

SYNCHRONOUS AND ASYNCHRONOUS REVERSIBLE MARKOV SYSTEMS⁽¹⁾

BY
D. A. DAWSON

ABSTRACT. The relationships between synchronous and asynchronous reversible Markov systems are investigated. It is shown that the invariant measure of such systems is a second order Markov random field. The conditions under which the invariant measure is a first order Markov random field are obtained.

1. Introduction. It has been shown by F. Spitzer [4] that a Markov system with continuous time parameter which is time reversible has an invariant measure which corresponds to a Markov random field. We introduce the concept of an asynchronous discrete time Markov system which has the same property. In this paper these asynchronous systems are compared and contrasted with the class of synchronous Markov systems which were introduced by the author in [2].

2. Basic terminology. Let A be a countable set, S a finite set and $\Gamma = S^A$. Γ will serve as the state space for the class of Markov chains to be discussed. Let \mathcal{G} be the σ -algebra of subsets of Γ generated by the coordinate functions and for $B \subset A$, let \mathcal{G}_B be the σ -algebra generated by the coordinates in B . A mapping $P: \Gamma \otimes \mathcal{G} \rightarrow [0, 1]$ is a *probability transition kernel* if for each $\gamma \in \Gamma$, $P(\gamma, \cdot)$ is a probability measure on \mathcal{G} and for every $G \in \mathcal{G}$, $P(\cdot, G)$ is a \mathcal{G} -measurable function.

To each $\alpha \in A$, there is an associated finite subset of A denoted by $N(\alpha)$ whose elements are called the *nearest neighbours* of α . We assume that the sets $N(\cdot)$ satisfy the condition:

$$(2.1) \quad \beta \in N(\alpha) \text{ if and only if } \alpha \in N(\beta).$$

For each α , a mapping

$$q_\alpha: S^{N(\alpha)} \otimes \mathcal{G}_{\{\alpha\}} \rightarrow [0, 1]$$

is a *local probability transition kernel* if for each $\xi \in S^{N(\alpha)}$, $q_\alpha(\xi, \cdot)$ is a probability mass function on $\mathcal{G}_{\{\alpha\}}$. $q_\alpha(\cdot, \cdot)$ is said to be *strictly stochastic* if for each $\xi \in S^{N(\alpha)}$ and $s \in S$, $q_\alpha(\xi, s) > 0$.

The kernel $P(\cdot, \cdot)$ is said to be the *synchronous kernel* associated with the collection $\{q_\alpha(\cdot, \cdot): \alpha \in A\}$ if for every finite subset $B \subset A$, $y \in \Gamma$ and $\{x_\alpha \in S: \alpha \in B\}$

$$(2.2) \quad P(y, \{\gamma: \gamma(\alpha) = x_\alpha \text{ for each } \alpha \in B\}) = \prod_{\alpha \in B} q_\alpha(\pi_{N(\alpha)} y, x_\alpha)$$

Received by the editors October 16, 1972 and, in revised form, September 5, 1973.

⁽¹⁾ This research was supported by the National Research Council of Canada.

where for $C \subset A$ and $y \in \Gamma$, $\pi_C y$ is the restriction of the function y to the domain C .

Let $\Gamma' \equiv \{0, 1\}^A$ and let \mathcal{G}' be the σ -algebra of subsets of Γ' generated by the coordinate functions. Let Q be a probability measure on \mathcal{G}' which satisfies the condition:

(2.3) for any two finite subsets D_1 and D_2 of A with $D_1 \cap D_2 = \emptyset$,

$$Q^*(D_1; D_2) \equiv Q(\{\gamma: \gamma(\alpha) = 1 \text{ for } \alpha \in D_1 \text{ and } \gamma(\alpha) = 0 \text{ for } \alpha \in D_2\}) > 0$$

if and only if the set D_1 contains no pair of distinct nearest neighbours of A , that is, α and β with $\alpha \in N(\beta)$ and $\alpha \neq \beta$.

For example let Q_1 be the measure which makes the coordinates independent and such that for each $\alpha \in A$, $Q_1(\{\gamma: \gamma(\alpha)=1\})=1/2$. Then let $F: \Gamma \rightarrow \Gamma$ be the mapping defined by

$$F(\gamma)(\alpha) = \gamma(\alpha) \text{ if } \gamma(\beta) = 0 \text{ for all } \beta \in N(\alpha) - \{\alpha\} \\ = 0 \text{ if } \gamma(\beta) = 1 \text{ for some } \beta \in N(\alpha) - \{\alpha\}.$$

Then if $Q_2(G) \equiv Q_1(F^{-1}(G))$, it is easy to verify that Q_2 satisfies condition (2.3).

Given a measure Q on \mathcal{G}' which satisfies condition (2.3) and a collection $\{q_\alpha: \alpha \in A\}$ of local probability transition kernels the associated *asynchronous probability transition kernel* $P(\cdot, \cdot)$ on Γ is defined as follows. For each finite subset $B \subset A$, $y \in \Gamma$ and $\{x_\alpha \in S: \alpha \in B\}$,

$$(2.4) \quad P(y, \{\gamma: \gamma(\alpha) = x_\alpha \text{ for each } \alpha \in B\}) \\ = \sum_{D \subset B} \left[Q^*(D, B-D) \prod_{\alpha \in D} q_\alpha(\pi_{N(\alpha)} y, x_\alpha) \prod_{\alpha \in B-D} \delta(y_\alpha, x_\alpha) \right]$$

where the summation runs over all subsets D of B and

$$\delta(y_\alpha, x_\alpha) \equiv 1 \text{ if } y_\alpha = x_\alpha \\ \equiv 0 \text{ if } y_\alpha \neq x_\alpha.$$

Note that the crucial property of asynchronous kernels is that simultaneous transitions at neighbouring sites are forbidden.

A probability measure μ on \mathcal{G} is said to be an *invariant measure* for the kernel $P(\cdot, \cdot)$ if for each $G \in \mathcal{G}$,

$$(2.5) \quad \mu(G) = \int P(\xi, G) \mu(d\xi).$$

A probability measure μ on G is said to be *locally invariant* for the collection $\{q_\alpha: \alpha \in A\}$ if for each α , it is invariant for the kernel defined by

$$P_\alpha(y, \{\gamma: \gamma(\beta) = x_\beta \text{ for each } \beta \in B\}) = q_\alpha(\pi_{N(\alpha)} y, x_\alpha) \prod_{\substack{\beta \neq \alpha \\ \beta \in B}} \delta(x_\beta, y_\beta) \text{ if } \alpha \in B \\ = \prod_{\beta \in B} \delta(x_\beta, y_\beta) \text{ if } \alpha \notin B.$$

In other words $P_\alpha(\cdot, \cdot)$ is the kernel in which a local transition is allowed at α but nowhere else.

3. **Reversible systems.** Given a finite subset $B \subset A$ we denote by ∂B the set of elements in $A - B$ which are nearest neighbours of elements of B . We define a sequence of boundaries $\partial_r B$ as follows:

$$\partial_{r+1} B = \partial_r B \cup \partial(B \cup \partial_r B).$$

In other words $\partial_r B$ is the set of points in $A - B$ whose "distance" from B is less than or equal to r .

A probability measure μ on \mathcal{G} is said to represent an *rth order Markov random field* if

(3.1) for any finite $B \subset A$ and $G \in \mathcal{G}_B$, $\mu(G) > 0$,

and

(3.2) if $B \equiv \{\alpha\}$ where $\alpha \in A$ and $G \in \mathcal{G}_B$, then $\mu(G \mid \mathcal{G}_{B^c}) = \mu(G \mid \mathcal{G}_{\partial_r B})$

with probability one.

Let $\Omega \equiv \Gamma^{Z^+}$ where $Z^+ \equiv \{0, 1, 2, 3, \dots\}$, let \mathcal{F} denote the product σ -algebra of subsets of Ω and for $n=0, 1, 2, \dots$ let $X_n(\omega)$ denote the n th coordinate of ω , that is, the state of the system at time n . Given an invariant initial probability measure μ on Γ and the kernel $P(\cdot, \cdot)$ we can construct a measure P on (Ω, \mathcal{F}) such that $\{X_n\}$ forms a time homogeneous Markov chain with transition probabilities given by $P(\cdot, \cdot)$. If $\{X_n\}$ is a Markov chain with invariant initial measure, then the Markov chain reversed in time is also a time homogeneous Markov chain. The Markov chain is said to be *reversible* if for $G \in \mathcal{G}$

(3.3) $P(X_0 \in G \mid \sigma(X_1)) = P(X_1 \in G)$ with probability one.

PROPOSITION 3.1. *Let $\{X_n: n \geq 0\}$ be a strictly stochastic Markov system (either synchronous or asynchronous) associated with a family of local kernels $\{q_\alpha(\cdot, \cdot): \alpha \in A\}$ and let μ be a measure which is locally invariant. Then μ represents a first order Markov random field and μ is invariant for the asynchronous system.*

Proof. Consider a fixed site $\alpha_0 \in A$ and fix the boundary conditions on $\partial(\{\alpha_0\})$. Since the measure is locally invariant, $\mu(\cdot \mid \mathcal{G}_{\{\alpha_0\}^c})$ is the invariant measure for the Markov chain obtained by allowing a sequence of transitions at the site α_0 . But since the system is strictly stochastic, $\mu(\cdot \mid \mathcal{G}_{\{\alpha_0\}^c})$ is uniquely determined and depends only on the boundary conditions on $\partial(\{\alpha_0\})$.

We now show that μ is invariant for the asynchronous system. Let B be a finite subset of A and let $G \equiv \{\gamma: \gamma(\alpha) = x_\alpha \text{ for each } \alpha \in B\} \in \mathcal{G}_B$. Then

$$\int P(y, G) \mu(dy) = \sum_{D \subset B} Q^*(D, B - D) \int \prod_{\alpha \in D} q_\alpha(\pi_{N(\alpha)} y, x_\alpha) \prod_{\alpha \in B - D} \delta(y_\alpha, x_\alpha) \mu(dy).$$

Hence it suffices to show that if $Q^*(D, B-D) > 0$, then

$$(3.4) \quad \int \prod_{\alpha \in D} q_\alpha(\pi_{N(\alpha)}y, x_\alpha) \prod_{\alpha \in B-D} \delta(y_\alpha, x_\alpha) \mu(dy | \mathcal{G}_{B^c}) = \mu(G | \mathcal{G}_{B^c}) \quad \text{a.s.}$$

But since $D = \{\alpha_1, \dots, \alpha_D\}$ contains no pair of nearest neighbours if $Q^*(D, B-D) > 0$,

$$\int \prod_{\alpha \in D} q_\alpha(\pi_{N(\alpha)}y, x_\alpha) \prod_{\alpha \in B-D} \delta(y_\alpha, x_\alpha) \mu(dy) = P_{\alpha_1} \circ \dots \circ P_{\alpha_D} \mu(G | \mathcal{G}_{B^c})$$

where

$$P_{\alpha_i} \mu(\cdot | \mathcal{G}_{B^c}) = \int P_{\alpha_i}(y, \cdot) \mu(dy | \mathcal{G}_{B^c}).$$

The result then follows by the local invariances.

The next result is a discrete time analogue of a result of F. Spitzer [4] for continuous time systems.

PROPOSITION 3.2. *Let $\{X_n\}$ be a Markov chain on the state space Γ whose probability transition kernel is the asynchronous kernel associated to a strictly stochastic family of local kernels $\{q_\alpha(\cdot, \cdot) : \alpha \in A\}$. Assume that μ is an invariant probability measure and assume that the corresponding Markov system is reversible. Then μ is a first order Markov random field and μ is locally invariant.*

Proof. Since the $q_\alpha(\cdot, \cdot)$ are strictly stochastic it is easy to verify that if G is a finite cylinder set, then $\mu(G) > 0$. Let $B_1 \equiv \{\alpha_0\}$ and for $j \geq 2$, let $B_j \equiv B_{j-1} \cup \partial B_{j-1}$. Let $\gamma_1, \gamma_2 \in \Gamma$ be such that

$$\pi_{B_1}(\gamma_1) \neq \pi_{B_1}(\gamma_2).$$

Let the events a_1, a_2, b, c be defined as follows

$$(3.4) \quad \begin{aligned} a_1 &\equiv \{\pi_{B_1} X_0 = \pi_{B_1} \gamma_1\} \\ a_2 &= \{\pi_{B_1} X_0 = \pi_{B_1} \gamma_2\} \\ b &\equiv \{\pi_{B_2-B_1} X_0 = \pi_{B_2-B_1} \gamma_1\} \\ c &\equiv \{\pi_{B_k-B_2} X_0 = \pi_{B_k-B_2} \gamma_1\} \text{ for some finite } k. \end{aligned}$$

Let a_1^+, a_2^+, b^+, c^+ be defined as in (3.4) but with X_0 replaced by X_1 .

By the definition of conditional probability and (2.4), we have

$$(3.5) \quad P(a_1, b, c | a_2^+, b^+) = \mu(a_1, b, c) q(a_1, b; a_2) Q^*(B_1, B_2 - B_1) / \mu(a_2, b).$$

The reversibility assumption implies that

$$(3.6) \quad \begin{aligned} P(a_1, b, c | a_2^+, b^+) &= P(a_1^+, b^+, c^+ | a_2, b) \\ &= Q^*(B_1, B_2 - B_1) q(a_2, b; a_1) P(c^+ | a_2, b; B_1) \end{aligned}$$

where

$$P(c^+ | a_2, b; B_1) = \int P_{B_1}(y, c) \mu(dy | a_2, b)$$

and $P_{B_1}(y, c)$ is defined as in (2.4) except that $Q^*(\cdot, \cdot)$ is replaced by

$$Q_{B_1}^*(D_1, D_2) \equiv Q(\{\gamma: \gamma(\alpha) = 1 \text{ for } \alpha \in D_1, \gamma(\alpha) = 0 \text{ for } \alpha \in D_2\} \mid \gamma\{\gamma(\alpha) = 1 \text{ for } \alpha \in B_1\}).$$

From (3.5) and (3.6)

$$(3.7) \quad P(c^+ \mid a_2, b; B_1) = \left[\frac{\mu(a_1, b)q(a_1, b; a_2)}{\mu(a_2, b)q(a_2, b; a_1)} \right] \mu(c \mid a_1, b).$$

But since $\sum_c P(c^+ \mid a_2, b; B_1) = \sum_c \mu(c \mid a_1, b) = 1$ where the summation is over all configurations over $B_k - B_2$, it follows that

$$(3.8) \quad \mu(a_1, b)q(a_1, b; a_2) / \mu(a_2, b)q(a_2, b; a_1) = 1.$$

We now consider two cases.

Case 1. $|S| > 2$ where $|S| = \text{number of elements in } S$. If $a_3 \equiv \{\pi_{B_1} X_0 = \pi_{B_1} \gamma_3\}$ where $\pi_{B_1}(\gamma_3) \neq \pi_{B_1}(\gamma_2)$ and $\pi_{B_1}(\gamma_3) \neq \pi_{B_1}(\gamma_1)$, then from (3.7) and (3.8),

$$(3.9) \quad \mu(c \mid a_1, b) = \mu(c \mid a_3, b) = P(c^+ \mid a_2, b; B_1).$$

Hence conditional on $\mathcal{G}_{B_2 - B_1}, \mathcal{G}_{B_k - B_2}$ and \mathcal{G}_{B_1} are conditionally independent. Since this is true for all k , μ represents a first order Markov random field.

Case 2. $|S| = 2, S = \{0, 1\}$.

By (3.7)

$$\begin{aligned} \mu(c \mid a_1, b) &= P(c^+ \mid a_2, b; B_1) = P(c^+, b^+ \mid a_2, b; B_1) = \frac{P(c^+, b^+, a_2, b \mid B_1)}{\mu(a_2, b)} \\ &= \int_{a_2 b} P_{B_1}(y, c) \mu(dy) / \mu(a_2, b) \\ &= \int_{cb} P_{B_1}(y, a_2) \mu(dy) / \mu(a_2, b) \text{ by reversibility} \\ &= P(c, b, a_2^+, b^+ \mid B_1) / \mu(a_2, b) \\ &= [\mu(a_1, b, c)q(a_1, b; a_2) + \mu(a_2, b, c)q(a_2, b; a_2)] / \mu(a_2, b) \text{ since } a_1 \cup a_2 = \Omega. \end{aligned}$$

Hence

$$\begin{aligned} \mu(c \mid a_1, b) &= \mu(c \mid a_2, b)q(a_2, b; a_2) + \mu(c \mid a_1, b)\mu(a_1, b)q(a_1, b; a_2) / \mu(a_2, b) \\ &= \mu(c \mid a_2, b)q(a_2, b; a_2) + \mu(c \mid a_1, b)q(a_2, b; a_1) \text{ by (3.8)}. \end{aligned}$$

But since $q(a_2, b; a_1) = 1 - q(a_2, b; a_2)$, this implies that

$$\mu(c \mid a_1, b) = \mu(c \mid a_2, b)$$

and the proof is completed as in Case 1.

Since μ represents a first order Markov random field, the local invariance of μ follows immediately from (3.8).

PROPOSITION 3.3. *Let μ represent a first order Markov random field. Then there exists an asynchronous Markov system having μ as invariant measure and such that the corresponding system is reversible.*

Proof. Let the local transition kernels $q(a_1, b; a_2)$ where a_1, a_2, b are defined as in Proposition 3.2 be defined as follows. If $a_1 \neq a_2$, let

$$(3.8) \quad \begin{aligned} q(a_1, b; a_2) &= \frac{1}{K} \frac{\mu(a_2 | b)}{\mu(a_1 | b)} \quad \text{if } \frac{\mu(a_2 | b)}{\mu(a_1 | b)} < 1 \\ &= \frac{1}{K} \quad \text{if } \frac{\mu(a_2 | b)}{\mu(a_1 | b)} \geq 1 \end{aligned}$$

and let

$$(3.9) \quad q(a_1, b; a_1) = 1 - \sum_{a_2 \neq a_1} q(a_1, b; a_2)$$

where K is the number of elements in the set S . For each b , it is easy to verify that $q(\cdot, b; \cdot)$ is a stochastic matrix and that

$$(3.10) \quad \mu(a_1 | b)q(a_1, b; a_2) = \mu(a_2 | b)q(a_2, b; a_1).$$

This implies that μ is locally invariant and hence μ is also invariant for the asynchronous system and the system is reversible.

COROLLARY. *Let μ represent a first order Markov random field. Given a finite set $B \subset A$ and $G \in \mathcal{G}_B$,*

$$(3.11) \quad \mu(G | \mathcal{G}_{B^c}) = \mu(G | \mathcal{G}_{\partial B})$$

with probability one.

Before stating the next result we must introduce a weaker form of a technical condition due to R. L. Dobrushin [3]. The measure μ on (Γ, \mathcal{G}) is said to satisfy the *D-condition* if there exists a version of the conditional distribution $\mu(\cdot | \mathcal{G}_{B_1}^c)$ such that

$$(3.12) \quad \text{Max}_{a \in \mathcal{G}_{B_1}} \text{Sup}_k |\mu(a | \pi_{B_1^c}^k \gamma) - \mu(a | \pi_{B_1^c}^k \gamma')| \equiv r(k) \rightarrow 0$$

as $k \rightarrow \infty$ where Sup_k is the supremum taken over all γ and γ' such that $\pi_{B_k - B_1} \gamma = \pi_{B_k - B_1} \gamma'$.

PROPOSITION 3.4. *Let $\{X_n\}$ be the strictly stochastic synchronous Markov system associated with the family of local transition kernels $\{q_\alpha(\cdot, \cdot); \alpha \in A\}$. Assume that μ is an invariant measure satisfying the D-condition and such that the corresponding Markov system is reversible. Then μ represents a second order Markov random field.*

Proof. Let $B_i, i \geq 1$ be defined as in the proof of Proposition 3.2. Let

$$\begin{aligned}
 a_1 &\equiv \{\pi_{B_1} X_0 = \gamma_{B_1}\} \text{ where } \gamma_{B_1} \in \pi_{B_1} \Gamma \\
 a_2 &\equiv \{\pi_{B_1} X_0 = \gamma'_{B_1}\} \text{ where } \gamma'_{B_1} \in \pi_{B_1} \Gamma \text{ and } \gamma_{B_1} \neq \gamma'_{B_1} \\
 b &\equiv \{\pi_{B_2-B_1} X_0 = \gamma_{B_2-B_1}\} \text{ where } \gamma_{B_2-B_1} \in \pi_{B_2-B_1} \Gamma \\
 c &\equiv \{\pi_{B_3-B_2} X_0 = \gamma_{B_3-B_2}\} \text{ where } \gamma_{B_3-B_2} \in \pi_{B_3-B_2} \Gamma \\
 d &\equiv \{\pi_{B_4-B_3} X_0 = \gamma_{B_4-B_3}\} \text{ where } \gamma_{B_4-B_3} \in \pi_{B_4-B_3} \Gamma
 \end{aligned}$$

and for some $k > 4$, let

$$e \equiv \{\pi_{B_k-B_4} X_0 = \gamma_{B_k-B_4}\}$$

where $\gamma_{B_k-B_4} \in \Gamma_{B_k-B_4}$ and

$$f \equiv \{\pi_{B_{k+1}-B_k} X_0 = \gamma_{B_{k+1}-B_k}\}$$

where

$$\gamma_{B_{k+1}-B_k} \in \pi_{B_{k+1}-B_k} \Gamma.$$

Let $a_1^+, a_2^+, b^+, c^+, d^+, e^+$ be defined in the same way but with X_0 replaced by X_1 .

Given a finite set $B \subset A$, we define

$q: S^B \cup \partial B \times S^B \rightarrow [0, 1]$ as follows:

$$q(x; y) = \prod_{a \in B} q_a(\pi_{N(a)} x; y_a).$$

By definition

$$\begin{aligned}
 (3.13) \quad &P(a_1, b, c, d, e \mid a_2^+, b^+, c^+, d^+, e^+) \\
 &= [\mu(a_1, b, c, d, e) / \mu(a_2, b, c, d, e)] \cdot q(a_1, b; a_2) \\
 &\quad \times q(a_1, b, c; b) q(b, c, d; c) P(d^+, e^+ \mid a_1, b, c, d, e).
 \end{aligned}$$

Because the system is reversible, we also have

$$\begin{aligned}
 (3.14) \quad &P(a_1, b, c, d, e \mid a_2^+, b^+, c^+, d^+, e^+) \\
 &= q(a_2, b; a_1) q(a_2, b, c; b) q(b, c, d; c) P(d^+ e^+ \mid a_2, b, c, d, e).
 \end{aligned}$$

From (3.13) and (3.14) we have

$$(3.15) \quad \frac{P(a_2, b, c, d, e, d^+, e^+)}{P(a_1, b, c, d, e, d^+, e^+)} = \frac{q(a_1, b; a_2) q(a_1, b, c; b)}{q(a_2, b; a_1) q(a_2, b, c; b)}$$

Because $\partial(B_4 - B_3) \subset (B_2 - B_1) \cup (B_k - B_4)$,

$$\frac{P(a_2, b, c, d, e, d^+, e^+)}{P(a_1, b, c, d, e, d^+, e^+)} = \frac{P(a_2, b, c, d, e, e^+)}{P(a_1, b, c, d, e, e^+)}$$

and therefore

$$(3.16) \quad \frac{P(a_2, b, c, d, e, e^+)}{P(a_1, b, c, d, e, e^+)} = \psi(a_1, a_2, b, c) \equiv \frac{q(a_1, b; a_2) q(a_1, b, c; b)}{q(a_2, b; a_1) q(a_2, b, c; b)}.$$

But

$$(3.17) \quad \frac{P(a_2, b, c, d, e, e^+)}{P(a_1, b, c, d, e, e^+)} = \frac{\sum_f \mu(b, c, d, e, f)q(d, e, f; e)\mu(a_2 | b, c, d, e, f)}{\sum_f \mu(b, c, d, e, f)q(d, e, f; e)\mu(a_1 | b, c, d, e, f)}$$

But by condition D , for $i=1, 2$

$$|\mu(a_i | b, c, d, e, f) - \mu(a_i | b, c, d, e)| < r(k)$$

where $r(k) \rightarrow 0$ as $k \rightarrow \infty$.

Hence for $i=1, 2$

(3.18)

$$\sum_f \mu(b, c, d, e, f)q(d, e, f; e)\mu(a_i | b, c, d, e, f) = \mu(a_i | b, c, d, e)P(b, c, d, e, e^+) + \theta_i$$

where

$$\theta_i = \sum_f \mu(b, c, d, e, f)q(d, e, f; e)[\mu(a_i | b, c, d, e, f) - \mu(a_i | b, c, d, e)]$$

and

$$|\theta_i| \leq P(b, c, d, e, e^+)r(k).$$

From (3.16), (3.17) and (3.18),

$$\mu(a_2 | b, c, d, e)$$

$$= \psi(a_1, a_2, b, c)\mu(a_1 | b, c, d, e) + (\psi(a_1, a_2, b, c)\theta_1 - \theta_2)/P(b, c, d, e, e^+).$$

But since

$$(3.19) \quad |(\psi(a_1, a_2, b, c)\theta_1 - \theta_2) P(b, c, d, e, e^+)| \leq (\psi(a_1, a_2, b, c) + 1)r(k),$$

$$\lim_{k \rightarrow \infty} (\mu(a_2 | b, c, d, e) / \mu(a_1 | b, c, d, e)) = \psi(a_1, a_2, b, c).$$

Since $\psi(a_1, a_2, b, c)$ does not depend on d or e ,

$$(3.20) \quad \mu(a_i | \mathcal{G}_{B_1^c}) = \mu(a_i | \mathcal{G}_{B_3 - B_1})$$

and μ represents a second order Markov random field.

COROLLARY 1. *Under the conditions of Proposition 3.4, μ represents a first order Markov random field if and only if*

$$(3.21) \quad \frac{q(a_1, b, c; b)}{q(a_2, b, c; b)}$$

does not depend on c .

Proof. By (3.19) and (3.20),

$$(3.22) \quad \frac{\mu(a_2 | b, c)}{\mu(a_1 | b, c)} = \frac{q(a_1, b; a_2)q(a_1, b, c; b)}{q(a_2, b; a_1)q(a_2, b, c; b)}.$$

But under the condition (3.21)

$$\mu(a_2 \mid b, c) / \mu(a_1 \mid b, c)$$

does not depend on c and hence $\mu(a_i \mid b, c) = \mu(a_i \mid b)$ and μ represents a first order Markov random field.

COROLLARY 2. Let $N^*(\alpha) = \partial_2(\{\alpha\})$ and consider the family of local transition kernels (defined with respect to $N^*(\cdot)$) defined by

$$(3.23) \quad q_\alpha^*(a_1, b, c; a_2) = \frac{q(a_1, b; a_2)q(a_1, b, c; b)}{q(a_2, b; a_1)q(a_2, b, c; b)}.$$

Then under the conditions of Proposition 3.4, μ is locally invariant with respect to the family $\{q_\alpha^*(\cdot, \cdot)\}$ and invariant with respect to the associated asynchronous family.

Proof. The local invariance with respect to the family $\{q_\alpha^*(\cdot, \cdot)\}$ follows immediately from (3.21). The invariance with respect to the associated asynchronous family then follows from Proposition 3.1.

4. Synchronous systems and Markov random fields. In this section we consider Markov systems in the following two cases.

Case I. $A = \mathbb{Z}^1$, the set of integers with $N(k) \equiv \{k-1, k, k+1\}$, $S = \{0, 1\}$.

Case II. $A = \mathbb{Z}^2$, the integer lattice in the plane with

$$N(k, l) \equiv \{(k-1, l), (k, l), (k+1, l), (l-1, k), (l+1, k)\}; \quad S = \{0, 1\}.$$

The family $\{q_\alpha(\cdot, \cdot) : \alpha \in A\}$ is translation invariant if for each $a \in \mathcal{G}_{\{\alpha\}}$ and $b \in \mathcal{G}_{N(\alpha)}$,

$$(4.1) \quad q_{\theta\alpha}(\theta^*b; \theta^*a) = q_\alpha(b; a)$$

where θ is a translation on A and θ^* is defined by

$$\theta^*b(\beta) \equiv b(\theta^{-1}\beta).$$

Similarly a Markov random field μ is homogeneous if for $a \in \mathcal{G}_{\{\alpha\}}$, $b \in \mathcal{G}_{N(\alpha)}$ and a translation θ ,

$$(4.2) \quad \mu(a \mid b) = \mu(\theta^*a \mid \theta^*b).$$

A homogeneous system of local transition kernels $\{q_\alpha(\cdot, \cdot) : \alpha \in A\}$ is symmetric if for $a \in \mathcal{G}_{\{0\}}$ and $b \in \mathcal{G}_{N(0)}$,

$$q_0(\phi^*b; a) = q_0(b; a)$$

where ϕ is a reflection about 0 in Case I and ϕ is any reflection in one of the axes in Case II and

$$\phi^*b(\alpha) \equiv b(\phi^{-1}\alpha).$$

A homogeneous system of local transition kernels $\{q_\alpha(\cdot, \cdot) : \alpha \in \mathbb{Z}^2\}$ is said to be degenerate if for all $a \in \mathcal{G}_{\{\alpha\}}$, $b \in \mathcal{G}_{N(\alpha)}$,

$$(4.3) \quad q_0(b; a) = q_0(\pi b, a)$$

where π is either $\pi_{\{(0, 1), (0, 0), (0, -1)\}}$ or $\pi_{\{(1, 0), (0, 0), (-1, 0)\}}$. In other words the system is degenerate if the interaction is essentially one dimensional. Similarly a Markov random field on Z^2 is said to be *degenerate* if $\mu(a | b) = \mu(a | \pi b)$.

THEOREM 4.1. *Let μ represent a homogeneous first order Markov random field on Z^1 . Then there exists a synchronous Markov system associated with a symmetric translation invariant family of local transition kernels $\{q_\alpha(\cdot, \cdot) : \alpha \in A\}$ for which μ is an invariant measure.*

Proof. F. Spitzer [4] has shown that such a Markov random field is characterized by

$$(4.4) \quad \mu(1 | k \text{ nearest neighbours occupied}) = 1/(1 + \alpha_0 r_0^k)$$

and

$$\mu(0 | k \text{ nearest neighbours occupied}) = \alpha_0 r_0^k / (1 + \alpha_0 r_0^k)$$

with $\alpha_0 > 0, r_0 > 0$ and where we say that a site is occupied if it assumes the value 1. In order to find the appropriate family of local transition kernels $\{q_\alpha(\cdot, \cdot) : \alpha \in A\}$ it suffices to find a family such that

$$(4.5) \quad \frac{q(0, b; 1) q(0, b, c; b)}{q(1, b; 0) q(1, b, c; b)} = \alpha_0 r_0^k, \quad k = 0, 1, 2$$

where k is the number of sites which are occupied in the configuration b over $\{-1, 1\}$. Hence in particular

$$(4.6) \quad q(0, b, c; b) / q(1, b, c; b) \text{ must be independent of } c.$$

where c is any configuration over the set $\{-2, 2\}$. Let

$$g(k) \equiv q(0, k; 0)$$

where k is defined as above and

$$h(k) \equiv q(1, k; 1).$$

Using the condition (4.6) for the various possible choices for the configurations b and c we obtain the equations

$$(4.7) \quad \begin{aligned} (g(2)/g(1)) &= (g(1)/g(0)), \\ (h(2)/h(1)) &= (h(1)/h(0)). \end{aligned}$$

Hence

$$(4.8) \quad \begin{aligned} g(k) &= g u^k, & k &= 0, 1, 2 \\ h(k) &= h v^k, & k &= 0, 1, 2 \end{aligned}$$

and

$$(4.9) \quad \frac{q(1, b, c; b)}{q(0, b, c; b)} = u^{2-k} v^k$$

where k is the number of sites in $\{-1, 1\}$ which are occupied in the configuration b .

Therefore by (4.5),

$$(4.10) \quad \frac{q(0, k; 1)}{q(1, k; 0)} = u^{2-k} v^k \alpha_0 r_0^k = \alpha r^k$$

where $\alpha = u^2 \alpha_0$, $r = (v/u)r_0$; $\alpha > 0$, $r > 0$. Then (4.10) implies that

$$q(0, k; 1) = 1 - g(k) = \alpha r^k q(1, k, 0) = \alpha r^k (1 - h(k)) \quad k = 0, 1, 2.$$

Hence

$$gu^k + \alpha r^k - (\alpha h)(rv)^k - 1 = 0, \quad k = 0, 1, 2.$$

Letting $d \equiv -\alpha h$, $w \equiv rv$, this becomes

$$(4.11) \quad \begin{aligned} gu^2 + \alpha r^2 + dw^2 &= 1 \\ gu + \alpha r + dw &= 1 \\ g + \alpha + d &= 1. \end{aligned}$$

The system (4.11) is a system of three equations in three unknowns which has a solution if and only if the Vandermonde determinant

$$(4.12) \quad \begin{vmatrix} u^2 & r^2 & w^2 \\ u & r & w \\ 1 & 1 & 1 \end{vmatrix} = (u-r)(u-w)(r-w) \neq 0$$

and is given by

$$(4.13) \quad g = \frac{(1-r)(1-w)}{(u-r)(u-w)}; \quad \alpha = \frac{(u-1)(1-w)}{(u-r)(r-w)}, \quad d = \frac{(u-1)(r-1)}{(u-w)(r-w)}.$$

Moreover in order for the solution to generate a probability transition kernel we must satisfy $g > 0$, $d < 0$. Because of the symmetry of the problem it suffices to show that this can be done in the case $r > 1$. We will show that a solution exists satisfying

$$(4.14) \quad w > r > 1 > u.$$

In this case (4.12) is automatically satisfied and also $g > 0$, $d < 0$ is satisfied.

The remaining condition is

$$(4.15) \quad (w-r)/(w-1) = (1/\alpha)((1-u)/(r-u)) \text{ where } \alpha \text{ is given.}$$

If we chose $u < 1$ so that

$$(4.16) \quad (1/\alpha)((1-u)/(r-u)) < 1$$

it is then possible to solve (4.15) with $w > r$. Hence if we consider the local transition function obtained from this solution, then the corresponding synchronous system has μ as an invariant measure and the proof is complete.

THEOREM 4.2. *Let μ represent a homogeneous non-degenerate first order Markov random field on Z^2 . Then there does not exist a reversible synchronous Markov system associated with a symmetric and translation invariant family of local transition kernels having μ as invariant measure.*

Proof. A homogeneous non-degenerate first order Markov random field in Z^2 is characterized (F. Spitzer [4]) as follows:

$$\begin{aligned} \mu(1 \mid k_1; k_2) &= 1/(1 + \tilde{\alpha}\tilde{r}_1^{k_1}\tilde{r}_2^{k_2}) \\ \mu(0 \mid k_1; k_2) &= \tilde{\alpha}\tilde{r}_1^{k_1}\tilde{r}_2^{k_2}/(1 + \tilde{\alpha}\tilde{r}_1^{k_1}\tilde{r}_2^{k_2}) \end{aligned}$$

with $\tilde{\alpha} > 0, \tilde{r}_1 > 0, \tilde{r}_2 > 0, \tilde{r}_1 \neq 1, \tilde{r}_2 \neq 1$ and where

- k_1 = number of neighbouring sites occupied in the vertical direction
- k_2 = number of neighbouring sites occupied in the horizontal direction.

From Corollary 1 to Proposition 3.4 it follows that a necessary condition for a reversible synchronous system with local transition kernels $\{q_\alpha(\cdot, \cdot); \alpha \in Z^2\}$ to have μ as an invariant measure is

$$(4.17) \quad \frac{q(0, b; 1) q(0, b, c; b)}{q(1, b; 0) q(1, b, c; b)} = \tilde{\alpha}\tilde{r}_1^k \tilde{r}_2^k.$$

If b is the configuration over the boundary of a single site, then b can assume 16 different values (see Figure 1).

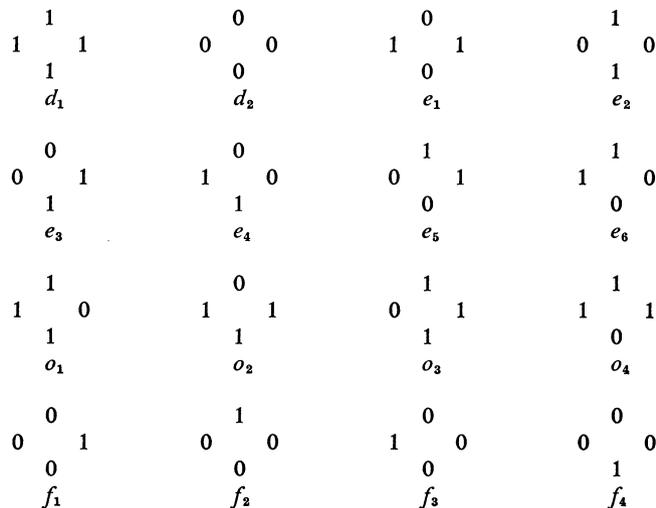


FIGURE 1.
Possible configurations for b .

Let

$$(4.18) \quad \begin{aligned} \varphi(b) &\equiv q(0, b; 0) \\ \varphi'(b) &\equiv q(1, b; 1). \end{aligned}$$

If we assume that the system $\{q_\alpha(\cdot, \cdot) : \alpha \in Z^2\}$ is symmetric, then

$$(4.19) \quad \begin{aligned} \varphi(f_1) &= \varphi(f_3); & \varphi(f_2) &= \varphi(f_4) \\ \varphi(e_3) &= \varphi(e_4) = \varphi(e_5) = \varphi(e_6) \\ \varphi(o_1) &= \varphi(o_3); & \varphi(o_2) &= \varphi(o_4). \end{aligned}$$

From (4.17)

$$(4.20) \quad R(c) \equiv \frac{q(0, b, c; b)}{q(1, b, c; b)} \text{ is independent of } c.$$

Let $b=d_2$ (c.f. Figure 1) and let us compute $R(c)$ for a number of choices for c . The different configurations chosen for c are listed in Figure 2. Then we obtain

$$(4.21) \quad R(A) = \varphi^4(d_2)/\varphi^2(f_1)\varphi^2(f_2)$$

$$(4.22) \quad R(B) = \varphi^2(o_1)\varphi^2(o_2)/\varphi^4(d_1)$$

$$(4.23) \quad R(C) = \varphi^3(d_2)/\varphi(e_2)\varphi^2(f_1)$$

$$(4.24) \quad R(D) = \varphi^3(d_2)/\varphi(e_2)\varphi^2(f_2)$$

$$(4.25) \quad R(E) = \varphi^2(d_2)/\varphi(o_1)\varphi(f_1)$$

$$(4.26) \quad R(F) = \varphi^2(d_2)/\varphi(o_2)\varphi(f_2)$$

$$(4.27) \quad R(G) = \varphi(f_1)\varphi(o_1)\varphi(d_2)/\varphi(d_1)\varphi(e_6)\varphi(e_4).$$

We then obtain

$$(4.28) \quad \varphi(o_1) = \varphi(f_1)\varphi^2(f_2)/\varphi^2(d_2)$$

from (4.21) and (4.25)

$$(4.29) \quad \varphi(o_2) = \varphi^2(f_1)\varphi(f_2)/\varphi^2(d_2)$$

from (4.21) and (4.26)

$$(4.30) \quad \varphi(d_1) = \varphi^2(f_1)\varphi^2(f_2)/\varphi^3(d_2)$$

from (4.21), (4.22), (4.28) and (4.29)

$$(4.31) \quad \varphi(e_1) = \varphi^3(f_1)/\varphi(d_2)$$

from (4.21) and (4.24)

$$(4.32) \quad \varphi(e_2) = \varphi^2(f_2)/\varphi(d_2)$$

from (4.21) and (4.23)

$$(4.33) \quad \varphi(e_3) = \varphi(f_1)\varphi(f_2)/\varphi(d_2)$$

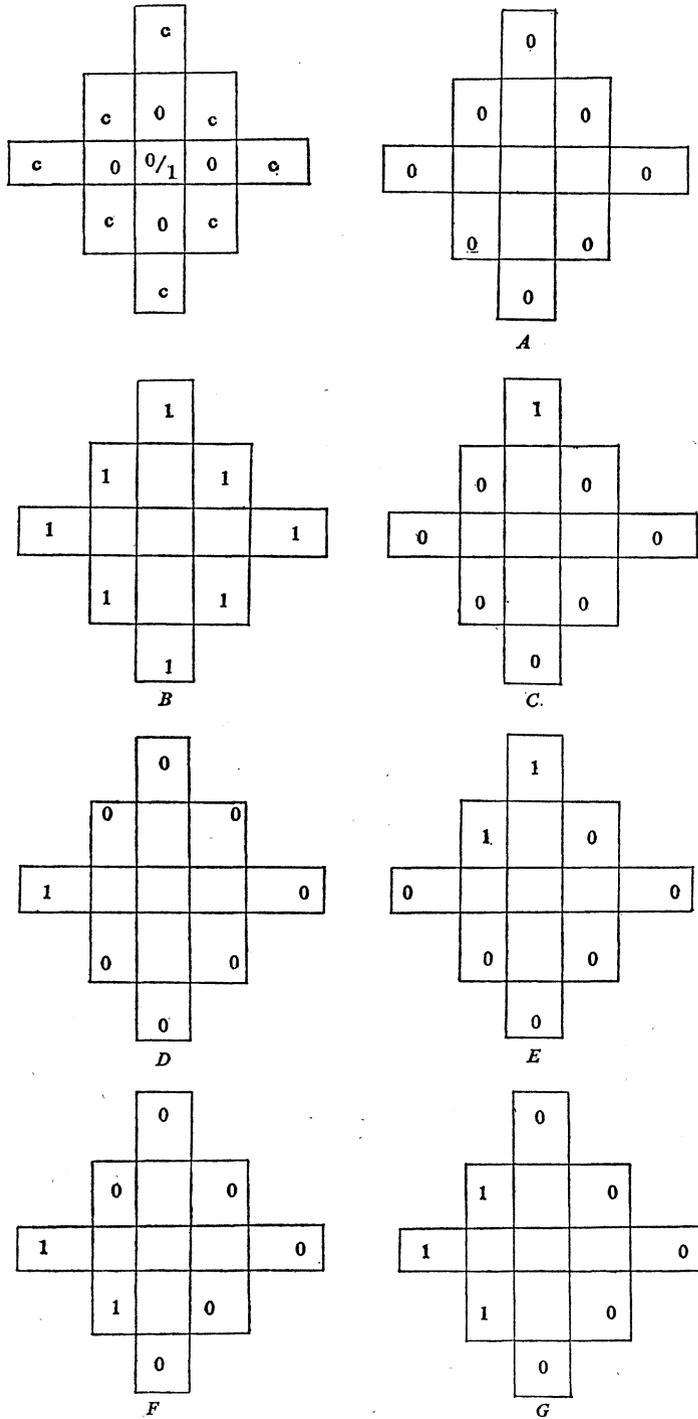


FIGURE 2.
Configurations for *c*.

from (4.21) and (4.27). Similar results are obtained in the same way for $\varphi'(b)$. Equations (4.28)–(4.33) imply that

$$(4.34) \quad \begin{aligned} q(0, b; 0) &= g u_1^{k_1} u_2^{k_2} \\ q(1, b; 1) &= h v_1^{k_1} v_2^{k_2} \end{aligned}$$

where

k_1 = number of sites in b in the vertical direction which are occupied
 k_2 = number of sites in b in the horizontal direction which are occupied.

Then

$$(4.35) \quad \frac{q(0, b, c; b)}{q(1, b, c; b)} = 1 / (v_1^{k_1} v_2^{k_2} u_1^{2-k_1} u_2^{2-k_2})$$

and therefore by (4.17)

$$(4.36) \quad \frac{q(0, b; 1)}{q(1, b; 0)} = \alpha r_1^{k_1} r_2^{k_2}$$

where $\alpha = \tilde{\alpha} u_1^2 u_2^2$, $r_1 = \tilde{r}_1 v_1 / u_1$, $r_2 = \tilde{r}_2 v_2 / u_2$. Therefore

$$(4.37) \quad 1 - q(0, b; 0) = \alpha r_1^{k_1} r_2^{k_2} (1 - q(1, b; 1)).$$

From (4.34) and (4.37) we obtain

$$(4.38) \quad \alpha r_1^{k_1} r_2^{k_2} + d w_1^{k_1} w_2^{k_2} + g u_1^{k_1} u_2^{k_2} = 1; \quad k_1, k_2 = 0, 1, 2$$

where $d = -\alpha h$, $w_1 = r_1 v_1$, $w_2 = r_2 v_2$. Letting $k_2 = 0$ in (4.38), we obtain

$$(4.39) \quad \begin{aligned} \alpha r_1^2 + d w_1^2 + g u_1^2 &= 1 \\ \alpha r_1 + d w_1 + g u_1 &= 1 \\ \alpha + d + g &= 1 \end{aligned}$$

and letting $k_2 = 1$ in (4.39), we obtain

$$(4.40) \quad \begin{aligned} (\alpha r_2) r_1^2 + (d w_2) w_1^2 + (g u_2) u_1^2 &= 1 \\ (\alpha r_2) r_1^2 + (d w_2) w_1 + (g u_2) u_1 &= 1 \\ (\alpha r_2) + (d w_2) + (g u_2) &= 1. \end{aligned}$$

But the systems (4.39) and (4.40) can only hold in two situations

$$(4.41) \quad r_2 = w_2 = u_2 = 1$$

or (4.42) the coefficient determinant

$$\begin{vmatrix} r_1^2 & w_1^2 & u_1^2 \\ r_1 & w_1 & u_1 \\ 1 & 1 & 1 \end{vmatrix} = (u_1 - r_1)(u_1 - w_1)(w_1 - r_1) = 0.$$

We must exclude (4.41) since it would imply that the system is degenerate. Therefore $u_1=r_1$ or $u_1=w_1$ or $w_1=r_1$. But in any of these three cases we arrive at a system of equations of the following form

$$\begin{aligned}
 &\alpha r_1^2 + (d+g)w_1^2 + e = 0 \\
 (4.43) \quad &\alpha r_1^2 + (d+g)w_1 + e = 0 \\
 &\alpha + (d+g) + e = 0 \quad \text{with } e = -1.
 \end{aligned}$$

But (4.43) has a non-trivial solution for $\alpha, (d+g), e$ only if

$$\begin{vmatrix} r_1^2 & w_1^2 & 1 \\ r_1 & w_1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = (1-r_1)(1-w_1)(w_1-r_1) = 0.$$

Hence $w_1=r_1$ or $w_1=1$ or $r_1=1$. But each of these cases again leads to a system of the form

$$\begin{aligned}
 (4.44) \quad &\alpha r_1 + (d+g+e) = 0 \\
 &\alpha + (d+g+e) = 0
 \end{aligned}$$

which implies that $r_1=1$ and hence $w_1=r_1=u_1=1$. Therefore equations (4.39) and (4.40) imply that either $u_1=w_1=r_1=1$ or $u_2=w_2=r_2=1$ which implies that the system is degenerate. Hence it is impossible to find a system of local transition kernels satisfying (4.17) and the proof is complete.

To complete this section we consider a phenomena discovered by R. P. Kindermann [4]. He discovered an example of a one dimensional synchronous Markov system whose invariant measure satisfies the condition

$$\mu(a \mid b, c) \text{ depends on } c \text{ but not on } b,$$

that is,

$$(4.45) \quad \mu(\cdot \mid \mathcal{G}_{B_1^c}) = \mu(\cdot \mid \mathcal{G}_{B_3-B_2})$$

where B_1, B_2, B_3 are defined as in the proof of Proposition 3.2. We call such a random field a *Kindermann field*. Consider a synchronous Markov system in Z^1 and let

$$\begin{aligned}
 g(k) &\equiv q(0, k; 0) \\
 h(k) &\equiv q(1, k; 1).
 \end{aligned}$$

For a reversible system which has μ as an invariant measure, we have

$$(4.46) \quad \frac{\mu(a_1 \mid b, c)}{\mu(a_2 \mid b, c)} = \frac{q(a_2, b, c; b) q(a_2, b; a_1)}{q(a_1, b, c; b) q(a_1, b; a_2)}.$$

If μ is a Kindermann field we obtain the following equations by allowing b and c to assume their various possible values in (4.46).

$$(4.47a) \quad \frac{g(0)(1-g(0))}{g(1)(1-g(1))} = \frac{h(0)(1-h(0))}{h(1)(1-h(1))}$$

$$(4.47b) \quad \frac{g(0)(1-g(1))}{g(1)(1-g(2))} = \frac{h(0)(1-h(1))}{h(1)(1-h(2))}$$

$$(4.47c) \quad \frac{g(1)(1-g(1))}{g(2)(1-g(2))} = \frac{h(1)(1-h(1))}{h(2)(1-h(2))}$$

$$(4.47d) \quad \frac{g(1)(1-g(0))}{g(2)(1-g(1))} = \frac{h(1)(1-h(0))}{h(2)(1-h(1))}$$

We can rewrite (4.47a) and (4.47c) as

$$\frac{g(0)(1-g(0))}{h(0)(1-h(0))} = \frac{g(1)(1-g(1))}{h(1)(1-h(1))} = \frac{g(2)(1-g(2))}{h(2)(1-h(2))} = \gamma.$$

The pair $(g(i), h(i))$ must be a solution of

$$x^2 - x = \gamma(y^2 - y).$$

In the case $\gamma=1$, we obtain $x=y$ or $x=1-y$ and in the latter case the remaining equations 4.47b and 4.47d reduce to

$$(4.48) \quad h(0)h(2)(1-h(1))^2 = h^2(1)(1-h(2))(1-h(0)).$$

In the case $h(1)=1/2$, (4.48) becomes

$$h(2) = 1-h(0).$$

Hence one solution of the system (4.47) is given by

$$\begin{aligned} h(i) &= 1-g(i), & i &= 0, 1, 2, \\ h(1) &= \frac{1}{2}, & h(2) &= 1-h(0). \end{aligned}$$

This example is the one discovered by R. P. Kindermann in the context of models of voting behaviour. Hence for a Kindermann field, $h(0), h(1), h(2)$ forms an arithmetic progression whereas for a Gibbs field $h(0), h(1), h(2)$ forms a geometric progression.

ACKNOWLEDGEMENT. I would like to thank R. Fischler for some helpful suggestions.

REFERENCES

1. M. B. Averintzev, *On a method of describing discrete parameter random fields*, *Problemy Peredaci Informacii* (6), 1970, 100-109.
2. D. A. Dawson, *Information flow in discrete Markov systems*, to appear in *J. Appl. Prob.*
3. R. L. Dobrushin, *The description of a random field by means of conditional probabilities and conditions of its regularity*, *Th. Prob. Appl.* (13), 1968, 197-224.
4. R. P. Kindermann, *Random fields; theorems and examples*, *J. Undergraduate Mathematics* (5), 1973, 25-34.
5. F. Spitzer, *Random fields and interacting particle systems*, Summer Seminar Notes, M.A.A., 1971.

CARLETON UNIVERSITY,
OTTAWA, CANADA