ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 4, Number 3, Summer 1974

# RECENT RESULTS ON THE STOCHASTIC ISING MODEL*

RICHARD HOLLEY

1. **Introduction.** The stochastic Ising model was first suggested by R. Glauber [7] as a simple model for studying the time evolution of the configuration of spins in a piece of iron in a heat bath. To describe the model we let **Z** be the integers and represent a configuration of spins by a function $\eta : \mathbf{Z}^3 \to \{-1, 1\}$. The interpretation is that if $\eta(x) = 1$ $(-1)$ the spin at $x$ is up (down). Notice that we are approximating a piece of iron, which has a very large but finite number of spin sites, with a model that has an infinite number of spin sites. This is a standard practice called taking the infinite volume or thermodynamic limit. The infinite number of spin sites in the model causes some technical difficulties but at the same time makes the model much more interesting from a mathematical point of view.

We let each of the spins interact with its neighbors in the following way. Let $U(x, \eta)$ be given by the formula $U(x, \eta) = -\eta(x)[\sum_y \eta(y) + H]$, where the summation is over those $y$ such that $|x - y| = 1$. $U(x, \eta)$ is to be thought of as the energy at the site $x$ in configuration $\eta$. The parameter $H$ is supposed to represent the external magnetic field. The idea is to have the spins at each site flipping back and forth, and the rate of flipping is to depend on the energy—high energy giving a high flip rate and low energy a low flip rate. Thus we let $c(x, \eta)$ be the rate that the spin at site $x$ flips when the entire configuration is $\eta$ and assume that $c(\cdot, \cdot)$ satisfies

(1.1) $c(x, \eta) = F(U(x, \eta))$ for some increasing function $F$, and

(1.2) $c(x, \eta)e^{-\beta U(x, \eta)} = c(x, {}_x\eta)e^{-\beta U(x, {}_x\eta)}$,

where $\beta > 0$ represents the reciprocal of the temperature and ${}_x\eta$ is the configuration given by

$$ {}_x\eta(y) = \begin{cases} \eta(y) \text{ if } y \neq x \\ -\eta(x) \text{ if } y = x. \end{cases} $$

One obvious choice of $F$ is $F(z) = e^{\beta z}$ and the one which is usually used in the physics literature is $F(z) = 1/(1 + e^{-2\beta z})$. The condition (1.2) is just a technical one to guarantee that the Gibbs states (which

*This work was partially supported by NSF Grant GP33314X.
Received by the Editors June 22, 1973.

will be described below) are equilibrium states for the process.

It can be shown (see [8] for a simple proof in this situation or [13] for a more general theorem) that there is a standard Markov process, $\eta_t$, whose state space is the set of all configurations of spins (denoted by $E$ in the future) and which is such that

$$P_\eta(\eta_t(x) = -\eta(x)) = tc(x, \eta) + o(t) \tag{1.3}$$

and

$$P_\eta(\eta_t(x) = -\eta(x) \text{ and } \eta_t(y) = -\eta(y)) = o(t) \tag{1.4}$$

if $x \neq y$. Here $P_\eta(\cdot)$ is the probability when the initial configuration is $\eta$. The Markov process $\eta_t$ is the stochastic Ising model.

Notice that the temperature, $1/\beta$, is held constant and that the process does not conserve energy. This is justified by saying that the process is in a heat bath which holds the temperature constant and exchanges energy in the form of heat.

The stochastic Ising model has received quite a bit of attention in the physics literature (see [4] for a bibliography). Most of the work there has been restricted to one dimensional configurations and is concerned with getting exact rates of convergence to equilibrium. For this type of problem it is obviously crucial which function $F$ one chooses. We will be concerned here with three dimensional configurations and will get qualitative results. For these results conditions (1.1) and (1.2) are all that we assume. The exact choice of the increasing function $F$ in (1.1) plays no role.

Most of the results in this paper have already appeared elsewhere. The only things that are new are Theorem (2.4) and the examples in section four. In section two we generalize the concept of one distribution being stochastically larger than another to lattices and prove a theorem which illustrates the main technique used in this paper. We then apply the theorem to derive some of the facts which we need about Gibbs states. Section three contains the ergodic theorem for the stochastic Ising model at high temperatures or nonzero magnetic field, and section four has some examples of what can happen at low temperatures and zero magnetic field.

2. **Probability distributions on lattices.** In this section we consider a finite lattice, $\Gamma$, and develop a way of deciding when one probability density is higher up on $\Gamma$ than another. To motivate this we suppose first that $\Gamma$ is a finite subset of the real line and that $\mu_1$ and $\mu_2$ are probability densities on $\Gamma$. Then $\mu_1$ is stochastically larger than $\mu_2$ (we write $\mu_1 >^s \mu_2$ in the future) if

$$\sum_{\substack{\alpha \leqq x \\ \alpha \in \Gamma}} \mu_1(\alpha) \leqq \sum_{\substack{\alpha \leqq x \\ \alpha \in \Gamma}} \mu_2(\alpha) \quad \text{for all } x \in \Gamma. \tag{2.1}$$

One reason that the concept of one density being stochastically larger than another is important is that if $\mu_1 >^s \mu_2$, then

$$\sum_{\alpha\in\Gamma} f(\alpha)\mu_1(\alpha) \geqq \sum_{\alpha\in\Gamma} f(\alpha)\mu_2(\alpha) \tag{2.2}$$

for all increasing functions $f$. In generalizing the definition from chains to lattices we want a definition which will imply (2.2). When $\Gamma$ is a general finite lattice (2.1) no longer implies (2.2), so a different definition is needed. The following definition is equivalent to (2.1) when $\Gamma$ is a finite chain and, as seen below, easily implies (2.2).

(2.3) Definition. Let $\Gamma$ be a finite lattice and let $\mu_1$ and $\mu_2$ be probability densities on $\Gamma$. Then $\mu_1$ is stochastically larger than $\mu_2$ ($\mu_1 >^s \mu_2$) if there is a density $\mu$ on $\Gamma \times \Gamma$ such that

(i) $\sum_{y\in\Gamma}\mu(x, y) = \mu_1(x)$ for all $x \in \Gamma$,

(ii) $\sum_{x\in\Gamma}\mu(x, y) = \mu_2(y)$ for all $y \in \Gamma$,

(iii) $\mu(x, y) = 0$ unless $x \geqq y$.

If $\mu_1 >^s \mu_2$ then it is clear that $\mu_1$ is higher up on the lattice than $\mu_2$. Indeed $\mu_2$ can be constructed from $\mu_1$ by dividing the mass $\mu_1(x)$ into pieces $\mu(x, y)$ for $y \leqq x$ and then moving the piece $\mu(x, y)$ down the lattice to $y$. If this is done for each $x \geqq y$, the resulting mass at $y$ will be $\mu_2(y)$.

To see how definition (2.3) implies (2.2) we have

$$\begin{aligned}\sum_{x\in\Gamma} f(x)\mu_1(x) &= \sum_{x,y\in\Gamma} f(x)\mu(x, y)\\ &= \sum_{x\geqq y} f(x)\mu(x, y) \geqq \sum_{x\geqq y} f(y)\mu(x, y)\\ &= \sum_{x,y\in\Gamma} f(y)\mu(x, y) = \sum_{y\in\Gamma} f(y)\mu_2(y).\end{aligned}$$

The more difficult problem is obviously to determine when one density is stochastically larger than another. In case the lattice $\Gamma$ is distributive the following theorem gives a sufficient condition. (Recall that $\Gamma$ is distributive if $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ or equivalently $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.)

(2.4) Theorem. *Let $\Gamma$ be a finite distributive lattice and let $\mu_1$ and $\mu_2$ be two strictly positive probability densities on $\Gamma$ such that*

$$\mu_1(x \vee y)\mu_2(x \wedge y) \geqq \mu_1(x)\mu_2(y). \tag{2.5}$$

*Then $\mu_1$ is stochastically larger than $\mu_2$.*

The proof is preceded by two lemmas.

(2.6) Lemma. *Let $\Gamma$ be a finite distributive lattice. Then $\Gamma$ is isomorphic to a sublattice, $\tilde{\Gamma}$, of the lattice of subsets of a finite set $\Lambda$.*

*Moreover, $\Lambda$ and $\tilde{\Gamma}$ can be chosen in such a way that $\phi$ and $\Lambda$ are in $\tilde{\Gamma}$ and for all $A, B \in \tilde{\Gamma}$ there is a sequence $A = A_0, A_1, \cdots, A_n = B$ in $\tilde{\Gamma}$ such that $|A_i \triangle A_{i+1}| = 1$ for all $i$. Here $|A|$ denotes the cardinality of $A$, and $A \triangle B = (A \backslash B) \cup (B \backslash A)$.*

Lemma (2.6) is essentially Corollary (2), page 59 in [1]. Actually Corollary (2) in [1] is not phrased this way; however, it is easily seen from the proof to be equivalent to Lemma (2.6).

Until the theorem is proved $\Lambda$ will be a fixed finite set and $\Gamma$ will be a sublattice of the lattice of subsets of $\Lambda$. It will also be assumed that $\Gamma$ has the properties of $\tilde{\Gamma}$ mentioned in Lemma (2.6).

Let $\nu$ be a strictly positive probability density on $\Gamma$. For $x \in \Lambda$ and $A \in \Gamma$ define a function $c(x, A)$ as follows:

$$c(x, A) = \begin{cases} 1 \text{ if } x \notin A \text{ and } A \cup \{x\} \in \Gamma, \\ \nu(A \backslash \{x\})/\nu(A) \text{ if } x \in A \text{ and } A \backslash \{x\} \in \Gamma, \\ 0 \text{ otherwise.} \end{cases} \tag{2.7}$$

If $A, B \in \Gamma$ and $A \neq B$ define

$$\Omega(A, B) = \begin{cases} c(x, A) \text{ if } A \triangle B = \{x\}, \\ 0 \text{ otherwise,} \end{cases} \tag{2.8}$$

and define $\Omega(A, A)$ so that $\sum_{B \in \Gamma} \Omega(A, B) = 0$. We think of $\Omega$ as a matrix and let $\Omega^n$ be the $n$th power of that matrix.

Now let

$$P_t(A, B) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \Omega^n(A, B).$$

Since $\Omega(A, B) \geqq 0$ if $A \neq B$ and $\sum_{B \in \Gamma} \Omega(A, B) = 0$, it follows that $P_t(\cdot, \cdot)$ is the transition function of a Markov process on $\Gamma$. The matrix $\Omega$ is the generator of the process.

(2.9) Lemma. *The Markov process on $\Gamma$ with generator $\Omega$ has $\nu$ as a stationary distribution (i.e., $\sum_{A \in \Gamma} \nu(A) P_t(A, B) = \nu(B)$ for all $t \geqq 0$). Moreover, this Markov process has only one stationary distribution.*

Proof. One easily checks from the definition of $\Omega$ that $\sum_{A \in \Gamma} \nu(A) \Omega(A, B) = 0$ for all $B \in \Gamma$. The stationarity of $\nu$ follows immediately from this. The Markov process has only one stationary distribution if $P_t(A, B) > 0$ for all $A, B \in \Gamma$ and all $t > 0$. In order to prove that $P_t(A, B) > 0$ for all $A, B \in \Gamma$ and all $t > 0$, it suffices to show that for all $A, B \in \Gamma$ there is a sequence $A = A_0, A_1, \cdots, A_n = B$ in $\Gamma$ such that $\prod_{i=1}^{n} \Omega(A_{i-1}, A_i) > 0$. But again this follows from the definition of $\Omega$ and the assumed structure of $\Gamma$.

The essential idea of the proof of Theorem (2.4) is to couple together two Markov processes having generators constructed as above using $\mu_1$ and $\mu_2$ instead of $\nu$. The coupling is done in such a way that the inequality (2.5) implies that if one of the coupled processes is ever higher up on the lattice than the other, then from that time on it stays higher on the lattice. This technique of coupling two processes together is also used in section three. We give the details here but only outline the later proofs.

PROOF OF THEOREM (2.4). Let $c_i(x, A)$ be defined as in (2.7) with $\mu_i$ in place of $\nu$ and let $\Omega_i$ be defined as in (2.8) with $c_i$ in place of $c$. For $A_1, A_2, B_1, B_2 \in \Gamma$ with either $A_1 \neq B_1$ or $A_2 \neq B_2$ define

$$\bar{\Omega}(A_1, A_2; B_1, B_2) =$$

$$= \begin{cases} \min(c_1(x, A_1), c_2(x, A_2)) \text{ if } x \in (A_1 \cap A_2) \cup (A_1^c \cap A_2^c) \\ \qquad\qquad \text{and } A_1 \triangle B_1 = A_2 \triangle B_2 = \{x\}, \\ c_1(x, A_1) - \min(c_1(x, A_1), c_2(x, A_2)) \text{ if } x \in (A_1 \cap A_2) \cup (A_1^c \cap A_2^c) \\ \qquad\qquad \text{and } A_1 \triangle B_1 = \{x\}, A_2 = B_2, \\ c_2(x, A_2) - \min(c_1(x, A_1), c_2(x, A_2)) \text{ if } x \in (A_1 \cap A_2) \cup (A_1^c \cap A_2^c) \\ \qquad\qquad \text{and } A_1 = B_1, A_2 \triangle B_2 = \{x\}, \\ \Omega_1(A_1, B_1) \text{ if } A_1 \triangle B_1 \subset A_1 \triangle A_2 \text{ and } A_2 = B_2, \\ \Omega_2(A_2, B_2) \text{ if } A_2 \triangle B_2 \subset A_1 \triangle A_2 \text{ and } A_1 = B_1, \\ 0 \text{ otherwise.} \end{cases}$$

Define $\bar{\Omega}(A_1, A_2; A_1, A_2)$ so that

$$\sum_{B_1, B_2 \in \Gamma} \bar{\Omega}(A_1, A_2; B_1, B_2) = 0.$$

We think of $\bar{\Omega}(A_1, A_2; B_1, B_2)$ as a matrix with rows indexed by $(A_1, A_2)$ and columns indexed by $(B_1, B_2)$. Then $\bar{\Omega}^n$ is just the matrix $\bar{\Omega}$ raised to the $n$th power. Similarly $\Omega_i^n$ is the matrix $\Omega_i$ raised to the $n$th power.

The following facts are easily checked by induction on $n$:

(i) $\sum_{B_2 \in \Gamma} \bar{\Omega}^n(A_1, A_2; B_1, B_2) = \Omega_1^n(A_1, B_1)$ for all $A_1, B_1 \in \Gamma$.

(ii) $\sum_{B_1 \in \Gamma} \bar{\Omega}^n(A_1, A_2; B_1, B_2) = \Omega_2^n(A_2, B_2)$ for all $A_2, B_2 \in \Gamma$.

(iii) If $A_1 \supset A_2$, then $\Omega^n(A_1, A_2; B_1, B_2) = 0$ unless $B_1 \supset B_2$.

In checking (iii) one needs the lattice structure of $\Gamma$ to guarantee for example that if $A_1 \supset A_2$ and $x \notin A_1$ but $A_2 \cup \{x\} \in \Gamma$, then $A_1 \cup \{x\} \in \Gamma$. Also the only place where the inequality (2.5) is used is in checking (iii).

Now $\bar{\Omega}$ is the generator of a Markov process on $\Gamma \times \Gamma$. We take $(\Lambda, \phi)$ as the initial state and apply the ergodic theorem to this finite state space Markov process to conclude the existence of

$$\mu(B_1, B_2) = \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \bar{\Omega}^n(\Lambda, \phi; B_1, B_2). \tag{2.10}$$

The ergodic theorem together with Lemma (2.9) applied to the Markov processes on $\Gamma$ with generators $\Omega_i$ imply that

$$\begin{aligned} \mu_1(B_1) &= \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \Omega_1{}^n(\Lambda, B_1) \\ \mu_2(B_2) &= \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \Omega_2{}^n(\phi, B_2). \end{aligned} \tag{2.11}$$

Since $\Lambda \supset \phi$, (iii) applied to (2.10) implies that $\mu(B_1, B_2) = 0$ unless $B_1 \supset B_2$.

Finally by applying (i), (ii) and (2.11) we get

$$\begin{aligned} \sum_{B_2 \in \Gamma} \mu(B_1, B_2) &= \sum_{B_2} \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \bar{\Omega}^n(\Lambda, \phi; B_1, B_2) \\ &= \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{B_2} \bar{\Omega}^n(\Lambda, \phi; B_1, B_2) \\ &= \lim_{t\to\infty} \sum_{n=0}^{\infty} \frac{t^n}{n!} \Omega_1{}^n(\Lambda, B_1) = \mu_1(B_1). \end{aligned}$$

Similarly $\sum_{B_1 \in \Gamma} \mu(B_1, B_2) = \mu_2(B_2)$.

This completes the proof of the theorem.

The rest of this section is devoted to applications of Theorem (2.4). The results we get are well known in the physics literature. They are usually obtained by applying Corollary (2.12), which was first proved by Fortuin, Kasteleyn, and Ginibre [5] using entirely different methods. However it is slightly easier to get most of the results by using Theorem (2.4) directly as we do.

(2.12) Corollary. (FKG) *Let $\Gamma$ be a finite distributive lattice and let $\mu$ be a probability density on $\Gamma$ satisfying*

(2.13) $$\mu(x \vee y)\mu(x \wedge y) \geqq \mu(x)\mu(y).$$

*Then if $f$ and $g$ are two increasing functions on* $\Gamma$,

(2.14) $$\sum_{x\in\Gamma} f(x)g(x)\mu(x) \geqq \sum_{x\in\Gamma} f(x)\mu(x) \sum_{y\in\Gamma} g(y)\mu(y).$$

PROOF. Inequality (2.13) implies that the $x \in \Gamma$ for which $\mu(x) > 0$ form a sublattice. By restricting our attention to this sublattice we may assume, without loss of generality, that $\mu$ is strictly positive. Also by adding a constant if necessary, we may assume that $g > 0$.

Now let $\mu_1(x) = g(x)\mu(x)/\sum_{y\in\Gamma} g(y)\mu(y)$ and let $\mu_2(y) = \mu(y)$.

Using (2.13) and the monotonicity of $g$ one easily checks that $\mu_1$ and $\mu_2$ satisfy (2.5). Thus $\mu_1 >^s \mu_2$. Applying (2.2) to $\mu_1$, $\mu_2$, and $f$ we conclude that

$$\sum_{x\in\Gamma} f(x)g(x)\mu(x) \Big/ \sum_{y\in\Gamma} g(y)\mu(y) \geqq \sum_{x\in\Gamma} f(x)\mu(x),$$

which is the same as (2.14) since $g$ is positive.

We now apply Theorem (2.4) to obtain some information about the equilibrium states (stationary distributions) of the stochastic Ising model. If $\{-1, 1\}$ is given the discrete topology and $E = \{-1, 1\}^{Z^3}$ is thought of as an infinite product of $\{-1, 1\}$ and is given the product topology, then the equilibrium states are certain probability measures on the Borel sets of $E$. They can be defined in several ways. The nicest way uses conditional probabilities and is due to Dobrushin [2]. We give here an operational definition.

Let $\Lambda$ be any finite subset of $Z^3$ and let $\partial\Lambda = \{y \in Z^3 \mid y \notin \Lambda$ but there is an $x \in \Lambda$ such that $|x - y| = 1\}$. If $\varphi : \partial\Lambda \to \{-1, 1\}$, we define a probability density $\mu_\Lambda{}^\varphi$ on $\{-1, 1\}^\Lambda$ by the formula

(2.15) $$\mu_\Lambda{}^\varphi(\varphi) = Z^1(\varphi, \Lambda)$$
$$\cdot \exp\Big\{ -\beta\Big[ -\sum_{\{x,y\}\subset\Lambda} \sigma(x)\sigma(y) - \sum_{\substack{\{x,y\}\\ x\in\Lambda\\ y\notin\Lambda}} \sigma(x)\varphi(y) - H \sum_{x\in\Lambda} \sigma(x) \Big]\Big\}.$$

Here the summations over $\{x, y\}$ are over the indicated pairs $x$ and $y$ with $|x - y| = 1$, and $Z(\varphi, \Lambda)$ is the normalizing constant which makes $\mu_\Lambda{}^\varphi$ a probability density. We also let $\mu_\Lambda{}^\varphi$ denote the probability measure on the Borel sets of $E$ which puts all of its mass on the finite set

$$A(\varphi, \Lambda) = \{\eta \mid \eta(x) = 1 \text{ if } x \in Z^3\backslash(\Lambda \cup \partial\Lambda) \text{ and } \eta(x) = \varphi(x) \text{ if } x \in \partial\Lambda\}$$

and has its density given by (2.15). If $f \in \mathcal{C}(E)$ (the continuous real valued functions on $E$), then $\langle f, \mu_\Lambda{}^\varphi \rangle$ will denote

$$\sum_{\eta \in A(\varphi,\Lambda)} f(\eta)\mu_\Lambda{}^\varphi(\eta) = \int_E f(\eta)\mu_\Lambda{}^\varphi(d\eta).$$

The equilibrium states for the stochastic Ising model are then all of the probability measures on the Borel sets of $E$ which can be constructed as follows. Let $\{\Lambda_n\}$ be an increasing sequence of finite subsets of $\mathbf{Z}^3$ whose union is all of $\mathbf{Z}^3$. For each $n$ let $\mu_n$ be a convex combination of the $\mu_{\Lambda_n}{}^\varphi$ 's, where the averaging is done over all possible $\varphi : \partial\Lambda_n \to \{-1, 1\}$. Since $E$ is compact there is at least one subsequence $\mu_{n_k}$, such that $\mu_{n_k}$ converges weakly to some probability measure. The set of all such weak limits is the set of equilibrium states. (Actually this last statement has not been completely proven. What is known is that the weak limits described above are always contained in the set of equilibrium states — see [**16**] and [**9**] — and as will be shown below they are actually equal in most cases. However at low temperatures and zero magnetic field the question is still open.) This set of weak limits is also called the set of infinite Gibbs states in the physics literature.

There are two choices of the sequence $\{\mu_n\}$ that play a particularly important role. They are the sequences that arise when all of the boundary values are either identically $+1$ or identically $-1$. If $\varphi$ is identically $+1$ we denote $\mu_\Lambda{}^\varphi$ by $\mu_\Lambda{}^+$, and if $\varphi$ is identically $-1$ we denote $\mu_\Lambda{}^\varphi$ by $\mu_\Lambda{}^-$.

In order to apply Theorem (2.4) to the densities $\mu_\Lambda{}^\varphi$ we must identify them as densities on a finite distributive lattice. The lattice is the set of functions from $\Lambda$ into $\{-1, 1\}$ with $\sigma \geqq \sigma'$ if and only if $\sigma(x) \geqq \sigma'(x)$ for all $x \in \Lambda$. It is easily checked that this lattice is distributive. Also from (2.15) it is clear that the density $\mu_\Lambda{}^\varphi$ is strictly positive.

(2.16) Corollary. *For all* $\varphi : \partial\Lambda \to \{-1, 1\}$

$$\mu_\Lambda{}^+ >^s \mu_\Lambda{}^\varphi \text{ and } \mu_\Lambda{}^\varphi >^s \mu_\Lambda{}^-. \tag{2.17}$$

Proof. We prove only the first inequality in (2.17). According to Theorem (2.4) it suffices to check that

$$\mu_\Lambda{}^+(\sigma \vee \sigma')\mu_\Lambda{}^\varphi(\sigma \wedge \sigma') \geqq \mu_\Lambda{}^+(\sigma)\mu_\Lambda{}^\varphi(\sigma'). \tag{2.18}$$

By first canceling the factors $Z(\varphi, \Lambda)Z(+, \Lambda)$, then taking logarithms, dividing $\beta$, and rearranging the summations, (2.18) is equivalent to

$$\sum_{\{x,y\}\subset\Lambda} [(\sigma \vee \sigma')(x)(\sigma \vee \sigma')(y) + (\sigma \wedge \sigma')(x)(\sigma \wedge \sigma')(y)$$

$$- \sigma(x)\sigma(y) - \sigma'(x)\sigma'(y)] \tag{2.19}$$

$$+ \sum_{\substack{\{x,y\} \\ x\in\Lambda \\ y\notin\Lambda}} [(\sigma \vee \sigma')(x) + (\sigma \wedge \sigma')(x)\varphi(y) - \sigma(x) - \sigma'(x)\varphi(y)]$$

$$+ H \sum_{x\in\Lambda} [ (\sigma \vee \sigma')(x) + (\sigma \wedge \sigma')(x) - \sigma(x) - \sigma'(x)] \geqq 0.$$

It is easily seen that each term in (2.19) is greater than or equal to zero; hence the entire summation is, and (2.19) is true.

In order to apply Corollary (2.16) we let S be a finite subset of $\mathbf{Z}^3$ and let $f_S$ be the function on $E$ defined by

$$f_S(\eta) = \begin{cases} 1 \text{ if } \eta(x) = 1 \text{ for all } x \in S, \\ 0 \text{ otherwise.} \end{cases}$$

Since $f_S$ is an increasing function on the lattice $E$, Corollary (2.16) implies that

$$(2.20) \qquad \langle f_S, \mu_\Lambda{}^+\rangle \geqq \langle f_S, \mu_\Lambda{}^\varphi \rangle \geqq \langle f_S, \mu_\Lambda{}^-\rangle.$$

(2.21) Corollary. *If* $S \subset \Lambda \subset \Lambda'$, *then*

$$\langle f_S, \mu_\Lambda{}^+\rangle \geqq \langle f_S, \mu_{\Lambda'}^+ \rangle \textit{ and } \langle f_S, \mu_\Lambda{}^-\rangle \leqq \langle f_S, \mu_{\Lambda'}^-\rangle.$$

Proof. We prove only the first inequality under the assumption that $\Lambda \cup \partial\Lambda \subset \Lambda'$. The second inequality and the modifications needed to handle the case when $\partial\Lambda \not\subset \Lambda'$ are left to the reader.

If $\sigma : \Lambda \to \{-1, 1\}$, $\varphi : \partial\Lambda \to \{-1, 1\}$ and $\psi : \mathbf{Z}^3\backslash(\Lambda \cup \partial\Lambda) \to \{-1, 1\}$, we let $[\psi, \varphi, \sigma]$ be the function which is equal to $\psi$ on $\mathbf{Z}^3\backslash(\Lambda \cup \partial\Lambda)$, $\varphi$ on $\partial\Lambda$, and $\sigma$ on $\Lambda$. Let $B(\Lambda', \Lambda)$ be the set of functions from $\mathbf{Z}^3\backslash(\Lambda \cup \partial\Lambda)$ into $\{-1, 1\}$ which are identically equal to 1 on $\mathbf{Z}^3\backslash\Lambda'$. Then

$$(2.22) \qquad \langle f_S, \mu_{\Lambda'}^+ \rangle = \sum_{\eta \in A(+,\Lambda')} f_S(\eta)\mu_{\Lambda'}^+(\eta)$$
$$= \sum_{\psi \in B(\Lambda',\Lambda)} \sum_{\varphi:\partial\Lambda\to\{-1,1\}} \sum_{\sigma:\Lambda\to\{-1,1\}} f_S([\psi, \varphi, \sigma])\mu_{\Lambda'}^+([\psi, \varphi, \sigma]).$$

Now if $\Theta(\psi, \varphi, \Lambda', \Lambda) = \sum_\sigma \mu_{\Lambda'}^+([\psi, \varphi, \sigma])$, then it follows immediately from the definition of $\mu_\Lambda{}^\varphi$ that

$$(2.23) \qquad \mu_{\Lambda'}^+([\psi, \varphi, \sigma]) = \mu_\Lambda{}^\varphi(\sigma)\Theta(\psi, \varphi, \Lambda', \Lambda).$$

Also $\sum_{\psi,\varphi} \Theta(\psi, \varphi, \Lambda', \Lambda) = 1$. Thus substituting (2.23) into (2.22) and using inequality (2.20) we have

$$\begin{aligned}\langle f_S, \mu^+_{\Lambda'}\rangle &= \sum_\psi\sum_\varphi\sum_\sigma f_S([\psi,\varphi,\sigma])\mu_\Lambda^\varphi(\sigma)\Theta(\psi,\varphi,\Lambda',\Lambda)\\ &= \sum_\psi\sum_\varphi \sum_{\eta\in A(\varphi,\Lambda)} f_S(\eta)\mu_\Lambda^\varphi(\eta)\Theta(\psi,\varphi,\Lambda',\Lambda)\\ &= \sum_\psi\sum_\varphi\langle f_S, \mu_\Lambda^\varphi\rangle\Theta(\psi,\varphi,\Lambda',\Lambda)\\ &\leqq \sum_\psi\sum_\varphi\langle f_S, \mu_\Lambda^{\ +}\rangle\Theta(\psi,\varphi,\Lambda',\Lambda) = \langle f_S, \mu_\Lambda^{\ +}\rangle.\end{aligned}$$

Notice that if $\langle f_S, \mu\rangle$ is known for all finite sets S, then the finite dimensional distributions of $\mu$ are uniquely determined, and hence $\mu$ is known. Hence if $\{\mu_n\}$ is a sequence of probability measures on the Borel sets of $E$, then to check whether or not $\{\mu_n\}$ converges weakly to $\mu$ it suffices to check whether or not $\langle f_S, \mu_n\rangle$ converges to $\langle f_S, \mu\rangle$ for all finite S. (We already know that the sequence is tight because $E$ is compact.) In view of this Corollaries (2.16) and (2.21) have several interesting consequences:

(2.24) Let $\{\Lambda_n\}$ be an increasing sequence of finite sets whose union is all of $\mathbf{Z}^3$, then there are measures $\mu^+$ and $\mu^-$ such that $\{\mu^+_{\Lambda_n}\}$ converges weakly to $\mu^+$ and $\{\mu^-_{\Lambda_n}\}$ converges weakly to $\mu^-$.

(2.25) The measures $\mu^+$ and $\mu^-$ are independent of the sequence $\{\Lambda_n\}$.

(2.26) If $\mu^+ = \mu^-$, then $\{\mu_{\Lambda_n}^{\varphi_n}\}$ converges weakly to $\mu^+$ for any choices of $\{\Lambda_n\}$ and $\{\varphi_n\}$.

(2.27) $\mu^+$ and $\mu^-$ satisfy $\langle f_S, \mu^+\rangle = \langle f_{S+a}, \mu^+\rangle$ and $\langle f_S, \mu^-\rangle = \langle f_{S+a}, \mu^-\rangle$ for all $a \in \mathbf{Z}^3$. Here $S + a = \{x + a \mid x \in S\}$.

Statements (2.24), (2.25), and (2.26) all follow immediately from the monotonicity implied by Corollaries (2.16) and (2.21). (2.27) is implied by (2.25) as follows. Pick any increasing sequence, $\{\Lambda_n\}$. Then

$$\langle f_S, \mu^+\rangle = \lim_{n\to\infty}\langle f_S, \mu^+_{\Lambda_n}\rangle = \lim_{n\to\infty}\langle f_{S+a}, \mu^+_{\Lambda_n+a}\rangle = \langle f_{S+a}, \mu^+\rangle.$$

Statement (2.27) implies that $\mu^+$ is invariant under shifts in $\mathbf{Z}^3$. Actually a stronger statement is true.

(2.28) Corollary. *$\mu^+$ is strongly mixing and hence ergodic.*

Proof. It suffices to show that $\lim_{|a|\to\infty}\langle f_S f_{T+a}, \mu^+\rangle = \langle f_S, \mu^+\rangle \cdot \langle f_T, \mu^+\rangle$ for all finite sets S and $T$. We indicate why this is so when $S = T = \{0\}$.

Just as in the proof of Corollary (2.16), Corollary (2.12) implies that

$$\langle f_{\{0\}}f_{\{a\}}, \mu_\Lambda^+\rangle \geqq \langle f_{\{0\}}, \mu_\Lambda^+\rangle\langle f_{\{a\}}, \mu_\Lambda^+\rangle$$

for any finite set $\Lambda$ containing 0 and $a$. Letting $\Lambda \uparrow \mathbf{Z}^3$ and using (2.27), it follows that $\langle f_{\{0\}}f_{\{a\}}, \mu^+\rangle \geqq \langle f_{\{0\}}, \mu^+\rangle^2$ for all $a \in \mathbf{Z}^3$.

Now let $\Lambda(a)$ be any finite set containing 0 and having $a \in \partial\Lambda(a)$. Then if $\Lambda \supset \Lambda(a)$ it follows from elementary computations using (2.15) that

$$(2.29) \quad \langle f_{\{0\}}f_{\{a\}}, \mu_\Lambda^+\rangle = \sum_{\substack{\varphi:\partial\Lambda(a)\to\{-1,1\}\\ \varphi(a)=1}} \langle f_{\{0\}}, \mu_{\Lambda(a)}^\varphi\rangle\mu_\Lambda^+(\{\eta : \eta(x) = \varphi(x) \text{ for all } x \in \partial\Lambda(a)\}).$$

Since $\mu^+ = \lim_{\Lambda\uparrow \mathbf{Z}^3}\mu_\Lambda^+$, taking the limit on $\Lambda$ in (2.29) yields

$$(2.30) \quad \langle f_{\{0\}}f_{\{a\}}, \mu^+\rangle = \sum_\varphi \langle f_{\{0\}}, \mu_{\Lambda(a)}^\varphi\rangle\mu^+(\{\eta \mid \eta(x) = \varphi(x) \text{ for all } x \in \partial\Lambda(a)\}).$$

We now use (2.20) and the observation that

$$\sum_{\substack{\varphi:\partial\Lambda(a)\to\{-1,1\}\\ \varphi(a)=1}} \mu^+(\{\eta \mid \eta(x) = \varphi(x) \text{ for all } x \in \partial\Lambda(a)\}) = \langle f_{\{a\}}, \mu^+\rangle = \langle f_{\{0\}}, \mu^+\rangle$$

to conclude from (2.30) that

$$\langle f_{\{0\}}f_{\{a\}}, \mu^+\rangle \leqq \langle f_{\{0\}}, \mu_{\Lambda(a)}^+\rangle\langle f_{\{0\}}, \mu^+\rangle.$$

Then by letting $\Lambda(a) \uparrow \mathbf{Z}^3$ as $|a| \to \infty$ we have

$$\limsup_{|a|\to\infty} \langle f_{\{0\}}f_{\{a\}}, \mu^+\rangle \leqq \langle f_{\{0\}}, \mu^+\rangle^2$$

and the proof is complete.

The rest of the results in this section are concerned with what happens when $\beta$ and $H$ are allowed to vary. Thus we need to include them in our notation. We do this by putting a subscript $\beta, H$ on the previous notation. As a result $\mu_\Lambda^\varphi$ becomes $\mu_{\Lambda,\beta,H}^\varphi$ and $\mu^+$ becomes $\mu_{\beta,H}^+$ etcetera.

(2.31) COROLLARY. *$\langle f_{\{0\}}, \mu_{\beta,H}^+\rangle$ is an increasing function of $H$.*

PROOF. Just as in the proof of Corollary (2.16) it can be shown that if $H' > H$ then $\mu_{\Lambda,\beta,H'}^+ >^s \mu_{\Lambda,\beta,H}^+$. Thus $\langle f_{\{0\}}, \mu_{\Lambda,\beta,H'}^+\rangle \geqq \langle f_{\{0\}}, \mu_{\Lambda,\beta,H}^+\rangle$, and the proof is completed by passing to the limit in $\Lambda$.

In fact it can be shown that $\langle f_{\{0\}}, \mu^+_{\beta,H}\rangle$ is a strictly increasing function of $H$, however a proof of this involves techniques which we have not talked about. The reader is referred to [14].

The following three facts are deeper than the ones which we have mentioned so far and their proofs are fairly involved. We omit the proofs. The reader is referred to [3], [12], [15], and [6].

(2.32) There is a critical number $\beta_c > 0$ such that if $\beta > \beta_c$ then $\mu^+_{\beta,0} \neq \mu^-_{\beta,0}$, and in fact $\langle f_{\{0\}}, \mu^-_{\beta,0}\rangle < 1/2 < \langle f_{\{0\}}, \mu^+_{\beta,0}\rangle$.

(2.33) If $\beta < \beta_c$ or if $H \neq 0$, then $\mu^+_{\beta,0} = \mu^-_{\beta,0}$. Henceforth if $\mu^+_{\beta,0} = \mu^-_{\beta,0}$ we will drop the $+$ or $-$ from the notation.

(2.34) There is a number $\beta' \geqq \beta_c$ such that if $\mu$ is a Gibbs state having parameters $\beta$ and $H$ with $\beta \geqq \beta'$ and $H = 0$ and if $\mu$ is also shift invariant (i.e., $\langle f_S, \mu\rangle = \langle f_{S+a}, \mu\rangle$ for all $S \subset \mathbf{Z}^3$ and all $a \in \mathbf{Z}^3$), then $\mu$ is a convex combination of $\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$.

3. **The Erogodic Theorem for the stochastic Ising model.** Again in this section we only sketch the proofs. For more detailed proofs the reader is referred to [10].

Let $\beta$ and $H$ be fixed and let $T_t$ denote the semigroup of operators on $\mathcal{C}(E)$ associated with the stochastic Ising model having parameters $\beta$ and $H$. (See [8] for a proof of the existence of $T_t$ and a proof that $T_t$ maps $\mathcal{C}(E)$ into itself.)

(3.1) THEOREM. *Let S be a finite subset of* $\mathbf{Z}^3$ *and let* $f_S$ *be as in section* 2. *Then for all* $\eta \in E$

$$\int f_S(\xi)\mu^-_{\beta,H}(d\xi) \leqq \liminf_{t\to\infty} T_t f_S(\eta) \leqq \limsup_{t\to\infty} T_t f_S(\eta) \leqq \int f_S(\xi)\mu^+_{\beta,H}(d\xi).$$

Since $\beta$ and $H$ are fixed throughout we occasionally omit them from the notation.

In view of Corollary (2.21) and (2.24), Theorem (3.1) is an immediate consequence of the following lemma.

(3.2) LEMMA. *Under the conditions of Theorem* (3.1)

$$\int f_S(\xi)\mu_\Lambda^-(d\xi) \leqq \liminf_{t\to\infty} T_t f_S(\eta) \leqq \limsup_{t\to\infty} T_t f_S(\eta) \leqq \int f_S(\xi)\mu_\Lambda^+(d\xi) \tag{3.3}$$

*for all finite* $\Lambda$.

The proof of Lemma (3.2) involves the same ideas as the proof of Theorem (2.4). We want to couple the stochastic Ising model to another process which we have some control over. In order to do this we let $\Lambda$ be fixed and concentrate on the last inequality in (3.3). Let $T_t^{(+,\Lambda)}$ be the semigroup corresponding to the Markov process, $\eta_t^{(+,\Lambda)}$, which evolves in exactly the same way that the stochastic Ising model does inside of $\Lambda$ but has the spins outside of $\Lambda$ fixed at $+1$ for all times. In other words, $\eta_t^{(+,\Lambda)}$ is the Markov process whose state space is the subset

$$A(+,\Lambda)\ (=\{\eta\in E \mid \eta(x)=1 \text{ if } x\notin\Lambda\})$$

of $E$ and which satisfies (1.3) and (1.4) for all $x\in\Lambda$ and $\eta\in A(+,\Lambda)$; but for $x\notin\Lambda$ it satisfies $P_\eta(\eta_t^{(+,\Lambda)}(x)=1)=1$ for all $t$.

The infinitesimal generator, $\mathfrak{A}$, of $\eta_t^{(+,\Lambda)}$ is given as follows. If $\varphi,\psi\in A(+,\Lambda)$ and $\varphi\neq\psi$, then

$$\mathfrak{A}(\varphi,\psi)=\begin{cases} c(x,\varphi) \text{ if } \varphi(y)=\psi(y) \text{ for } y\neq x \\ \qquad \text{and } \varphi(x)=-\psi(x), \\ 0 \text{ otherwise.}\end{cases}$$

$\mathfrak{A}(\varphi,\varphi)$ is defined so that $\sum_{\psi\in A(+,\Lambda)}\mathfrak{A}(\varphi,\psi)=0$.

Using (1.2) and (2.15) it is easily checked that

$$(3.4)\qquad \sum_{\varphi\in A(+,\Lambda)}\mu_\Lambda^+(\varphi)\,\mathfrak{A}(\varphi,\psi)=0 \text{ for all } \psi\in A(+,\Lambda).$$

Thus $\mu_\Lambda^+$ is a stationary distribution for $\eta_t^{(+,\Lambda)}$, and since the Markov process $\eta_t^{(+,\Lambda)}$ can clearly get from any configuration in $A(+,\Lambda)$ to any other configuration in $A(+,\Lambda)$, $\mu_\Lambda^+$ is the unique stationary distribution for $\eta_t^{(+,\Lambda)}$.

The next step is to couple the two Markov processes $\eta_t^{(+,\Lambda)}$ and $\eta_t$ together in such a way that if only $\eta_t^{(+,\Lambda)}$ is watched, it does not appear to be coupled to anything, and similarly for $\eta_t$. However the coupling should be such that if $\eta_0^{(+,\Lambda)}(x)\geqq\eta_0(x)$ for all $x$, then for all $t>0$ we have $\eta_t^{(+,\Lambda)}(x)\geqq\eta_t(x)$ for all $x$ with probability one. Once this is done the last inequality in (3.3) follows immediately by applying the ergodic theorem to the finite state space process $\eta_t^{(+,\Lambda)}$.

The coupling is practically the same as in Theorem (2.4). Rather than attempt to describe it we content ourselves with pointing out the crucial fact that makes it work. The interested reader is referred to [10] for details.

Note that from the definition of $U(x,\eta)$ given in the introduction it follows that the more spins that are lined up with the spin at $x$ the smaller $U(x, \eta)$ is. Thus, because of (1.1), the more neighbors a spin is lined up with, the lower its flip rate is. This attraction of like spins for each other allows one to define a coupling which has the desired properties.

The following corollary of Theorem (3.1) is the ergodic theorem for the stochastic Ising model.

(3.5) COROLLARY. *If $\beta < \beta_c$ or $H \neq 0$, then for all $f \in \mathcal{C}(E)$ and all $\eta \in E$*

$$\lim_{t \to \infty} T_t f(\eta) = \int_E f(\xi) \mu_{\beta,H}(d\xi). \tag{3.6}$$

PROOF. As we have already noticed it suffices to prove (3.6) for $f$ of the form $f_S$. But then (3.6) is an immediate consequence of Theorem (3.1) and (2.33).

(3.7) THEOREM. *If $\beta > \beta_c$ and $H = 0$, then there exists $f \in \mathcal{C}(E)$ and $\eta \in E$ for which* $\lim_{t \to \infty} T_t f(\eta)$ *does not exist.*

If we recall (2.32), then Theorem (3.7) is hardly surprising; however, a proof seems to be more involved than we want to get into here. We leave it to the interested reader to give a proof by using Lemma (3.8) below together with the facts that $\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$ are both stationary distributions for $T_t$ and $T_t f_{\{0\}}(\eta)$ is continuous in $\eta$.

In the final section we want to give some physically interesting examples of what can happen when $\beta > \beta_c$ and $H = 0$. The next lemma is useful in this regard.

(3.8) LEMMA. *If $\eta$ and $\eta'$ are two initial configurations of spins such that $\eta(x) \geqq \eta'(x)$ for all $x$, then for all $S \subset \mathbf{Z}^3$ and all $t$, $T_t f_S(\eta) \geqq T_t f_S(\eta')$.*

The attraction of like spins for each other makes Lemma (3.8) seem very reasonable. A proof can be given by coupling two infinite systems together using the same type of coupling that is used in the proof of Lemma (3.2).

4. **The stochastic Ising model at low temperatures and zero external magnetic field.** We imagine a stochastic Ising model with parameters $\beta_0$ and $H_0$ which has been running for a long time and is in equilibrium. Suppose now that the parameters are suddenly changed to values $\beta$ and $H$ for which the stochastic Ising model does not have a unique equilibrium state. The problem we are interested in is to which, if any, of the possible equilibrium states does the process then converge. We look at two cases. One case is when $\beta_0 > \beta_c$ and $H_0 >$

0. $\beta$ is then left unchanged at $\beta_0$ but $H$ is changed to zero. We will see that in this case the distribution at time $t$ converges to $\mu^+_{\beta,0}$.

To understand what this means physically note that $\int \eta(0)\mu^+_{\beta,0}(d\eta)$ may be thought of as the average magnetization at the site 0 in state $\mu^+_{\beta,0}$. By (2.27) this is the same as the average magnetization at any other site in state $\mu^+_{\beta,0}$. Thus $\langle f_{\{0\}}, \mu^+_{\beta,0}\rangle - [1 - \langle f_{\{0\}}, \mu^+_{\beta,0}\rangle] = \int \eta(0)\mu^+_{\beta,0}(d\eta)$ is the average magnetization per site of the state $\mu^+_{\beta,0}$. But from (2.32) we know that $\langle f_{\{0\}}, \mu^+_{\beta,0}\rangle > 1/2$. Thus the average magnetization per site of the equilibrium state $\mu^+_{\beta,0}$ is strictly positive. Hence if one puts the stochastic Ising model at low temperatures in a positive external magnetic field, lets it come to equilibrium, and then suddenly turns off the external field, the result is that when it comes to equilibrium again it will be magnetized in the up direction.

The second case is when $\beta_0 < \beta_c$ and $H_0 = 0$ are changed to $\beta > \beta'$ ($\beta'$ is as in (2.34)) and $H = 0$. This corresponds to suddenly lowering the temperature from a temperature above the critical temperature to one sufficiently below the critical temperature. We prove that in this case the distribution at time $t$ converges to $(1/2)\mu^+_{\beta,0} + (1/2)\mu^-_{\beta,0}$. However, more than this is true. We show that as time goes on the Markov process $\eta_t$ begins to form simultaneously regions where most of the spins are up and regions where most of the spins are down. These regions continue growing in size and for large times roughly half of the sites are in regions of up spins and half are in regions of down spins. The shape and location of these regions is of course random. It is hard to see what the physical interpretation of this should be in the spin language; however, if we switch to lattice gas language and identify a $+1$ with a particle and a $-1$ with empty space, then it is clear that what we have here is condensation as the temperature is lowered below the critical temperature.

Before beginning the proofs of these facts we need to introduce one more bit of notation. Let ${}_{\beta,H}T_t$ be the semigroup of operators corresponding to the stochastic Ising model with parameters $\beta$ and $H$. Then ${}_{\beta,H}T_t^*$ will be the dual semigroup of operators on measures given by $\langle f, {}_{\beta,H}T_t^*\mu\rangle = \langle {}_{\beta,H}T_t f, \mu\rangle$.

(4.1) THEOREM. *If $H > 0$, then ${}_{\beta,0}T_t^*\mu_{\beta,H}$ converges weakly to $\mu^+_{\beta,0}$ as t goes to infinity.*

PROOF. The semigroup ${}_{\beta,0}T_t$ will be fixed during the proof and we drop the $(\beta, 0)$ from the notation.

As before it suffices to show that for all finite $S \subset \mathbf{Z}^3$

$$\lim_{t\to\infty} \langle T_t f_S, \mu_{\beta,H}\rangle = \langle f_S, \mu^+_{\beta,0}\rangle.$$

But $\langle T_t f_S, \mu_{\beta,H}\rangle = \int T_t f_S(\eta)\mu_{\beta,H}(d\eta)$, and $T_t f_S$ is bounded; therefore, by Theorem (3.1) $\lim\sup_{t\to\infty}\langle T_t f_S, \mu_{\beta,H}\rangle \leqq \langle f_S, \mu^+_{\beta,0}\rangle$. Hence we need only show that $\langle T_t f_S, \mu_{\beta,H}\rangle \geqq \langle f_S, \mu^+_{\beta,0}\rangle$ for all $t$. Since $\mu^+_{\beta,0}$ is an equilibrium state for the semigroup $T_t$, this is the same as showing that $\langle T_t f_S, \mu_{\beta,H}\rangle \geqq \langle T_t f_S, \mu^+_{\beta,0}\rangle$. Now

$$\begin{aligned}(4.2)\quad &\langle T_t f_S, \mu_{\beta H}\rangle - \langle T_t f_S, \mu^+_{\beta,0}\rangle = \\ &= \lim_{\Lambda\uparrow Z^3}\left[\int T_t f_S(\eta)\mu^+_{\Lambda,\beta,H}(d\eta) - \int T_t f_S(\xi)\mu^+_{\Lambda,\beta,0}(d\xi)\right].\end{aligned}$$

Recall that $\mu^+_{\Lambda,\beta,H}$ and $\mu^+_{\Lambda,\beta,0}$ are measures concentrated on $A(+,\Lambda) \subset E$ and that $\mu^+_{\Lambda,\beta,H} >^s \mu^+_{\Lambda,\beta,0}$ (see the proof of Corollary (2.31)). Thus there is a measure $\mu_\Lambda$ on $A(+,\Lambda) \times A(+,\Lambda)$ which puts all of its mass on the set

$$B(\Lambda) = \{(\eta, \xi) \in A(+,\Lambda) \times A(+,\Lambda) \mid \eta(x) \geqq \xi(x) \text{ for all } x\}$$

and has $\mu^+_{\Lambda,\beta,H}$ as its first marginal and $\mu^+_{\Lambda,\beta,0}$ as its second marginal. Using this and (4.2) we get

$$\langle T_t f_S, \mu_{\beta,H}\rangle - \langle T_t f_S, \mu^+_{\beta,0}\rangle = \lim_{\Lambda\uparrow Z^3}\int [T_t f_S(\eta) - T_t f_S(\xi)]\,\mu_\Lambda(d\eta, d\xi).$$

Since $\mu_\Lambda$ is concentrated on $B(\Lambda)$, Lemma (3.8) implies that

$$\int [T_t f_S(\eta) - T_t f_S(\xi)]\,\mu_\Lambda(d\eta, d\xi) \geqq 0$$

for all $\Lambda$ and all $t$.

Thus $\langle T_t f_S, \mu_{\beta,H}\rangle \geqq \langle T_t f_S, \mu^+_{\beta,0}\rangle$ for all $t$, and the proof is complete.

We now turn to the second case.

(4.3) Theorem. *If* $\tilde\beta < \beta_c$ *and* $\beta > \beta'$ *then* ${}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}$ *converges weakly to* $(1/2)\mu^+_{\beta,0} + (1/2)\mu^-_{\beta,0}$ *as t goes to infinity.*

Proof. According to (2.27) and (2.33) $\mu_{\tilde\beta,0}$ is shift invariant. Now Corollary (3.17) in [11] says that if $\mu$ is shift invariant, then ${}_{\beta,0}T_t{}^*\mu$ converges weakly to the set of shift invariant Gibbs states with parameters $\beta$ and 0. Since $\beta > \beta'$, (2.34) implies that the shift invariant Gibbs states with parameters $\beta$ and 0 are all of the form $\alpha\mu^+_{\beta,0} + (1-\alpha)\mu^-_{\beta,0}$ for some $0 \leqq \alpha \leqq 1$.

Now when $H = 0$, the symmetry in (2.15) together with the equalities $\mu_{\tilde\beta,0} = \mu^+_{\tilde\beta,0} = \mu^-_{\tilde\beta,0}$ imply that $\langle f_{\{0\}}, \mu_{\tilde\beta,0}\rangle = \langle(1 - f_{\{0\}}), \mu_{\tilde\beta,0}\rangle$. Hence

$$(4.4)\qquad \langle f_{\{0\}}, {}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}\rangle = \langle(1 - f_{\{0\}}), {}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}\rangle$$

for all $t$. The last statement is intuitively plausible due to the symmetry in the generator of the semigroup ${}_{\beta,0}T_t$; however, we omit the proof. A proof can be given by making use of Theorem (2.6) in [8].

Now from the definitions of $\mu^+$ and $\mu^-$, (2.15) and (2.32), it is easily seen that $\langle f_{\{0\}}, \mu^+_{\tilde\beta,0}\rangle = \langle (1 - f_{\{0\}}), \mu^-_{\tilde\beta,0}\rangle \neq \langle (1 - f_{\{0\}}), \mu^+_{\tilde\beta,0}\rangle$. Thus the only $\alpha$ for which

$$\langle f_{\{0\}}, \alpha\mu^+_{\tilde\beta,0} + (1-\alpha)\mu^-_{\tilde\beta,0}\rangle = \langle (1 - f_{\{0\}}), \alpha\mu^+_{\tilde\beta,0} + (1-\alpha)\mu^-_{\tilde\beta,0}\rangle$$

is $\alpha = 1/2$.

From this observation, (4.4), and the convergence of ${}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}$ to the set of shift invariant Gibbs states it follows that ${}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}$ converges weakly to $(1/2)\mu^+_{\tilde\beta,0} + (1/2)\mu^-_{\tilde\beta,0}$.

The proof of the next lemma can be found in [8].

(4.5) Lemma. *If $\mu$ is a probability measure on the Borel sets of $E$ which is ergodic for shifts in space, then ${}_{\beta,0}T_t{}^*\mu$ is also ergodic for shifts in space.*

Since $\tilde\beta < \beta_c$, Lemma (4.5) and Corollary (2.28) imply that ${}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}$ is ergodic for all $t$. We also know from Theorem (4.3) that ${}_{\beta,0}T_t{}^*\mu_{\tilde\beta,0}$ converges weakly to $(1/2)\mu^+_{\tilde\beta,0} + (1/2)\mu^-_{\tilde\beta,0}$, and from (2.32) we know that $\mu^+_{\tilde\beta,0}$ is magnetized up while $\mu^-_{\tilde\beta,0}$ is magnetized down. We leave it to the reader to use these facts and convince himself of the truth of the statements which we made at the beginning of this section concerning the formation of regions of mostly up spins and simultaneously the formation of regions of mostly down spins.

## References


1. G. Birkhoff, *Lattice Theory*, Amer. Math. Soc., Providence (1967).

2. R. L. Dobrushin, *Gibbsian random fields for lattice systems with pairwise interactions*, Functional Analysis and its Applications 2 (1968), 292–302.

3. ——, *The problem of uniqueness of a Gibbsian random field and the problem of phase transitions*, Functional Analysis and its Applications 2 (1968), 302–312.

4. B. U. Felderhof, *Spin relaxation of the Ising chain*, Rep. Math. Phys. 1 (1970), 215–234.

5. C. M. Fortiun, P. W. Kasteleyn, and J. Ginibre, *Correlation inequalities on some partially ordered sets*, Commun. Math. Phys. 22 (1971), 89–103.

6. G. Gallavotti and S. Miracle-Sale, *Equilibrium states of the Ising model in the two phase region*, to appear.

7. R. Glauber, *Time dependent statistics of the Ising model*, J. Math. Phys. 4 (1963), 294–307.

8. R. Holley, *Markovian interaction processes with finite range interactions*, to appear in Annals of Math. Stat.

9. ——, *A class of interactions in an infinite particle system*, Adv. in Math. **5** (1970), 291-309.

10. ——, *An ergodic theorem for interacting systems with attractive interactions*, to appear in Z. Wahrscheinlichkeitstheorie verw. Geb.

11. ——, *Free energy in a Markovian model of a lattice spin system*, Commun. Math. Phys. **23** (1971), 87-99.

12. J. Liebowitz and A. Martin Löf, *On the uniqueness of the equilibrium state for Ising spin systems*, Commun. Math. Phys. **25** (1972), 276-282.

13. T. M. Liggett, *Existence theorems for infinite particle systems*, Trans. Amer. Math. Soc. **165** (1972), 471-481.

14. D. Ruelle, *Statistical Mechanics*, Benjamin, New York (1969).

15. ——, *On the use of "small external fields" in the problem of symmetry breakdown in statistical mechanics*, to appear in Commun. Math. Phys.

16. F. Spitzer, *Interaction of Markov processes*, Adv. in Math. **5** (1970), 246-290.

PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08540