum_i a_{ik}<1$

implies local convergence, and that it is actually enough to require

(1.10) $\lim_n \sum_i a_{ik}^{(n)}=0 \quad (k\in I)$

for the powers of the matrix (a_{ik}). Theorem (3.11) shows that the Dobrushin condition even implies local mixing, and here it is enough to require

(1.11) $\sum_{i,k} a_{ik}^{(n)}<\infty \quad (k\in I)$.

The second application concerns the *boundary process* which arises if we observe the process (X_n) only via the σ-field $\hat{\mathscr{E}}=\bigcap_{V\in\mathscr{V}}\mathscr{E}_{I-V}$ of events which do not depend on any finite number of sites. Let $\hat{\mathscr{A}}$ be the corresponding boundary tail field. Taking $\mathscr{E}_0=\hat{\mathscr{E}}$ and $\mathscr{A}_1=\mathscr{A}$ in (1.5), we obtain different characterizations of the condition that

(1.12) $\mathscr{A}=\hat{\mathscr{A}} \bmod P_\mu$

for any initial μ; see Theorem (4.4). Here, the analogue of condition (1.1) is

(1.13) $\mu=v$ on $\hat{\mathscr{E}}\Rightarrow\lim_n \|\mu P^n - vP^n\|_{\hat{\mathscr{E}}}=0.$

The potential theoretic version says that the best prediction of any asymptotic event does not require the detailed description of the present state X_n but only the boundary information. Theorem (4.6) shows that the process does have these properties if the dual version

(1.14) $\lim_n \sum_k a_{ik}^{(n)}=0 \qquad (i\in I)$

of the Wasserstein condition (1.10) holds.

In part, the results were obtained during a visit at Cornell, and thanks are due to E.B. Dynkin, H. Kesten and F. Spitzer for several discussions.

2. An Extension of the Blackwell-Orey Equivalence

Let $(E,\mathscr{E})$ be a measurable space. We denote by $\mathscr{E}^b$ the class of bounded $\mathscr{E}$-measurable functions on E and by $\mathscr{M}$ the class of probability measures on $\mathscr{E}$. For $f\in\mathscr{E}^b$ we write

$$\|f\|=\sup_{x\in E}|f(x)|,$$

and for μ, $v\in\mathscr{M}$ we define

$$\|\mu-v\|_{\mathscr{E}}=\sup_{A\in\mathscr{E}}|\mu(A)-v(A)|=(\mu-v)^+_{\mathscr{E}}(E)$$

where $(\mu-v)^+_{\mathscr{E}}$ is the positive part in the Hahn decomposition of the signed measure $\mu-v$ on $\mathscr{E}$.

(2.1) *Remark*. Suppose we have σ-fields $\mathscr{E}_n\subseteq\mathscr{E}$ which either decrease to $\mathscr{E}_\infty=\bigcap_n\mathscr{E}_n$ or increase to $\mathscr{E}_\infty=\bigvee_n\mathscr{E}_n$. Then

(2.2) $\lim\|\mu-v\|_{\mathscr{E}_n}=\|\mu-v\|_{\mathscr{E}_\infty}.$

Here is a short proof (a more involved argument appears in [4]): Defining $\gamma=\frac{1}{2}(\mu+v)$ and denoting by φ_n and ψ_n the densities of μ and v relative to γ on $\mathscr{E}_n$, we have

$$\|\mu-v\|_{\mathscr{E}_n}=\int(\varphi_n-\psi_n)^+\,d\gamma \qquad (n=1,\ldots,\infty),$$

and (2.2) is now obvious by martingale convergence backwards respectively forwards.

A *transition probability* from $(E, \mathscr{E})$ to some measurable space $(E', \mathscr{E}')$ is a map $P: E \times \mathscr{E}' \to [0, 1]$ such that

i) $P(x, .)$ is a probability measure on $\mathscr{E}'$ $(x \in E)$,
ii) $P(., A) \in \mathscr{E}^b$ $(A \in \mathscr{E}')$.

For $\mu \in \mathscr{M}$ and $f \in (\mathscr{E}')^b$ we denote by μP the induced probability measure $\mu P(.) = \int \mu(dx) P(x, .)$ on $\mathscr{E}'$ and by Pf the induced function $Pf(.) = \int P(., dy) f(y) \in \mathscr{E}^b$.

Let P be a transition probability on $(E, \mathscr{E})$, and let us recall the canonical model for the associated Markov chain. Define the set $\Omega = E^{\{0, 1, \ldots\}}$ of all trajectories $\omega: \{0, 1, \ldots\} \to E$ and for each $n \geqq 0$ the coordinate map X_n with $X_n(\omega) = \omega(n)$. Introduce the σ-fields

$$\mathscr{F}_n = \sigma(X_0, \ldots, X_n), \qquad \mathscr{F}_n^* = \sigma(X_n, X_{n+1}, \ldots),$$
$$\mathscr{F} = \bigvee_n \mathscr{F}_n, \qquad \mathscr{A} = \bigcap_n \mathscr{F}_n^*$$

on Ω. $\mathscr{A}$ is called the σ-field of asymptotic events or the *tail field*. For each $\mu \in \mathscr{M}$, the stochastic evolution starting with μ and governed by P is then given by the unique probability measure P_μ on $(\Omega, \mathscr{F})$ such that

i) $P_\mu[X_0 \in A] = \mu(A) \quad (A \in \mathscr{E})$,
ii) $P_\mu[X_{n+1} \in A \mid \mathscr{F}_n] = P(X_n, A) \quad (A \in \mathscr{E}, n \geqq 0)$.

$E_\mu[.]$ resp. $E_x[.]$ will denote the expectation with respect to P_μ resp. $P_x = P_{\varepsilon_x}$. Let us also introduce the class

$$\mathscr{H} = \{h = (h_n)_{n \geqq 0} \mid h_n \in \mathscr{E}^b, h_n = P h_{n+1} (n \geqq 0), \sup_n \|h_n\| < \infty\}$$

of all bounded *space-time harmonic functions* of P. Then we have the following well known characterization of asymptotic loss of memory; see [6] or [7] Chaps. 6, 3.4.

(2.3) **Theorem** (Blackwell, Orey et al.). *The following conditions are equivalent:*

(a) $\mu, \nu \in \mathscr{M} \Rightarrow \lim \|\mu P^n - \nu P^n\|_{\mathscr{E}} = 0$,
(b) $\mu, \nu \in \mathscr{M} \Rightarrow P_\mu = P_\nu$ on $\mathscr{A}$,
(c) $\mu \in \mathscr{M} \Rightarrow P_\mu = 0 - 1$ on $\mathscr{A}$,
(d) *Each* $h \in \mathscr{H}$ *is constant.*

(2.4) *Remarks.* 1) One usually proves (a)$\Leftrightarrow$(c)$\Leftrightarrow$(d). But the Markov property implies

$$(2.5) \qquad \|\mu P^n - \nu P^n\|_{\mathscr{E}} = \|P_\mu - P_\nu\|_{\sigma(X_n)} = \|P_\mu - P_\nu\|_{\mathscr{F}_n^*},$$

and so the equivalence of (a) and (b) is clear from (2.2).

2) Let us recall for later purposes the correspondence between space-time harmonic functions $h \in \mathscr{H}$ and functions $\varphi \in \mathscr{A}^b$ which is behind the equivalence (c)$\Leftrightarrow$(d). Let θ_n be the shift map on Ω defined by $(\theta_n \omega)(k) = \omega(n+k)$. For each

$\varphi \in \mathscr{A}^b$ there are functions $\varphi_n \in \mathscr{A}^b$ such that

(2.6) $\varphi = \varphi_n \circ \theta_n, \qquad \varphi_{n+1} \circ \theta = \varphi_n \qquad (n=0, 1, \ldots).$

Defining

(2.7) $h_n^\varphi(x) = E_x[\varphi_n] \qquad (x \in E, n \geqq 0)$

we obtain a space-time harmonic function $h^\varphi \in \mathscr{H}$ such that

(2.8) $\varphi = \lim h_n^\varphi(X_n) \qquad P_\mu$-a.s. $(\mu \in \mathscr{M})$,

and each $h \in \mathscr{H}$ is of that form; see e.g. [7] Chaps. 6, 2.3.

In view of applications to the infinite particle case let us now generalize the equivalence (b)⇔(c)⇔(d) in (2.3). Let $\mathscr{E}_0$ be a sub-σ-field of $\mathscr{E}$ which is *saturated* with respect to some equivalence relation $\sim$ on E; cf. [1]. This means that for any function $f \in \mathscr{E}^b$ we have

(2.9) $f \in \mathscr{E}_0^b \Leftrightarrow [x \sim y \Rightarrow f(x) = f(y)].$

It is shown in [1] that countably generated sub-σ-fields of a standard Borel space, tail fields, invariant fields, symmetric fields are of this type. Let

$$\mathscr{A}_0 = \bigcap_n \bigvee_{k \geqq n} X_k^{-1}(\mathscr{E}_0) \subseteq \mathscr{A}$$

be the corresponding tail field on Ω. Let $\mathscr{A}_1 \subseteq \mathscr{A}$ be some other σ-field of asymptotic events which is *stable* with respect to the representation (2.6), i.e., for $\varphi \in \mathscr{A}_1^b$ and $n \geqq 0$ there exists $\varphi_n \in \mathscr{A}_1^b$ such that $\varphi = \varphi_n \circ \theta_n$. Let

(2.10) $\mathscr{H}_1 = \{h^\varphi \in \mathscr{H} \mid \varphi \in \mathscr{A}_1^b\}$

be the class of bounded space-time harmonic functions associated to $\mathscr{A}_1$ in the sense of (2.7).

(2.11) **Theorem.** *Consider the following conditions:*

(b) $\mu, \nu \in \mathscr{M}$, $\mu = \nu$ *on* $\mathscr{E}_0 \Rightarrow P_\mu = P_\nu$ *on* $\mathscr{A}_1$,
(c) $\mu \in \mathscr{M} \Rightarrow \mathscr{A}_0 = \mathscr{A}_1 \bmod P_\mu$,
(d) *Each* $h \in \mathscr{H}_1$ *is* $\mathscr{E}_0$*-measurable.*

We have (b)⇔(d)⇒(c), *and if* P *is a transition probability on* $(E, \mathscr{E}_0)$ *then all three conditions are equivalent.*

(2.12) *Remark.* For $\mathscr{E}_0 = \{\emptyset, \Omega\}$ and $\mathscr{A}_1 = \mathscr{A}$ (2.11) together with (2.5) reduces to the classical result (2.3). Example (3.15) shows that (c)⇒(d) does not hold in general.

Proof. 1) For $h = (h_n) \in \mathscr{H}_1$ and $n \geqq 0$ we have $h_n(x) = E_x[\varphi_n](x \in E)$ for some $\varphi_n \in \mathscr{A}_1^b$. If $x \sim y$ then $\varepsilon_x = \varepsilon_y$ on $\mathscr{E}_0$ so that (b) implies $h_n(x) = h_n(y)$. Thus (b) implies $h_n \in \mathscr{E}_0^b$ due to (2.9), and this is (d).

2) Take $\varphi\in\mathscr{A}_1^b$ and the associated space-time harmonic function $h^\varphi=(h_n^\varphi)\in\mathscr{H}_1$. For any $\mu\in\mathscr{M}$ we have

$$E_\mu[\varphi]=E_\mu[E_{X_0}[\varphi]]=\int\mu(dx)\,h_0^\varphi(x).$$

Thus (d) implies (b).

3) Let us show (d)⇒(c). Take $\varphi\in\mathscr{A}_1^b$ and the associated $h^\varphi\in\mathscr{H}_1$. We have

$$\varphi=\limsup_n h_n^\varphi(X_n) \qquad P_\mu\text{-a.s. } (\mu\in\mathscr{M})$$

due to (2.8). But under (d) the right side is $\mathscr{A}_0$-measurable, and this yields (c).

4) In order to show (c)⇒(d) we have to assume

(2.13) $\quad A\in\mathscr{E}_0\Rightarrow P(.,A)\in\mathscr{E}_0^b.$

For $h\in\mathscr{H}_1$ and $n\geqq 0$ we have $h_n(x)=E_x[\varphi_n]$ $(x\in E)$ for some $\varphi_n\in\mathscr{A}_1^b$. Take $x,y\in E$. Condition (c) with $\mu=\frac{1}{2}(\varepsilon_x+\varepsilon_y)$ implies that there is some $\psi_n\in\mathscr{A}_0^b$ with $h_n(x)=E_x[\psi_n]$ and $h_n(y)=E_y[\psi_n]$. But (2.13) guarantees that the function $z\to E_z[\psi_n]$ is $\mathscr{E}_0$-measurable. This implies $h_n(x)=h_n(y)$ for $x\sim y$, and so we have shown $h_n\in\mathscr{E}_0^b$ due to (2.9).

3. Local Mixing

Let I be a countable set of *sites* or *particles*, and let $\mathscr{V}$ denote the class of finite subsets $V\subseteq I$. Suppose that each $i\in I$ can assume states in some measurable state space $(E_i,\mathscr{E}_i)$, and introduce the *microscopic state space*

$$(E,\mathscr{E})=\prod_{i\in I}(E_i,\mathscr{E}_i).$$

For each $V\subseteq I$ we denote by $\mathscr{E}_V$ the σ-field on E generated by the projection maps $x\to x(i)$ $(i\in V)$.

In the sequel we fix a transition probability P on this product space $(E,\mathscr{E})$.

(3.1) *Example.* Suppose that for each $i\in I$ we have a transition probability P_i from $(E,\mathscr{E})$ to $(E_i,\mathscr{E}_i)$ which describes the behavior of particle $i\in I$ in reaction to its environment. The corresponding *synchronous interaction* is the kernel P on $(E,\mathscr{E})$ which to each $x\in E$ associates the product measure

$$P(x,.)=\prod_{i\in I}P_i(x,.).$$

See for example [2, 3]. Often each $i\in I$ has some "neighborhood" $N_i\subseteq I$, and P_i is in fact a kernel from $(E,\mathscr{E}_{N_i})$ to $(E_i,\mathscr{E}_i)$. If N_i is finite for each $i\in I$ then the interaction is *local in the passive sense* as defined in (3.9) below. If the set $\{j\mid i\in N_j\}$ is finite for each $i\in I$, then the interaction is *local in the active sense* as defined in (4.2).

Our purpose is to investigate two different versions of loss of memory for the Markov chain associated to P. Both versions will coincide with the classical condition (2.3) if I is finite. In this section we are going to *localize* (2.3). For $V \subseteq I$ we define the σ-fields

$$\mathscr{F}_{n,V} = \bigvee_{k=0}^{n} X_k^{-1}(\mathscr{E}_V), \qquad \mathscr{F}_{n,V}^{*} = \bigvee_{k \geq n} X_k^{-1}(\mathscr{E}_V),$$
$$\mathscr{F}_V = \bigvee_{n} \mathscr{F}_{n,V}, \qquad \mathscr{A}_V = \bigcap_{n} \mathscr{F}_{n,V}^{*}$$

on Ω. The σ-field

$$\mathscr{A}_{\text{loc}} = \bigvee_{V \in \mathscr{V}} \mathscr{A}_V$$

will be called the *local tail field*. Let us also introduce the class $\mathscr{H}_{\text{loc}}$ of all space-time harmonic functions $h \in \mathscr{H}$ which correspond to some $\varphi \in \mathscr{A}_{\text{loc}}^b$ in the sense of (2.7).

(3.2) **Theorem.** *The following conditions are equivalent:*

(a) $\mu, \nu \in \mathscr{M} \Rightarrow \lim \|P_\mu - P_\nu\|_{\mathscr{F}_{n,V}^{*}} = 0 \ (V \in \mathscr{V})$,

(b) $\mu, \nu \in \mathscr{M} \Rightarrow P_\mu = P_\nu$ on $\mathscr{A}_{\text{loc}}$,

(c) $\mu \in \mathscr{M} \Rightarrow P_\mu = 0-1$ on $\mathscr{A}_{\text{loc}}$,

(d) Each $h \in \mathscr{H}_{\text{loc}}$ is constant.

Let us say that P is *locally mixing* if these conditions hold.

(3.3) *Remark.* Condition (a) implies *local convergence* in the sense that

(3.4) $\mu, \nu \in \mathscr{M} \Rightarrow \lim \|\mu P^n - \nu P^n\|_{\mathscr{E}_V} = 0 \qquad (V \in \mathscr{V})$.

But (3.4) does not imply condition (a) as in the classical version (2.3); see example (3.15).

Proof. Recalling (2.2) we see that (a) is equivalent to $P_\mu = P_\nu$ on $\mathscr{A}_V$ for each $V \in \mathscr{V}$, and this is equivalent to (b). The rest follows from (2.11), taking $\mathscr{E}_0 = \{\emptyset, \Omega\}$ and $\mathscr{A}_1 = \mathscr{A}_{\text{loc}}$. We have only to check that $\mathscr{A}_{\text{loc}}$ is stable. But the class

$$\Phi \equiv \{\varphi \in \mathscr{A}_{\text{loc}}^b \mid \varphi = \varphi_n \circ \theta_n, \ \varphi_n \in \mathscr{A}_{\text{loc}}^b (n \geq 0)\}$$

is a linear space closed under pointwise convergence, and it contains the class

$$\Phi_0 \equiv \bigcup_{V \in \mathscr{V}} \mathscr{A}_V^b$$

since each $\mathscr{A}_V$ is stable. This implies $\Phi = \mathscr{A}_{\text{loc}}^b$ by [5] *I*, *T*20.

The usual convergence theorems for interactive Markov chains give criteria for local convergence (3.4), and this is weaker than local mixing (3.2). Let us now look specifically at the convergence theorem of Wasserstein [8], and let us turn it into a criterium for local mixing. For $f \in \mathscr{E}^b$ define the *oscillation coefficients*

$$\rho(f) = \sup\{|f(x) - f(y)| \mid x, y \in E\},$$
$$\rho_i(f) = \sup\{|f(x) - f(y)| \mid x, y \in E, \ x(j) = y(j) \ (j \neq i)\},$$
$$\rho_\infty(f) = \inf_{V \in \mathscr{V}} \sup\{|f(x) - f(y)| \mid x, y \in E, \ x(j) = y(j) \ (j \in V)\}.$$

Then we have

(3.5) $\rho(f) \leqq \sum_{i \in I \cup \{\infty\}} \rho_i(f) \quad (f \in \mathscr{E}^b)$.

(3.6) *Definition.* A matrix $A = (a_{ik})_{i \in I,\, k \in I \cup \{\infty\}}$ will be called a Wasserstein matrix for P if

(3.7) $f \in \mathscr{E}^b \Rightarrow \rho_i(Pf) \leqq \sum_{k \in I \cup \{\infty\}} a_{ik}\, \rho_k(f) \quad (i \in I)$.

(3.8) *Example.* For a synchronous interaction as in (3.1) define

$$a_{ik} = \sup \|P_k(x,.) - P_k(y,.)\|_{\mathscr{E}_k} \quad (i, k \in I)$$

where the supremum is taken over all $x, y \in E$ with $x(j) = y(j)$ $(j \neq i)$, and

$$a_{i\infty} = \inf_{V \in \mathscr{V}} \sup \|P(x,.) - P(y,.)\|_{\mathscr{E}_{I-V}}$$

where the supremum is taken over all $x, y \in E$ with $x(j) = y(j)$ $(j \neq i)$. The arguments in [8] show that this gives us a Wasserstein matrix for P. The coefficients $a_{i\infty}$ are actually irrelevant for this section because here we shall apply (3.6) only to functions f with $\rho_\infty(f) = 0$. For general kernels P there are various ways of constructing a Wasserstein matrix; one such construction is given in [8].

In the sequel we fix a Wasserstein matrix A for P. Let us say that the interaction is *local in the passive sense* if P has the Feller property

(3.9) $f \in C(E) \Rightarrow Pf \in C(E)$

where $C(E) = \{f \in \mathscr{E}^b \mid \rho_\infty(f) = 0\}$ (= the class of all continuous functions with respect to the compact product topology if each E_i is finite). Under (3.9), Wasserstein [8] has shown that we have local convergence (3.4) as soon as

(3.10) $\lim_n \sum_{i \in I} a_{ik}^{(n)} = 0 \quad (k \in I)$

where $(a_{ik}^{(n)})$ is the n-th power of the matrix $(a_{ik})_{i,\, k \in I}$. Note that the Dobrushin condition (1.9) implies both (3.10) and the following condition (3.12).

(3.11) **Theorem.** *Suppose that P satisfies* (3.9) *and*

(3.12) $\sum_{n, i} a_{i,k}^{(n)} < \infty \quad (k \in I)$.

Then we have local mixing in the sense of (3.2).

Proof. We shall verify condition (a) of (3.2). Take $V \in \mathscr{V}$, $\mu, \nu \in \mathscr{M}$, $n, p \geqq 0$ and $\varphi \in \left(\bigvee_{k=n}^{n+p} X_k^{-1}(\mathscr{E}_V)\right)^b$. Due to (2.3) (in the case of increasing σ-fields) it is enough to show that

(3.13) $|E_\mu[\varphi] - E_\nu[\varphi]| \leqq \rho(\varphi)\, c_{n,V}$

where $c_{n,V}$ does not depend on φ and p and satisfies $\lim_n c_{n,V} = 0$.

1) We have $\varphi=\psi\circ\theta_n$ for some $\psi\in\mathscr{F}_V^b$. Let us define $f(x)=E_x[\psi]$ $(x\in E)$. By induction on p the Feller property of P yields $f\in C(E)$, and so we get

$$|E_\mu[\varphi]-E_\nu[\varphi]|=|\int P^n f\,d\mu-\int P^n f\,d\nu|$$
$$\leqq \rho(P^n f)\leqq \sum_i \rho_i(P^n f)\leqq \sum_{i,k} a_{ik}^{(n)}\,\rho_k(f)$$

due to (3.5) and (3.7). But (3.14) below implies

$$\rho_i(f)\leqq \sum_{m=0}^{p}\sum_{k\in V} a_{ik}^{(m)}\,\rho(\varphi).$$

This yields (3.13) with

$$c_{n,V}=\sum_{k\in V}\sum_{m\geqq n}\sum_i a_{i,k}^{(m)},$$

and we have $\lim_n c_{n,V}=0$ due to (3.12).

2) Let us show, by induction on p, that

$$\text{(3.14)}\quad \rho_i(E_x[g(X_0,\ldots,X_p)])\leqq \sum_{m=0}^{p}\sum_k a_{ik}^{(m)}\,\rho_k^{(m)}(g)$$

for each $g\in\left(\prod_{m=0}^{p}\mathscr{E}_V\right)^b$ where $\rho_k^{(m)}(g)$ is the oscillation coefficient of g for space-time coordinate (k,m). For $p=0$ both sides reduce to $\rho_i(g)$. Now take $p\geqq 1$ and suppose that (3.14) holds for $p-1$. Note that

$$E_x[g(X_0,\ldots,X_p)]=E_x[h(X_0,\ldots,X_{p-1})]$$

with

$$h(x_0,\ldots,x_{p-1})=\int P(x_{p-1},dy)\,g(x_0,\ldots,x_{p-1},y),$$
$$\rho_k^{(m)}(h)\leqq\rho_k^{(m)}(g)\qquad(0\leqq m<p-1),$$
$$\rho_k^{(p-1)}(h)\leqq\rho_k^{(p-1)}(g)+\sum_j a_{kj}\,\rho_j^{(p)}(g).$$

Applying (3.14) for $p-1$ and h, we obtain (3.14) for p and g.

(3.15) *Example* (see e.g. [8]). Take $I=Z$, $E_i=\{0,1\}$ $(i\in I)$ and the synchronous interaction $P(x,.)=\prod_{i\in I}P_i(x,.)$ where P_i is defined by

$$P_i(x,\{1\})=\begin{cases}p_i & \text{if } x(i+1)=1\\ 1-p_i & \text{if } x(i+1)=0\end{cases}$$

with some $p_i\in(\frac{1}{2},1)$. It is easy to check that for each $r\in[0,1]$ the product measure ν_r on E with

$$\nu_r[x(i)=1]=\tfrac{1}{2}+\prod_{j\geqq i}(2p_j-1)(r-\tfrac{1}{2})$$

is invariant for P. These measures all coincide with the coin tossing measure $\nu_{\frac{1}{2}}$ if and only if

$$(3.16) \quad \prod_{j \geqq 0} (2p_j - 1) = 0,$$

and this is just the Wasserstein condition (3.10). Let us now show that this is not enough for local mixing. Let μ_0 resp. μ_1 be the Dirac measure on the state $x \equiv 0$ resp. $x \equiv 1$. The 0-coordinates $(X_{n,0})$ of the chain (X_n) are independent both under P_{μ_0} and P_{μ_1}. Define

$$a_n = P_{\mu_1}[X_{n,0} = 1], \qquad b_n = P_{\mu_0}[X_{n,0} = 1].$$

Then we have

$$a_n - b_n = \prod_{0 \leqq k < n} (2p_k - 1) \qquad (n \geqq 1),$$

and so (3.16) holds if and only if $a_n - b_n$ goes to 0. Now construct (p_n) such that $a_n - b_n$ goes to 0 so slowly that P_{μ_0} and P_{μ_1} are singular on the σ-field $\mathscr{F}^*_{1,\{0\}}$ generated by the process $(X_{n,0})_{n \geqq 1}$ (recall Kakutani's criterium for the singularity of product measures). Then P_{μ_0} and P_{μ_1} must also be singular to each other on $\mathscr{A}_{\{0\}}$, and this contradicts local mixing.

4. Sufficiency of the Boundary Process

Let us introduce the *boundary* $(E, \hat{\mathscr{E}})$ where

$$\hat{\mathscr{E}} = \bigcap_{V \in \mathscr{V}} \mathscr{E}_{I-V}$$

is the spatial tail field on E, and the corresponding σ-fields

$$\hat{\mathscr{F}}_n = \bigvee_{k=0}^{n} X_k^{-1}(\hat{\mathscr{E}}), \qquad \hat{\mathscr{F}}_n^* = \bigvee_{k=n}^{\infty} X_k^{-1}(\hat{\mathscr{E}})$$

$$\hat{\mathscr{F}} = \bigvee_n \hat{\mathscr{F}}_n, \qquad \hat{\mathscr{A}} = \bigcap_n \hat{\mathscr{F}}_n^*$$

on Ω. Each coordinate map $\omega \to \omega(n)$ may be viewed as a measurable map from $(\Omega, \hat{\mathscr{F}}_n)$ to $(E, \hat{\mathscr{E}})$, and in that case we denote it by $\hat{X}_n$. $(\hat{X}_n)_{n \geqq 0}$, as an adapted process over $(\Omega, \hat{\mathscr{F}}, (\hat{\mathscr{F}}_n), (P_\mu))$ with values in $(E, \hat{\mathscr{E}})$, may be called the *boundary process* associated to the underlying microscopic process $(X_n)_{n \geqq 0}$ with values in $(E, \mathscr{E})$. Note that the boundary process is again a Markov chain if

(4.1) P is a transition probability on $(E, \hat{\mathscr{E}})$,

i.e., if $P(., A)$ is $\hat{\mathscr{E}}$-measurable for each $A \in \hat{\mathscr{E}}$.

(4.2) *Remarks.* 1) Condition (4.1) means that the interaction is *local in the active sense:* no individual particle can influence the boundary situation in one step. This implies that it also won't be able to influence events in the boundary tail field $\hat{\mathscr{A}}$. Nevertheless it may influence events in the σ-field

(4.3) $\mathscr{A}_\infty = \bigcap_{V \in \mathscr{V}} \mathscr{A}_{I-V}$;

see example (4.8). Such a long run effect is excluded by the conditions of the following theorem.

2) Recall the coefficients $a_{i\infty}$ defined in (3.8) and note that the definition makes sense for a general P. If $a_{i\infty}=0$ for each $i \in I$ then the boundary process has the Markov property (4.1).

(4.4) **Theorem.** *Consider the following conditions:*

(a) $\mu, \nu \in \mathscr{M}$, $\mu=\nu$ on $\hat{\mathscr{E}} \Rightarrow \lim_n \|\mu P^n - \nu P^n\|_{\mathscr{E}} = 0$,

(b) $\mu, \nu \in \mathscr{M}$, $\mu=\nu$ on $\hat{\mathscr{E}} \Rightarrow P_\mu = P_\nu$ on $\mathscr{A}$,

(c) $\mu \in \mathscr{M} \Rightarrow \mathscr{A} = \hat{\mathscr{A}} \bmod P_\mu$,

(d) *Each* $h \in \mathscr{H}$ *is* $\hat{\mathscr{E}}$*-measurable.*

We have (a)$\Leftrightarrow$(b)$\Leftrightarrow$(d)$\Rightarrow$(c), *and under* (4.1) *all four conditions are equivalent.*

Proof. The equivalence of (a) and (b) is clear from (2.2) and (2.5). The rest follows from (2.11), taking $\mathscr{A}_1 = \mathscr{A}$ and $\mathscr{E}_0 = \hat{\mathscr{E}}$.

(4.5) *Remark.* Condition (d) implies

$$E_\mu[\varphi \mid \mathscr{F}_n] = E_\mu[\varphi \mid \hat{X}_n] \qquad (\varphi \in \mathscr{A}^b,\ \mu \in \mathscr{M}),$$

i.e., it is sufficient to observe the boundary process in order to make the best predictions for the asymptotic behavior of the underlying microscopic process. Let us say that the *boundary process is sufficient* if all conditions hold.

Let us now show that the boundary process is indeed sufficient if a dual version of the Wasserstein condition (3.10) holds.

(4.6) **Theorem.** *Let A be a Wasserstein matrix for P with $a_{i\infty}=0$ for each $i \in I$. If*

(4.7) $\lim_n \sum_{k \in I} a_{ik}^{(n)} = 0 \qquad (i \in I)$

then all conditions in (4.4) *are satisfied.*

Proof. Let us verify condition (d). Iterating (3.7) we obtain

$$\rho_i(P^n f) \leqq \sum_{k \in I} a_{ik}^{(n)} \rho_k(f) \qquad (i \in I)$$

for any $f \in \mathscr{E}^b$. Now take $h=(h_n) \in \mathscr{H}$. For $i \in I$ and $m, n \geqq 0$ we get

$$\begin{aligned}\rho_i(h_m) = \rho_i(p^n h_{m+n}) &\leqq \sum_{k \in I} a_{ik}^{(n)} \rho_k(h_{m+n}) \\ &\leqq \sum_{k \in I} a_{ik}^{(n)} 2 \sup_p \|h_p\|.\end{aligned}$$

Thus (4.7) implies $\rho_i(h_m)=0$ $(i \in I)$, hence $h_m \in \hat{\mathscr{E}}^b$.

(4.8) *Examples.* 1) In example (3.15) our condition (4.7) is equivalent to

(4.9) $\prod_{k \leq 0} (2p_k - 1) = 0,$

and this is also necessary for property (d) of (4.4): For $c \in [0, 1]$ define

$$c_n = \tfrac{1}{2} + \prod_{k<n} (2p_k - 1)(c - \tfrac{1}{2})$$

and $h_n(x) = c_{-n}$ resp. $= 1 - c_{-n}$ if $x(-n) = 1$ resp. $= 0$. Then $h = (h_n)$ is in $\mathscr{H}$, but not $\hat{\mathscr{E}}$-measurable if $c \neq \frac{1}{2}$ and if (4.9) does not hold. This shows, by the way, that the σ-field $\mathscr{A}_\infty$ defined in (4.3) does not coincide with $\hat{\mathscr{A}}$ since $\varphi = \lim h_n(X_n)$ is $\mathscr{A}_\infty$-but not $\hat{\mathscr{A}}$-measurable. The criteria (4.9) and (3.16) show that local mixing and sufficiency of the boundary process are independent of one another: each can appear with or without the other one.

2) Let us show that without (4.1) our condition (c) in (4.4) is not enough to guarantee condition (d). Consider the synchronous interaction with $I = \{0, 1, \ldots\}$, $E_i = \{0, 1\}$ and $P_i(x, .) = \mathscr{E}_{x(0)}$ for $i \neq 0$, $E_0 = \{0, \frac{1}{2}, 1\}$ and

$$P_0(x, .) = \begin{cases} \mathscr{E}_{x(0)} & \text{if } x(0) \in \{0, 1\} \\ \frac{1}{3}(\varepsilon_0 + \varepsilon_{\frac{1}{2}} + \varepsilon_1) & \text{if } x(0) = \frac{1}{2}. \end{cases}$$

Then $P(x, .) = \prod_{i \in I} P_i(x, .)$ depends only on the 0-coordinate $x(0)$, and this implies $h_n = P h_{n+1} \in \mathscr{E}^b_{\{0\}}$ for each $h = (h_n) \in \mathscr{H}$. For instance, $h = (h_n)$ with $h_n(x) = x(0)$ is in $\mathscr{H}$ but not $\hat{\mathscr{E}}$-measurable so that (d) does not hold. Now let us verify condition (c). For any $x \in E$ we have

$$\rho(y) = \lim \frac{1}{n} \sum_{k=0}^{n-1} y(k) = x(0)$$

for $P(x, .)$-almost all $y \in E$. Take $\varphi \in \mathscr{A}^b$. The associated $h^\varphi \in \mathscr{H}$ is of the form $h_n(x) = f_n(x(0))$, and so we have

$$\varphi = \lim h_n(X_n) = \lim f_n(\rho(X_{n+1})) \qquad P_\mu\text{-a.s.}$$

for any $\mu \in \mathscr{M}$. But the right side is $\hat{\mathscr{A}}$-measurable.

Let us finally mention some further applications of (2.11). Taking $\mathscr{E}_0 = \hat{\mathscr{E}}$ and $\mathscr{A}_1 = \mathscr{A}_{\text{loc}}$ so that condition (b) becomes

(4.10) $\mu, \nu \in \mathscr{M}, \quad \mu = \nu \text{ on } \hat{\mathscr{E}} \Rightarrow P_\mu = P_\nu \text{ on } \mathscr{A}_{\text{loc}},$

we obtain a theorem which characterizes *local sufficiency of the boundary process*. A modification of the proof of (3.11) shows that (4.10) does hold if the Wasserstein matrix satisfies

(4.11) $\sum_n a_{ik}^{(n)} < \infty \qquad (i, k \in I),$

assuming either the Feller property (3.9) or the dual condition $a_{i\infty} = 0$ for each $i \in I$.

We can also replace in these three variants the σ-fields $\mathscr{A}$, $\mathscr{A}_{\text{loc}}$, $\mathscr{A}$ by the corresponding σ-fields $\mathscr{I}$, $\mathscr{I}_{\text{loc}}$, $\mathscr{I}$ of *invariant* events, and the classes $\mathscr{H}$, $\mathscr{H}_{\text{loc}}$, $\mathscr{H}$

by the corresponding classes of *harmonic* functions. Taking for example $\mathscr{E}_0 = \hat{\mathscr{E}}$ and $\mathscr{A}_1 = \mathscr{I}_{\text{loc}}$ in (2.11), we get a characterization of the condition

(4.12) $\mu, \nu \in \mathscr{M}, \quad \mu = \nu$ on $\hat{\mathscr{E}} \Rightarrow P_\mu = P_\nu$ on $\mathscr{I}_{\text{loc}}$.

This last (and weakest) condition is related to the following question discussed by Dawson [2]. Suppose that $\mu_i \in \mathscr{M}$ is *invariant* under P for $i = 1, 2$. Under which conditions are we able to conclude

(4.13) $\mu_1 = \mu_2$ on $\hat{\mathscr{E}} \Rightarrow \mu_1 = \mu_2$?

Since μ_i is invariant, the ergodic theorem yields

$$\int f \, d\mu_i = E_{\mu_i}\left[\lim \frac{1}{n} \sum_{k=0}^{n-1} f(X_k)\right] \qquad (f \in \mathscr{E}^b).$$

For $V \in \mathscr{V}$ and $f \in \mathscr{E}_V^b$ the integrand is $\mathscr{I}_{\text{loc}}$-measurable. Thus condition (4.12) allows to conclude $\int f \, d\mu_1 = \int f \, d\mu_2$, and this implies $\mu_1 = \mu_2$. The argument in [2] seems to suggest (in identifying $\hat{\mathscr{F}}$ and $\bigcap_{V \in \mathscr{V}} \mathscr{F}_{I-V}$) that even condition (b) of (4.4) is automatic for synchronous interactions of a type including example (3.15). We have seen that this is not so, and so the argument does need some additional assumption.

References

1. Blackwell, D., Dubins, L.E.: On existence and non-existence of proper, regular, conditional distributions. Ann. Probability **3**, 741–752 (1975)
2. Dawson, D.A.: Information flow in discrete Markov systems. J. Appl. Probability **10**, 63–83 (1973)
3. Dobrushin, R.L. et al. (Ed).: Locally interacting systems and their application in biology. Lecture Notes in Mathematics **653**. Berlin-Heidelberg-New York: Springer 1978
4. Goldstein, S.: Maximal coupling. Z. Wahrscheinlichkeitstheorie verw. Gebiete **46**, 193–204 (1979)
5. Meyer, P.A.: Probability and Potentials. Waltham: Blaisdell 1966
6. Orey, S.: Limit theorems for Markov chain transition probabilities. New York: Van Nostrand 1971
7. Revuz, D.: Markov Chains. Amsterdam: North Holland 1975
8. Wasserstein, L.N.: Markov processes over denumerable products of spaces describing large systems of automata. Problemy Peredači Informacii, **5**, no. 3, 64–72 (1969)

Received February 2, 1979; in revised form August 28, 1979