The Annals of Probability
1975, Vol. 3, No. 3, 387-398

SPECIAL INVITED PAPER

MARKOV RANDOM FIELDS ON AN INFINITE TREE[1]

BY FRANK SPITZER

Cornell University

Phase transition is studied on the infinite tree T_N in which every point has exactly $N+1$ neighbors. For every assignment of conditional probabilities which are invariant under graph isomorphism there is a Markov chain with these conditional probabilities and the main results ascertain for which ones of these chains there are other Markov random fields with the same conditional probabilities.

Let T_N, $N \geqq 1$ be the infinite tree with $N+1$ branches emanating from every vertex. When $N = 1$ this means that $T_1 = \mathbb{Z}_1$, the integers. When $N \geqq 2$, T_N is the connected infinite graph without loops. Two points $x \neq y$ in T_N are neighbors if they are connected by a branch. For any two points $x \neq y$ there is a unique path $x = x_1, x_2, \cdots, x_{k+1} = y$ such that x_i and x_{i+1} are neighbors for $i \leqq 1 \leqq k$. Our goal is to discuss certain probability measures μ on the space $\Omega = \{0, 1\}^{T_N}$ (with the σ-algebra generated by the finite dimensional cylinders). We are interested in those probability measures (called Markov random fields) which reduce to ordinary 0, 1 valued stationary Markov chains in the case when $N = 1$. These questions are of far greater importance in the setting of equilibrium statistical mechanics, where the graph $\mathbb{Z}_N$ is of principal interest rather than T_N. Indeed all the methods we shall use here to obtain rather complete results were first developed to solve the analogous, much more difficult problems for $\mathbb{Z}_N$, which are still not completely solved. For recent surveys see [4], [5], [8], [11]. The infinite trees T_N were first studied by Preston ([8] pages 97–105), who proved Theorems 1, 2, 3, and 6 which follow.

We begin by stating the principal definitions and results.

DEFINITION 1. A Markov random field (MRF) is a probability measure μ on $\Omega = \{0, 1\}^{T_N}$, with strictly positive values for finite cylinder sets, and such that conditional probabilities of the form $\mu[\omega(x) = 1 \mid \omega(\cdot) \text{ on } T_N \backslash x]$ depend only on the values of ω at the neighbors of x. Finally these conditional probabilities are assumed invariant under graph isomorphism (but not μ itself!). The set of all MRF's is denoted $\mathscr{G}$.

It follows from the invariance requirement that the conditional probabilities

Received August 30, 1974.

[1] This paper was presented as a Special Invited Paper at the IMS annual meeting in Edmonton, Canada, August 15, 1974.

AMS 1970 *subject classifications.* 60J10, 60K35, 82A25.

Key words and phrases. Phase transition, Markov chains on infinite trees, Markov random fields.

are determined by $N+2$ parameters,

(1) $\alpha_k = \mu[\omega(x) = 1 \mid \omega = 1$ at exactly k of the neighbors of $x]$,
$0 \leqq k \leqq N+1$.

Not all possible vectors $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_{N+1})$ are realizable, of course, by a MRF. A familiar result concerning the equivalence of MRF's and so called Gibbs states ([8] Theorem 4.1) describes exactly the class of realizable α.

THEOREM 1 [8]. *The vector α is realized by a* MRF *if and only if there exists a pair of positive numbers x and y, such that*

(2) $$\alpha_k = [1 + y \cdot x^{2k-(N+1)}]^{-1}, \qquad 0 \leqq k \leqq N+1.$$

Therefore we make

DEFINITION 2. $\mathscr{G}_\alpha \subset \mathscr{G}$ is the class of MRF's with a particular α satisfying (2), and one has the decomposition

(3) $$\mathscr{G} = \bigcup_\alpha \mathscr{G}_\alpha .$$

Note that each $\mathscr{G}_\alpha$ may consist of one or of many MRF's. Our goal is to describe which is the case, for all possible α, when $N \geqq 2$. When $N = 1$, it is known ([2] Theorem 3, [10] Theorem 3.22) that $|\mathscr{G}_\alpha| = 1$ for all α. This study will begin by showing that each $\mathscr{G}_\alpha$ contains a particularly simple and elegant type of MRF which we shall call a Markov chain. (The theory on $\mathbb{Z}_N$ is much deeper primarily because it contains no analogue of these simple objects.)

DEFINITION 3. For every strictly positive stochastic 2×2 matrix $M = \{M(i,j)\}$, $i, j = 0, 1$, a probability measure μ_M on Ω is defined as follows: First let $\pi = \{\pi(0), \pi(1)\}$ be the unique invariant probability measure for M ($\pi M = \pi$). Then, for any finite connected subset $A \subset T_N$, let ε be a function from A to $\{0, 1\}$, and define a simple ordering $\mathbf{A} = \{x_1, x_2, \cdots, x_k\}$ of A with the property that each x_j with $j > 1$ is the neighbor of exactly one $x_i \in \{x_1, x_2, \cdots, x_{j-1}\}$. Denote this index $i = i(j)$. Thus $i(2) = 1$. Define the cylinder set probabilities of μ_M by

(4) $$\mu_M[\omega(t) = \varepsilon(t), t \in A] = \pi(x_1) \prod_{j=2}^{k} M(\varepsilon(x_{i(j)}), \varepsilon(x_j)) .$$

Note that when $T_n = T_1 = \mathbb{Z}$ this definition obviously gives a stationary Markov chain if we let $\mathbf{A}$ be the usual ordering of A, which is an interval of integers. In fact (4) is independent of the ordering $\mathbf{A}$ of A chosen. This is an easy consequence of the time reversibility of two valued stationary Markov chains with strictly positive transition matrix, i.e. of

(5) $$\pi(i)M(i,j) = \pi(j)M(j,i), \quad i, j \in \{0, 1\} .$$

In fact an easy induction on the cardinality of A, shows

THEOREM 2 [8]. *Definition* 3 *defines unique consistent cylinder set probabilities (independent of the ordering* $\mathbf{A}$ *for every finite connected $A \subset T_n$) and hence a unique probability measure μ_M on Ω.*

DEFINITION 4. For each strictly positive M, μ_M is called a Markov chain (MC) and $\mathscr{M}$ is the class of all Markov chains.

An easy calculation shows that every MC is an MRF, or in other words that $\mathscr{M} \subset \mathscr{G}$. In fact the class $\mathscr{M}$ is large enough so that every $\mathscr{G}_\alpha$ contains at least one element of $\mathscr{M}$. This and subsequent results will now be stated in rapid succession, and the proofs will follow.

THEOREM 3 [8]. *$\mathscr{M} \subset \mathscr{G}$, and for every α satisfying* (2), *the cardinality $|\mathscr{M} \cap \mathscr{G}_\alpha|$ is either* 1, 2, *or* 3 (*depending on α*). *When $N = 1$, $|\mathscr{M} \cap \mathscr{G}_\alpha| = 1$. When $N > 1$, $|\mathscr{M} \cap \mathscr{G}_\alpha|$ can take all three values*, 1, 2, *and* 3.

In order to further elucidate the role of $\mathscr{M}$ as a subset of $\mathscr{G}$ we make

DEFINITION 5. $\mathscr{H}$ is the class of all homogeneous probability measures on Ω, i.e. those which are invariant under graph isomorphisms of T_N (translation and reflection of $\mathbb{Z}$ when $N = 1$). Let $\mathscr{T}$ be the class of all probability measures on Ω with trivial tail field.

For each α, $\mathscr{G}_\alpha$ is a compact and convex set (in fact a Choquet simplex [8] Proposition 5.2, [5], [6]). Its extreme points are denoted $\operatorname{Ext}(\mathscr{G}_\alpha)$. Part (i) of the following theorem is well known ([8] Theorem 11.1, [5], [6]).

THEOREM 4. (i) $\operatorname{Ext}(\mathscr{G}_\alpha) = \mathscr{T} \cap \mathscr{G}_\alpha$;
(ii) $\mathscr{M} \subset \mathscr{T} \cap \mathscr{H}$;
(iii) $\operatorname{Ext}(\mathscr{G}_\alpha) \cap \mathscr{H} = \mathscr{M} \cap \mathscr{G}_\alpha$.

Combining Theorems 3 and 4 we see that $\mathscr{G}_\alpha$ has always at least one homogeneous extreme point, and more than one if and only if $|\mathscr{M} \cap \mathscr{G}_\alpha| > 1$. To find useful conditions it is more convenient to parametrize the problem by use of M instead of α.

DEFINITION 6. For every strictly positive transition matrix

$$M = \begin{pmatrix} s & 1-s \\ 1-t & t \end{pmatrix}, \qquad s, t \in (0, 1) \tag{6}$$

let $\mathscr{G}_M = \mathscr{G}_\alpha$ with α chosen (uniquely) so that $\mu_M \in \mathscr{G}_\alpha$. Let φ be the rational function

$$\varphi(x) = \frac{tx^N + 1 - t}{(1 - s)x^N + s}. \tag{7}$$

If M and M' are two matrices of the type in (6) it may happen that they give rise to MRF's which lie in the same $\mathscr{G}_\alpha$. We shall characterize when this happens.

THEOREM 5. *For each M satisfying* (6), *$|\mathscr{G}_M \cap \mathscr{M}| = 1$ if and only if the equation $\varphi(x) = x$ has only one positive real root* (*namely $x = 1$*). *When $N = 1$, this is always the case. When $N \geq 2$, $\mathscr{G}_M \cap \mathscr{M}$ always consists of one, two, or three* MC's, *μ_M being one of them. When $N = 2$, here is a detailed classification: divide the unit square $0 < s < 1$, $0 < t < 1$ into the three regions defined by*

$$R_1 = \{D(s, t) < 0\} \cup \{s = t = \tfrac{3}{4}\}$$
$$R_2 = \{s + t = \tfrac{3}{2} \ \ and \ \ s \neq \tfrac{3}{4}\} \cup \{D(s, t) = 0 \ \ and \ \ s \neq \tfrac{3}{4}\}$$
$$R_3 = \{D(s, t) > 0 \ \ and \ \ s + t \neq \tfrac{3}{2}\},$$

where $D(s, t) = (s - t)^2 + 2(s + t) - 3$.
Then $|\mathscr{M} \cap \mathscr{G}_M| = k$ *on* R_k, $k = 1, 2, 3$.

Theorem 5 still does not tell us the cardinality of $\mathscr{G}_M$, even when $|\mathscr{G}_M \cap \mathscr{M}| = 1$. To understand the connection between $|\mathscr{G}_M|$ and $|\mathscr{G}_M \cap \mathscr{M}|$ we have to introduce a classification familiar from statistical physics.

DEFINITION 7. The matrix M in (6) is attractive if $s + t \geqq 1$, repulsive if $s + t < 1$.

In the attractive case C. Preston showed how one can sharpen Theorem 5.

THEOREM 6 [8]. *If M is attractive, then $|\mathscr{G}_M| = 1$ if and only if the equation $\varphi(x) = x$ has only one positive real root (namely $x = 1$).*

In the repulsive case it follows immediately from Theorem 5 that $|\mathscr{M} \cap \mathscr{G}_M| = 1$. But it may happen, nevertheless, that $|\mathscr{G}_M| > 1$.

THEOREM 7. *In the repulsive case $|\mathscr{G}_M| = 1$ if and only if the equation $\varphi \circ \varphi(x) = x$ has only one positive solution (namely $x = 1$). When $N = 2$ this happens if and only if $s + t \geqq \frac{1}{2}$.*

The proof of Theorem 7 will depend on a new class of non-homogeneous Markov chains exhibiting the symmetry break-down into even and odd states associated with the repulsive (anti-ferro-magnetic) case in statistical mechanics [3].

DEFINITION 8. Let M^e and M^0 be two stochastic matrices as in (6) and π^e, π^0 two probability vectors on $\{0, 1\}$ such that

$$\pi^e(i)M^e(i, j) = \pi^0(j)M^0(j, i), \qquad i, j \in \{0, 1\},\ M^e \neq M^0. \tag{8}$$

Decompose $T_N = E \cup 0$ where E are the even sites (points which can be reached by an even number of branches from some fixed site) and $0 = T_N \backslash E$. Define the probability measure μ_{M^e, M^0} as in Definition 3, using π^e for even sites, π^0 for odd sites, M^e for transitions from E to 0, and M^0 for transitions from 0 to E.

Just as in Theorem 2 it can be shown that this defines consistent cylinder set probabilities, which define an MRF μ_{M^e, M^0}. These probability measures enter the picture in the following way.

THEOREM 8. *In the repulsive case $|\mathscr{G}_M| > 1$ if and only if $\mathscr{G}_M$ contains an* MRF *μ_{M^e, M^0}, with $M^e \neq M^0$.*

In the attractive case it is easy to see that $\varphi(x)$ is monotone increasing on $x > 0$. Therefore the positive solutions of $\varphi(x) = x$ are exactly the positive solutions of $\varphi \circ \varphi(x) = x$. Hence Theorems 5, 6, and 7 can be combined into

THEOREM 9. *For each $M = \begin{pmatrix} s & 1-s \\ 1-t & t \end{pmatrix}$, $s, t \in (0, 1)$, $\mathscr{G}_M$ consists of a single probability measure (namely μ_M) if and only if the equation $\varphi \circ \varphi(x) = x$ has only one positive solution (namely $x = 1$). When $N = 1$ this is always the case, and when $N = 2$ in the unshaded region, where the repulsive shaded region is the open set*

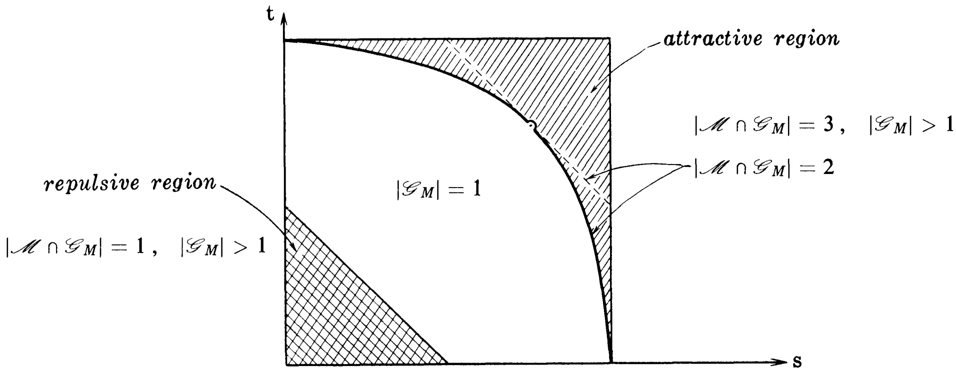

FIG. 1.

$s > 0$, $t > 0$, $s + t < \frac{1}{2}$, *while the attractive shaded region, described by* $s < 1$, $t < 1$, $(s - t)^2 + 2(s + t) \geqq 3$ *and* $(s, t) \neq (\frac{3}{4}, \frac{3}{4})$ *is neither open nor closed.*

PROOF OF THEOREM 1. It follows from [8], Theorem 4.1, that an MRF is an infinite Gibbs state with homogeneous nearest neighbor pair potential U and vice versa. Let $U(x, x) = u_0$ and $U(x, y) = U(y, x) = u_1$ when x and y are neighbors. Otherwise $U(x, y) = 0$. If we use U to define Gibbs states by the Boltzman formula

$$\mu(A) = Z_\Lambda^{-1} \exp -\tfrac{1}{2} \textstyle\sum_{x \in A} \sum_{y \in A} U(x, y)\,, \qquad A \subset \Lambda\,, \tag{9}$$

then any infinite Gibbs state with potential U will have the conditional probabilities

$$\alpha_k = \frac{1}{1 + \exp\left\{\dfrac{u_0}{2} + ku_1\right\}} = \frac{1}{1 + y \cdot x^{2k-(N+1)}}\,, \qquad 0 \leqq k \leqq N + 1\,, \tag{10}$$

if

$$x = \exp\left(\frac{u_1}{2}\right), \qquad y = \exp \tfrac{1}{2}[u_0 + (N + 1)u_1]\,. \tag{11}$$

□

PROOF OF THEOREM 2. Formula (5) shows that the cylinder set probabilities are well defined (independent of the ordering **A** in Definition 3) when A consists of two neighboring points. Next, we shall show that the choice of x_1 in **A** is immaterial for finite connected A of any cardinality. The rest of the product in (4) is uniquely determined by the choice of x_1, since every vertex $x \neq x_1$ in A can only be reached by one uniquely oriented sequence of branches. Equation (5) may be used, step by step, to move x_1 from any site of A to any other, without changing the value of μ_M in (4). The cylinder set probabilities in (4) are obviously consistent, since M is a stochastic matrix. By Kolmogorov's extension theorem they therefore define a unique MC μ_M on Ω. □

PROOF OF THEOREM 3. It follows readily from Definitions 3 and 4 that μ_M is an MRF for every M. Hence $\mathscr{M} \subset \mathscr{G}$. Now suppose α satisfies (2) for some pair $x > 0$, $y > 0$. We shall show that there always exists a matrix M, satisfying (6), such that $\mu_M \in \mathscr{G}_\alpha$, and that the number of possible choices for M is always either 1, 2, or 3. By Theorem 1, $\mu_M \in \mathscr{G}_\alpha$ if and only if

$$\mu_M[\omega(x) = 1 \mid \omega = 1 \text{ at } N+1 \text{ neighbors}] = [1 + yx^{N+1}]^{-1}\,,$$
$$\mu_M[\omega(x) = 1 \mid \omega = 1 \text{ at } N \text{ neighbors}] = [1 + yx^{N-1}]^{-1}\,.$$

If $M = \binom{s \quad 1-s}{1-t \quad t}$, then, using (4), these two equations become

$$yx^{N+1} = \frac{(1-t)(1-s)^N}{t^{N+1}}\,, \qquad yx^{N-1} = \frac{s(1-s)^{N-1}}{t^N}\,. \tag{12}$$

The system (12) is equivalent to

$$\frac{1-s}{s} = \frac{tx^2}{1-t}\,, \tag{13}$$

and

$$tx^2 + 1 - t = \left(\frac{1-t}{t}\right)^{1/N} x(xy)^{-1/N}\,. \tag{14}$$

Theorem 3 will therefore hold if the number of solutions $t \in (0, 1)$ of (14) is always one, two or three. This is easily verified, since the left side in (14) changes linearly from 1 to x^2 as t goes from 0 to 1, and the right side decreases from ∞ to 0, and has exactly one point of inflection in (0, 1) when $N > 1$. When $N = 1$, (14) always has a unique solution. □

Proof of Theorem 4. Part (i) is the well-known result that the extremal Gibb's states with a given potential may be characterized by the property that their tail field is trivial. This will be essential for the proof of (iii). Part (ii) consists of two assertions. $\mathscr{M} \subset \mathscr{H}$ is immediate from the definition of $\mathscr{M}$. The fact that $\mathscr{M} \subset \mathscr{T}$ is well known if $N = 1$ ([9] Chapter 5), for then every $\mu_M \in \mathscr{M}$ is a positive, irreducible, ergodic, stationary Markov chain. Therefore it is strongly mixing and hence it has a trivial tail field. This proof is easily adapted to the infinite tree T_N with $N \geqq 2$. Let U, V be finite subsets of T_N and A, B the cylinder sets in Ω defined by

$$A = \{\omega : \omega = u \text{ on } U\}\,, \qquad B = \{\omega : \omega = v \text{ on } V\}\,.$$

Let $B^{(x)} = \{\omega : \omega = v \text{ on } V + x\}$, $x \in T_N$, and let $|x|$ be the distance from x to a fixed point a of A (the number of branches from x to a). Then a short computation based on (4) and the ergodic theorem for finite positive stochastic matrices shows that

$$\lim_{|x|\to\infty} \mu_M(A \cap B^{(x)}) = \mu_M(A)\mu_M(B)\,. \tag{15}$$

By a standard approximation argument ([1] Theorem 8.1.1) (15) continues to hold for arbitrary events A and B. If we choose $A = B$ in the tail field, then $B^{(x)} = A$ for each x, and (15) shows that $\mu_M(A) = 0$ or 1, so that the tail field is trivial.

To prove (iii) suppose that $\mu \in \operatorname{Ext}(\mathscr{G}_\alpha) \cap \mathscr{H}$. Let 0 be a fixed point in T_N, and let $\mathbb{Z}$ be a subgraph of T_N which is graph isomorphic to the integers, and which contains 0. The first step of the proof of (iii) will be to show that the projection $\tilde{\mu}$ of μ on $\{0, 1\}^{\mathbb{Z}}$ is an MRF (when $N = 1$, $T_N = \mathbb{Z}$, and then this is obvious). Let $\mathscr{F}_n$ be the σ-algebra generated by $\omega(x)$, $|x| \geqq n$, and $\mathscr{F}_\infty = \bigcap_{n=1}^\infty \mathscr{F}_n$ the trivial tail field. Let $E[\cdot \mid \cdot]$ denote conditional expectation with respect to μ. Let $j > 0$ be an element of $\mathbb{Z}$. Since μ is a nearest neighbor Gibbs state, the conditional expectations for finite sets depend only on the values on the boundary ([8] page 26). Hence

$$E[\omega(0) \mid \omega(-1), \omega(1), \mathscr{F}_n] = E[\omega(0) \mid \omega(k) \ \text{ for } \ |k| \leqq j, k \neq 0; \mathscr{F}_n]$$

for every $1 \leqq j \leqq n$.
Letting $n \to \infty$, and using the fact that $\mathscr{F}_\infty$ is trivial,

$$E[\omega(0) \mid \omega(-1), \omega(1)] = E[\omega(0) \mid \omega(k) \ \text{ for } \ |k| \leqq j, k \neq 0] .$$

Since $\mu \in \mathscr{H}$, we have

$$E[\omega(n) \mid \omega(n-1), \omega(n+1)] = E[\omega(n) \mid \omega(n+k) \ \text{ for } \ |k| \leqq j, k \neq 0] ,$$

and therefore $\tilde{\mu}$ is an MRF. It follows from [2], Theorem 3, or [10], Theorem 3.2.2, that $\tilde{\mu}$ is a stationary Markov chain. Let M denote its transition matrix. It follows that μ has the cylinder set probabilities specified by (4) for any finite set $A \subset \mathbb{Z}$. For finite sets A which cannot be imbedded in a subgraph isomorphic to $\mathbb{Z}$, a simple induction argument (on the cardinality of A) establishes that (4) holds. For example, take $N = 2$ and A the set $\{x, y, z, u\}$ where x, z, u are the neighbors of y, and think of x, y, z as imbedded in $\mathbb{Z}$. Then

$$\begin{aligned} \mu_M[\omega = \varepsilon \ \text{ on } \ A] &= \mu_M[\omega = \varepsilon \ \text{ on } \ \{x, y, z\}]\mu_M[\omega = \varepsilon \ \text{ at } \ u \mid \omega \ \text{ on } \ \{x, y, z\}] \\ &= \pi(\varepsilon(x))M(\varepsilon(x), \varepsilon(y))M(\varepsilon(y), \varepsilon(z)) \\ &\qquad \times \lim\nolimits_{n\to\infty} \mu_M[\omega = \varepsilon \ \text{ at } \ u \mid \omega \ \text{ on } \ \{x, y, z\}, \mathscr{F}_n] . \end{aligned}$$

The above limit is

$$\begin{aligned} \lim\nolimits_{n\to\infty} \mu_M[\omega = \varepsilon \ \text{ at } \ u \mid \omega \ \text{ at } \ y, \mathscr{F}_n] &= \mu_M[\omega = \varepsilon \ \text{ at } \ u \mid \omega \ \text{ at } \ y] \\ &= M(\varepsilon(y), \varepsilon(u)) . \end{aligned}$$

This gives the cylinder set probability required by (4). □

PROOF OF THEOREM 5. We start with M given by (6) and look for a transition matrix $\tilde{M}$ such that $\mu_{\tilde{M}} \in \mathscr{G}_M$. Let

(16) $$\frac{\tilde{M}_{10}}{\tilde{M}_{11}} = \xi , \qquad \frac{\tilde{M}_{01}}{\tilde{M}_{01}} = \eta .$$

A simple calculation shows that $\mu_{\tilde{M}} \in \mathscr{G}_M$ if and only if

(17) $$yx^{N+1} = \xi \left[\frac{\eta(1+\xi)}{1+\eta} \right]^N , \qquad x^2 = \xi\eta ,$$

where x, y are given by (12). Equation (17) can be written

$$(18) \qquad yx^{N-1}\eta = \left[\frac{\eta + x^2}{\eta + 1}\right]^N, \qquad x^2 = \xi\eta .$$

Using (12) to express yx^{N-1} and x^2 in terms of s and t gives

$$(19) \qquad \frac{\eta s}{1-s} = \left[\frac{t\frac{\eta s}{1-s} + 1 - t}{(1-s)\frac{\eta s}{1-s} + s}\right]^N, \qquad \frac{(1-t)(1-s)}{st} = \xi\eta .$$

Letting

$$(20) \qquad u = \left(\frac{\eta s}{1-s}\right)^{1/N},$$

(19) induces to

$$(21) \qquad u = \varphi(u), \qquad \xi = \frac{1-t}{t}u^{-N} .$$

Equation (21) always has the solution $u = 1$ which gives $\tilde{M} = M$. There are other possibilities for $\tilde{M}$ if and only if $\varphi(u) = u$ has a positive solution $u \neq 1$. When $N = 1$ this is never the case. When $N \geqq 2$ it is easy to check that $\varphi(x) = x$ can only have one, two or three positive solutions, but this also follows from Theorem 3. The detailed results for the case $N = 2$ follow from a careful analysis of the equation $\varphi(x) = x$. One obtains

$$(22) \qquad x - \varphi(x) = [(1-s)x^2 + s]^{-1}(x-1)[x^2(1-s) + x(1-s-t) + (1-t)]$$

from which one readily deduces that (22) has one, two or three positive zeros in the regions R_1, R_2, R_3 respectively. (Note that $(s, t) \in R_1$ whenever M is repulsive.) □

PROOF OF THEOREM 6. We take M given by (6), with $s + t \geqq 1$, and (s, t) such that $\varphi(x) = x$ has only the positive root $x = 1$, and we have to show that every MRF with the same conditional probabilities as μ_M must be μ_M itself. Let us then suppose that μ is such an MRF. In other words it is a Gibbs state with potential U as in (9), (10), (11) and the condition $s + t \geqq 1$, which by (11) and (13) is equivalent to $u_1 \leqq 0$, means that U is an attractive pair potential. There is an elegant criterion for the absence of phase transition for such a potential ([7], [8] Theorem 8.1): Let Λ_n be a sequence of finite subsets of T_N which increase to T_N. We shall take $\Lambda_n = \{x : X \in T_n, |x| > n\}$. Then the boundary $\partial\Lambda_n = \{x : |x| = n\}$. Let μ_n^+ and μ_n^- be the restrictions of μ_M to Λ_n, and conditioned by $\omega \equiv +1$ on $\partial\Lambda_n$ in the case of μ_n^+, and by $\omega \equiv 0$ on $\partial\Lambda_n$ in the case of μ_n^-. Then

$$(23) \qquad \mu_n^-[\omega(0) = 1] \leqq \mu[\omega(0) = 1] \leqq \mu_n^+[\omega(0) = 1], \qquad n \geqq 1$$

for every $\mu \in \mathscr{G}_M$, and $\mathscr{G}_M = \{\mu_M\}$ if and only if

$$(24) \qquad \lim_{n\to\infty} \mu_n^-[\omega(0) = 1] = \lim_{n\to\infty} \mu_n^+[\omega(0) = 1] .$$

Fortunately these limits can be explicitly calculated. Let

$$\rho_n^+ = \mu_n^+[\omega(0) = 1], \qquad \rho_n^- = \mu_n^-[\omega(0) = 1].$$

Decompose $\Lambda_n \cup \partial\Lambda_n$ into N isomorphic pieces, each starting at 0. Call one of these S_n. Thus every branch of S_n which starts at 0 has length n and there are 2^{n-1} of these. Let V_n be the set of 2^{n-1} end vertices

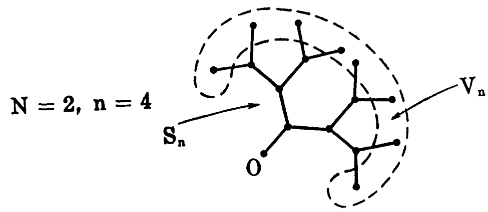

FIG. 2.

$$\begin{aligned}
\pi(0)R_n^+(0) &= \mu_M[\omega(0) = 0, \omega = 1 \text{ on } V_n], \\
\pi(0)R_n^-(0) &= \mu_M[\omega(0) = 0, \omega = 0 \text{ on } V_n], \\
\pi(1)R_n^+(1) &= \mu_M[\omega(0) = 1, \omega = 1 \text{ on } V_n], \\
\pi(1)R_n^-(1) &= \mu_M[\omega(0) = 1, \omega = 0 \text{ on } V_n].
\end{aligned}$$

Then

(25)
$$\rho_n^+ = \left\{1 + \frac{\pi(0)}{\pi(1)}\left(\frac{R_n^+(0)}{R_n^+(1)}\right)^{N+1}\right\}^{-1}$$
$$\rho_n^- = \left\{1 + \frac{\pi(0)}{\pi(1)}\left(\frac{R_n^-(0)}{R_n^-(1)}\right)^{N+1}\right\}^{-1}.$$

The definition of μ_M shows that

(26)
$$\begin{aligned}
R_{n+1}^+(1) &= t[R_n^+(1)]^N + (1-t)[R_n^+(0)]^N \\
R_{n+1}^+(0) &= (1-s)[R_n^+(1)]^N + s[R_n^+(0)]^N,
\end{aligned}$$

with two similar recursion formulas for R_n^-. If we define

(27)
$$r_n^+ = \frac{R_n^+(1)}{R_n^+(0)}, \qquad r_n^- = \frac{R_n^-(1)}{R_n^-(0)},$$

then (26) shows that, for φ defined as in (7),

(28)
$$r_{n+1}^+ = \varphi(r_n^+), \qquad r_{n+1}^- = \varphi(r_n^-).$$

Now it follows from (25) and (28) that (24) will hold provided

(29)
$$x_{n+1} = \varphi(x_n), \qquad n \geqq 0 \Rightarrow \lim_{n\to\infty} \varphi(x_n) = 1 \qquad \text{for every } x_0 > 0.$$

But (29) is true when $\varphi(x) = x$ has only the positive root $x = 1$. In fact $\varphi(x_n) \searrow 1$ when $x_0 > 0$ and $\varphi(x_n) \nearrow 1$ when $x_0 < 1$ because $\varphi(x) \nearrow$ as $x \nearrow$. Hence $\mathscr{G}_M = \{\mu_M\}$. □

PROOF OF THEOREMS 7 AND 8. The proof is divided into three parts. In Part I we show that $|\mathscr{G}_M| > 1$ when $\varphi \circ \varphi(u) = u$ has a positive root $u \neq 1$, and that

$\mathscr{G}_M$ will then contain a probability measure μ_{M^e,M^0} with $M^e \neq M^0$. In Part II we assume that $\varphi \circ \varphi(u) = u$ has only the positive root $u = 1$, and show that $\mathscr{G}_M = \{\mu_M\}$. Finally, in Part III we take $N = 2$ and show that $\varphi \circ \varphi(u) = u$ has a positive root $u \neq 1$ if and only if $s + t < \frac{1}{2}$.

PART I. Let us assume M is defined by (6) with conditional probabilities given by (2). Let us call the probability measures defined in Definition (8) even-odd Markov chains (EOMC's). We begin by looking for an EOMC with the same conditional probabilities as μ_M. The proof will be complete when we show that there is one if and only if $\varphi \circ \varphi(u) = u$ has a positive root $u \neq 1$. The computation will be essentially the same as in (16) through (21). We get

$$\tag{30} \begin{aligned} \alpha_{N+1} &= (1 + yx^{N+1})^{-1} = \left[1 + \frac{\pi^e(0)}{\pi^e(1)}\left(\frac{M^e(0,1)}{M^e(1,1)}\right)^{N+1}\right]^{-1} \\ \alpha_N &= (1 + yx^{N-1})^{-1} = \left[1 + \frac{\pi^e(0)}{\pi^e(1)}\left(\frac{M^e(0,1)}{M^e(1,1)}\right)^{N} \frac{M^e(0,0)}{M^e(1,0)}\right]^{-1} \end{aligned}$$

and two more equations with M^e, π^e replaced by M^0, π^0. If we define

$$\tag{31} \frac{M^e(1,0)}{M^e(1,1)} = \xi\,, \qquad \frac{M^e(0,1)}{M^e(0,0)} = \eta\,, \qquad \frac{M^0(1,0)}{M^0(1,1)} = \tilde{\xi}\,, \qquad \frac{M^0(0,1)}{M^0(0,0)} = \tilde{\eta}\,,$$

then (30) becomes, after some algebra,

$$\tag{32} \begin{gathered} x^2 = \xi\eta = \tilde{\xi}\tilde{\eta}\,, \\ \tilde{\eta} y x^{N-1} = \left[\frac{\eta + x^2}{\eta + 1}\right]^N, \qquad \eta y x^{N-1} = \left[\frac{\tilde{\eta} + x^2}{\tilde{\eta} + 1}\right]^N. \end{gathered}$$

Note that this is the analogue of (18). Just as was done there, use (12) to express yx^{N-1} and x^2 in terms of s and t, and define

$$\tag{33} u = \left[\frac{\eta s}{1 - s}\right]^{1/N}, \qquad \tilde{u} = \left[\frac{\tilde{\eta} s}{1 - s}\right]^{1/N}.$$

Then (12), (32), and (33) yield

$$\tag{34} \tilde{u} = \varphi(u)\,, \qquad u = \varphi(\tilde{u})\,.$$

It follows that we have found an EOMC if and only if (34) has a solution with $u > 0$, $\tilde{u} > 0$, $u \neq \tilde{u}$. This happens if and only if $\varphi \circ \varphi(u)$ has a positive solution $u \neq 1$. □

PART II. This part is analogous to the proof of Theorem 6. We assume

$$\tag{35} \varphi \circ \varphi(u) = u\,, \qquad u > 0 \Rightarrow u = 1\,,$$

and that M is defined by (6) with $s + t < 1$. Thus μ_M is a Gibbs state with self potential u_0 and repulsive pair potential $u_1 > 0$. The proof will depend on the mapping $\rho : \Omega \to \Omega$ defined by

$$\rho \circ \omega(x) = \omega(x)\,, \qquad x \in E\,, \qquad \rho \circ \omega(x) = 1 - \omega(x)\,, \qquad x \in \mathscr{O}\,,$$

where E are the even sites (containing the origin of T_N) and $\mathscr{O}$ the odd sites. The point of this mapping is that it transforms μ_M, by the formula

$$(\rho \circ \mu_M, f) = (\mu_M, f \circ \rho)\,, \qquad f \in C(\Omega)$$

into a Gibbs state $\rho \circ \mu_M = \mu_M'$ which is again a nearest neighbor Gibbs state. Thus its potential U' (cf. [8] page 56) satisfies $U'(x, y) = u_1' = -u_1 \leqq 0$ when x is a neighbor of y. Thus μ_M' is an attractive Gibbs state. Its self potential $u_0'(x)$ is non-homogeneous, but this does not affect the theorem, used in the proof of Theorem 6, that there is a unique Gibbs state with potential U' if and only if the one point probabilities are the same in the limit, whether one uses the boundary condition $\omega \equiv 1$ or $\omega \equiv 0$ on $\partial\Lambda_n$. But of course there is a unique Gibbs state for U' if and only if there is a unique one for U. We shall carry out the evaluation with μ_M instead of μ_M'. Then we must take for ρ_n^+ the μ_M-probability that $\omega(0) = 1$ with the boundary condition that $\omega \equiv 1$ on $\partial\Lambda_n$ when n is even and $\omega \equiv 0$ on $\partial\Lambda_n$ when n is odd. In the definition of ρ_n^-, 0 and 1 are reversed.

The recursion formula (26) now becomes

$$\begin{aligned} R_{n+1}^+(1) &= t[R_n^-(1)]^N + (1-t)[R_n^-(0)]^N \\ R_{n+1}^+(0) &= (1-s)[R_n^-(1)]^N + s[R_n^-(0)]^N \end{aligned} \tag{36}$$

and two more equations with $+$ and $-$ interchanged. Let us define r_n^+ and r_n^- exactly as in (27). Then one obtains, just as in (25),

$$\begin{aligned} \rho_n^+ &= \left[1 + \frac{\pi(0)}{\pi(1)}(r_n^+)^{-N}\right]^{-1} \\ \rho^{n-} &= \left[1 + \frac{\pi(0)}{\pi(1)}(r_n^-)^{-N}\right]^{-1}, \end{aligned} \tag{37}$$

while (36) gives

$$r_{n+1}^+ = \varphi(r_n^-)\,, \qquad r_{n+1}^- = \varphi(r_n^+)\,. \tag{38}$$

Thus ρ_n^+ and ρ_n^- will have the same limit (and hence $|\mathscr{G}_M| = 1$) provided

$$\begin{aligned} &a_{n+1} = \varphi(b_n)\,, \quad b_{n+1} = \varphi(a_n)\,, \quad n \geqq 0 \\ &\Rightarrow \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 1 \quad \text{for every pair} \quad a_0 > 0\,, \quad b_0 > 0\,. \end{aligned} \tag{39}$$

To prove (39) observe that $a_{n+1} = \varphi(b_n)$ and $b_{n+1} = \varphi(a_n)$ implies

$$a_{n+2} = \varphi \circ \varphi(a_n)\,, \qquad b_{n+2} = \varphi \circ \varphi(b_n)\,. \tag{40}$$

Also $s + t < 1$ implies that $\varphi(x)$ is strictly decreasing for $x > 0$, so that $\varphi \circ \varphi$ is strictly increasing. Thus (35) and (40) imply

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 1\,. \qquad \square$$

Part III. We assume $N = 2$ and investigate the positive roots of $\varphi \circ \varphi(x) = x$. The set of zeros of $\varphi \circ \varphi(x) - x$ contains the set of zeros of $\varphi(x) = x$. Therefore $\varphi \circ \varphi(x) - x$ must contain as a factor the cubic polynomial

$$P(x) = [(1-s)x^2 + s][x - \varphi(x)]\,.$$

See (22) for an explicit formula for $P(x)$. We may write

$$\varphi \circ \varphi(x) - x = [(1-s)(tx^2 + 1 - t)^2 + s((1-s)x^2 + s)^2]^{-1}Q(x)$$

where Q is a polynomial of degree 5. Hence P is a factor of Q, and one may verify that

$$Q(x) = P(x)R(x)\,, \qquad R(x) = x^2(t^2 + s - s^2) + x(t + s - 1) + s^2 + t - t^2\,.$$

Since $s + t < 1$ we know that P has no positive zeros other than $x = 1$. Hence every positive root of $\varphi \circ \varphi(x) = x$ with $x \neq 1$ must be a zero of $R(x)$. The zeros of $R(x)$ are given by

$$(41) \qquad x = \frac{1 - s - t}{2(t^2 + s - s^2)} \pm \frac{1}{2(t^2 + s - s^2)}[(2s + 2t - 1)[(s - t)^2(2s + 2t - 1) - 1]]^{\frac{1}{2}}\,.$$

In the region $s > 0$, $t > 0$, $s + t < 1$ the discriminant in (40) is positive if and only if $s + t < \frac{1}{2}$ and zero exactly when $s + t = \frac{1}{2}$. The latter case gives $x = 1$. Therefore $\varphi \circ \varphi(x) = x$ has a positive zero $x \neq 1$ if and only if $s + t < \frac{1}{2}$. □

REFERENCES

[1] CHUNG, K. L. (1968). *A Course in Probability Theory*. New York.

[2] DOBRUSHIN, R. L. (1968a). Description of a random field by means of conditional probabilities and the conditions governing its regularity. *Theor. Probability Appl.* **10** 193–213.

[3] DOBRUSHIN, R. L. (1968b). The problem of uniqueness of a Gibbsian random field, and the problem of phase transition. *Functional Anal. Appl.* **2** 302–312.

[4] GALLAVOTTI, G. (1972). Instabilities and phase transition in the Ising model. *Riv. del Nuovo Cimento* **2** 133–169.

[5] GEORGII, H. O. (1973). Stochastische Felder und ihre Anwendung auf Interaktionssysteme. Preprint, Heidelberg.

[6] LANFORD, D. E. and RUELLE, D. (1969). Observables at infinity, and states with short range correlations in statistical mechanics. *Comm. Math. Phys.* **13** 194–215.

[7] LEBOWITZ, J. L. and MARTIN-LÖF, A. (1972). On the uniqueness of the equilibrium state for Ising spin systems. *Comm. Math. Phys.* **25** 276–282.

[8] PRESTON, C. J. (1974). *Gibbs States on Countable Sets*. Cambridge Univ. Press.

[9] ROSENBLATT, M. (1962). *Random Processes*. Oxford Univ. Press.

[10] SPITZER, F. (1971). Random fields and interacting particle systems. MAA lecture notes, Williamstown.

[11] SPITZER, F. (1974). St. Flour Lectures, *Springer Lecture notes in Math.* **390**.

DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850